Fast-Tracked to Success: Evidence on the Returns to Vocational Education in Switzerland

Fernando Saltiel*

McGill University

April 3, 2021

Abstract

I estimate the returns to vocational education in Switzerland, which has the highest share of secondary students in firm-based vocational education across developed countries. I present a discrete choice model of secondary tracking decisions and tertiary education. I rely on longitudinal data encompassing students' age fifteen test scores, non-cognitive skills, educational progression and age-30 labor market outcomes. The model considers specific upper-secondary tracking choices along with tertiary degree completion. Students sort into academic tracks based on their cognitive and non-cognitive ability. There are negative returns to academic studies relative to vocational tracks, yet these vary by the specific vocational track under consideration. There are positive returns to higher education for students in all upper-secondary tracks, with larger returns for those who pursued vocational tracks. I find a negative continuation value of academic studies relative to vocational education along with evidence of dynamic substitutability, with varying significance across specific vocational tracks. The strong linkage between vocational tracks and tertiary schooling drives positive outcomes for students in these tracks.

Keywords: Vocational Education, Returns to Skills, Unobserved Heterogeneity.

*Email: fernando.saltiel@mcgill.ca. Department of Economics, McGill University, Leacock Building 426, 855 Sherbrooke St. West, Montreal, QC H3A 2T7. I would like to thank the editor Arnaud Maurel and two anonymous referees for many helpful comments and suggestions. I would also like to thank participants at the Southeastern Micro Labor Workshop and in particular my discussant Perihan Saygin. This paper also benefited from conversations with Jorge Rodriguez. All errors are my own.

1 Introduction

Human capital investments in secondary school directly shape students' subsequent educational attainment and labor market outcomes (Altonji et al., 2012). This fact is particularly pronounced in countries with academic and vocational tracks in high school, as the former provide students with general skills while the latter focus on occupation-specific training. As a result, the existing literature has generally posited a trade-off across these options: vocational tracks may facilitate labor market entry while worsening long-term labor market outcomes. A growing literature has thus considered the employment and wage effects of pursuing academic and vocational tracks (Hanushek et al., 2017; Hampf and Woessmann, 2017), though there is scant evidence on the returns to specific tracks *within* vocational education. Furthermore, to the best of my knowledge, the existing literature has not yet examined the extent to which the returns to specific vocational tracks are shaped through the higher education system.

In this paper, I estimate the returns to secondary school tracking choices in Switzerland, which has the highest share of firm-based vocational education at the upper-secondary level in the world (Hoeckel and Schwartz, 2010). I take advantage of longitudinal data linking fifteen year old PISA 2000 exam takers to their secondary and tertiary educational attainment along with their age 30 labor market outcomes. This data includes detailed information on students' background characteristics, their performance in the PISA assessment along with a battery of non-cognitive skill measures. I first classify students by whether they completed an academic or vocational upper-secondary degree, but also examine heterogeneity in vocational education, by separately considering students who pursued trades- or service-focused occupations.[1] Raw data comparisons show that students in academic tracks have higher baseline test scores and an increased likelihood of tertiary degree completion relative to their vocational-track peers, yet differences in hourly wages are muted across tracks.

[1]Given my interest in estimating the returns to specific vocational tracks, I focus the analysis on males, as only 6.7% of women pursue trades-based vocational studies. Bertrand et al. (2019) document similar differences in vocational track choices by gender in Norway.

To recover the returns to vocational education and understand the (potential) dynamic complementarity in human capital accumulation in this context, I estimate a discrete choice model of educational attainment, fitting in with recent work by Heckman et al. (2018), Rodríguez et al. (2018) and Humphries et al. (2018). In this model, students first decide whether to pursue an academic or vocational upper-secondary degree. Vocational track students further select training in a trades- or services-based occupation. Upon completing secondary schooling, students in all three tracks may complete a higher education degree. I examine their labor market outcomes using hourly wage outcomes at age 30. In the model, a vector of low-dimensional latent abilities, unobserved to the econometrician, affect both decisions and outcomes (Carneiro et al., 2003; Hansen et al., 2004; Heckman et al., 2006). The model additionally incorporates exclusion restrictions at each decision node. Following the literature on latent factor models, I identify the distribution of latent cognitive and non-cognitive ability using a measurement system of age fifteen test score and non-cognitive skill measures (Williams, 2019). The structure of the model allows me to recover the returns to various human capital investment decisions across the ability distribution.

I document sorting into upper-secondary tracks across both latent factors, as academic students outpace their counterparts in vocational tracks by 0.42 and 0.34 standard deviations (σ) in cognitive and non-cognitive ability, respectively. Important differences emerge *within* vocational tracks, as well: students in services-focused occupations have higher non-cognitive skills vis-a-vis their peers in service-based tracks. Using the estimated model parameters, I estimate the returns to academic studies relative to vocational tracks. The average treatment effect indicates a negative return to academic studies, in the range of -2.2%, exhibiting limited heterogeneity across the latent skill distribution. On the other hand, I find heterogeneous returns to academic studies depending on the alternative vocational track in consideration, as the returns relative to tracks in trades are larger than -3%, while remaining in the range of -1% for service-based tracks.

I further analyze the extent to which higher education drives the estimated returns to upper-secondary tracks. I find positive returns to tertiary education degrees for students in both academic and vocational tracks, exhibiting larger returns for higher-cognitive-ability students. Moreover, the estimated returns to tertiary education are largest for students in trades-based secondary tracks, in excess of 20%. To understand the drivers of the wage returns to academic tracks, I decompose the estimated impacts into its direct effects — assuming no tertiary degree completion — and the continuation value, which captures the extent to which the returns to academic tracks are driven by outcomes in higher education. I find a small negative continuation value of upper-secondary academic studies relative to vocational tracks, in the range of -0.5%. The magnitude and sign of the continuation value parameter further varies across the vocational track under consideration. As such, these findings differ from the existing evidence showing the returns to schooling in the United States are largely driven through positive continuation values (Heckman et al., 2018).

To further understand the drivers of these findings, I follow Rodríguez et al. (2018) and relate the continuation value to dynamic complementarities in secondary and tertiary schooling, which capture whether academic tracking increases the returns to tertiary degree completion irrespective of whether students pursue higher education. The returns to upper-secondary academic studies exhibit dynamic substitutability, since the relative returns to tertiary education are larger for students in vocational tracks. These findings fit in with Rodríguez et al. (2018)'s results in the context of job training participation in Chile. As such, educational investment decisions in Switzerland may more closely track the structure of the returns to training than to schooling choices in the U.S. Lastly, counterfactual simulations demonstrate the critical role played by the strong linkage between vocational tracks and tertiary education in driving the 'positive' outcomes associated with vocational education.

This paper makes various contributions to the literature on vocational education, and more broadly to the literature on the returns to secondary school investments. To the best of my knowledge, this is the first paper to examine how the returns to specific vocational tracks are mediated by subsequent human capital accumulation. Humphries et al. (2018) estimate a Roy model of secondary and higher education choices in Sweden. I extend their work by incorporating different tracking choices *within* vocational education and documenting the extent of dynamic complementarities in such investments. This paper further fits in with recent work by Eckardt (2019) who documents heterogeneous returns to apprenticeships in Germany, finding larger returns for individuals working in occupations related to their apprenticeships. As a result, this paper contributes to an extensive literature on sequential human capital investment decisions, including Altonji (1995); Stange (2012); Stinebrickner and Stinebrickner (2012); Eisenhauer et al. (2015); Arcidiacono et al. (2016); Rodríguez et al. (2018); Heckman et al. (2018), along with previous work examining the linkage between secondary, higher education and labor market outcomes (Levine and Zimmerman, 1995; Rose and Betts, 2004; Joensen and Nielsen, 2009; Cortes et al., 2015; Goodman, 2019). Lastly, by directly modeling the returns to specific vocational education tracks, I contribute to a growing literature on the returns to vocational education (Oosterbeek and Webbink, 2007; Malamud and Pop-Eleches, 2010; Hall, 2012, 2016; Hanushek et al., 2017; Hampf and Woessmann, 2017; Brunello and Rocco, 2017; Golsteyn and Stenberg, 2017; Silliman and Virtanen, 2019; Bertrand et al., 2019; Korber, 2019; Neyt et al., 2020).

2 Background, Data Sources and Summary Statistics

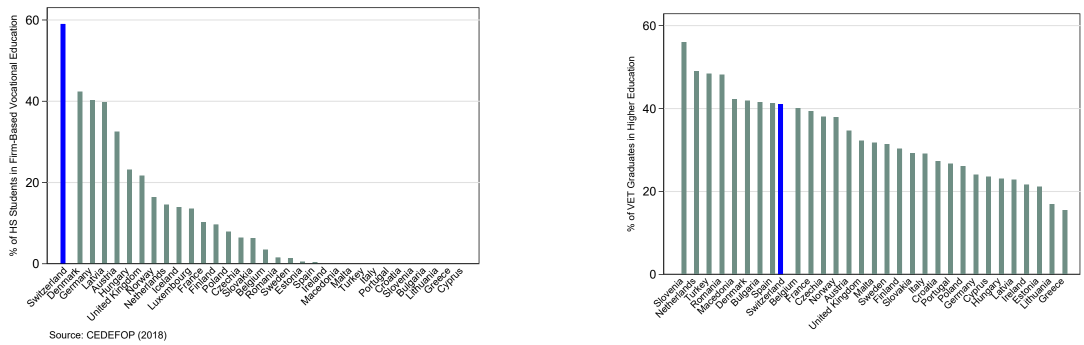
2.1 Institutional Background

Both the prevalence and the content of vocational education vary significantly across developed countries: vocational education can be generally classified by the relative importance of school-based learning vis-a-vis work-based training (Eichhorst et al., 2015). In school-based systems, vocational education is integrated into formal schooling as an alternative for academic education, largely amounting to a tracking system. School-based vocational education is prevalent in Belgium, Netherlands, Spain, the UK and Nordic countries. On the other hand, the dual-, or firm-based, vocational system combines class-based education with on-the-job training provided in private companies. Firm-based programs, which are prevalent in Austria, Denmark, Germany and Switzerland, exhibit a high degree of formalization, as detailed regulations often define the type and length of training covered in such tracks.

As shown in the first panel of Figure 1, Switzerland has the highest share of upper-secondary students enrolled in firm-based vocational education across European countries. The Swiss educational system is structured as follows. Students attend nine years of compulsory education, composed of six years of primary school along with three years of lower-secondary schooling. At the end of compulsory schooling, students may pursue upper-secondary studies, which are broadly classified into academic and vocational tracks. About one-third of students pursue academic studies, which includes baccalaureate schools, along with specialized schools in business and teaching. The remaining two thirds enroll in vocational tracks, and 90% of these students follow firm-based education. Training in upwards of 250 occupations is offered in this track, and one third of firms in Switzerland provide such training (Hoeckel et al., 2009).

Switzerland's tertiary education system is directly linked to its upper-secondary system. There are two distinctive levels of higher education. The first largely corresponds to vocational-based tertiary education, designed for individuals seeking to climb up the occupational ladder. As such, it encompasses upwards of fifty professional degree programs in engineering, health care, and design, among others, as well as professional examinations. These programs last at least two years and are directly integrated with the secondary-level vocational system, as employers and professional organizations define the content of degree programs. Graduation rates in this sector more than doubled the OECD average in 2009 (Mihály and Simon, 2013). The second sector encompasses cantonal universities, federal institutes of technology and universities of applied sciences, which are generally attended by students completing academic studies in secondary schooling. Vocational track students may enroll in this sector by completing a one-year preparatory course for the Federal Vocational Baccalaureate. All in all, vocational education plays an important role in Swiss higher education, as half of degrees at this level were granted in professional colleges in 2009 (Hoeckel et al., 2009). As a result, Switzerland exhibits high levels of tertiary education completion for vocational track students within Europe, as well (Figure 1, Panel B).

Figure 1: Vocational Education in Switzerland



(a) % in Firm-Based Vocational Education **(b)** % of VET Graduates in Higher Education

Source: CEDEFOP (2018). The first panel of Figure 1 shows the share of secondary school students enrolled in firm-based vocational education across different European countries. The second panel shows the share of vocational education graduates who complete a tertiary education degree in these countries.

2.2 Data Sources

In this paper, I use data from the Transitions from Education to Employment (TREE) survey, a longitudinal study of students who had completed lower secondary schooling in 2000. TREE used the PISA 2000 exam as the baseline survey round, and has followed this cohort of students through 2015, when the majority had turned 30.[2] The initial survey round included 6,343 students who provided detailed information on their background characteristics, such as their age, gender and nationality, family composition and parents' educational attainment, as well as geographic information on the canton of residence and whether their region is German, French or Italian. Moreover, TREE data includes information on the lower secondary track they had completed.[3] The 2000 round of the PISA exam focused on assessing students'

[2]This cohort of students was surveyed annually in 2001-2007, capturing detailed information on their upper secondary educational attainment. Two additional rounds have been completed in 2010 and 2015.

[3]Lower secondary education options include pre-gymnasial, or following tracks with basic or extended academic requirements. Tracking is more prevalent in the German parts of the country, yet students across all lower secondary tracks enroll in both academic and vocational upper secondary schooling.

reading performance as the main subject domain, while also assessing mathematics and science as minor domains.[4] Moreover, PISA included information on students' non-cognitive skills across multiple domains, including measures of their perseverance, memorization, control strategies, self-efficacy, control expectation and elaboration strategies.[5] As such, I can examine the extent to which non-cognitive skills drive sorting into secondary school tracks.

In the follow-up survey rounds, TREE collected detailed information on respondents' educational attainment. I use data from the 2010 survey to first classify students by whether they completed an academic or vocational upper-secondary degree. For vocational graduates, TREE includes information on their occupational track, yet to reduce the dimensionality of the model, I follow Bertrand et al. (2019) and classify students by whether they completed a trades-based or a services-based track.[6] I additionally consider whether students had completed a tertiary education degree by age 30.[7] Students are thus classified into one of six potential outcomes, encompassing secondary schooling (academic and two vocational tracks) and higher education completion. I recover information on labor market outcomes from the last survey round, as respondents had turned 30 and largely exited out of formal schooling. TREE data provides information on whether respondents had been employed in the past year, their annual earnings and hourly wages in (up to) seven jobs. I focus on respondents' highest paid job and examine the returns to educational choices on hourly wages.

The sample used in the paper is constructed as follows. 6,176 of the 6,343 students included in the baseline round provided information on their background characteristics and participated in PISA. 3,142 respondents were successfully followed through the endline survey round (Gomensoro and Meyer, 2017), of whom 2,876 had completed an upper-secondary

[4]The implementation of the exam in Switzerland was such that one-fourth of the test-taking sample was randomly assigned to complete assessments in reading, math and science (Ray et al., 2003). Table A.1 shows the likelihood of taking all three assessments is uncorrelated with students' baseline characteristics, lending credence to the missing-at-random assumption considered in the empirical model presented in Section 3.

[5]Table A.2 presents an inventory of the questions used to construct the six non-cognitive skill measures used in the paper, which fit in with the existing literature on non-cognitive skills (Kautz et al., 2014).

[6]Trades include students in industrial, electrical, construction or agricultural occupations. Service-based tracks encompass students in sales, services, business, arts, design, health and social care.

[7]TREE includes information on the type of tertiary degree completed, yet given the small sample size available in the data, I focus the analysis on a binary higher education completion outcome.

degree at age 30 while providing valid information on their educational attainment and labor market outcomes. I restrict the sample to males, due to the small share of women pursuing vocational tracks in trades, yielding a sample of 1,212 males. The final sample includes 1,001 males who were employed at age 30.

2.3 Reduced Form Evidence

Table 1: Summary Statistics by HS Track

	Full Sample (1)	Academic (2)	Trades (3)	Services (4)
Background Characteristics				
Age	15.517	15.372	15.597***	15.676***
Swiss	0.918	0.921	0.922	0.903
Pre-Gymnasial Track	0.441	0.715	0.225***	0.286***
German Region	0.489	0.385	0.544***	0.606***
French Region	0.390	0.464	0.346***	0.314***
Both Parents	0.845	0.837	0.858	0.834
Parents' Ed: Tertiary	0.430	0.555	0.331***	0.360***
Test Scores				
PISA Reading Score	-0.144	0.296	-0.485***	-0.399***
Control Strategies	-0.074	0.199	-0.340***	-0.107**
Perseverance	-0.040	0.116	-0.214***	-0.006
Memorization	-0.056	0.128	-0.213***	-0.129**
Self-Efficacy	0.180	0.421	-0.013***	0.055***
Control Expectation	0.057	0.294	-0.144***	-0.041***
Elaboration Strategies	0.078	0.269	-0.063***	-0.049***
Outcomes				
College Graduate	0.627	0.801	0.483***	0.549***
Hourly Wage at Age 30	30.40	31.89	29.51*	28.88
Observations	1,001	418	408	175

Source: TREE (Transitions from Education to Employment). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Table 1 displays summary statistics by high school track, including vocational tracks in Trades and Services. The stars in the third and fourth columns capture the statistical significance of a t-test comparing the means of the variables of academic-track students to each of the vocational-track groups, respectively.

Descriptive Statistics. In Table 1, I present summary statistics for the sample used in the paper. 44% of the sample had followed a pre-gymnasial track in lower secondary education and 43% had parents with a completed tertiary education degree. By the last survey round, 63% of the sample had completed a higher education degree and earned an average hourly wage of 30.40 CHF. In columns (2)-(4), I compare students across their secondary school track choice. There are significant differences in background characteristics across tracks: students in academic tracks are more likely to have attended a pre-gymnasial track and have higher educated parents. There are sizable differences in PISA reading test scores, as well. Students in academic tracks outperformed their peers in trades and services-based tracks by

0.77 and 0.69 σ, respectively. Similar patterns emerge in the non-cognitive dimension, as respondents pursuing academic studies significantly outpace their peers in both vocational tracks across the six non-cognitive measures. These differences persist through age 30, as 80% of students in academic tracks completed a tertiary degree, compared to 48% and 55% of their peers in trades and services, respectively. On the other hand, differences in hourly wages across tracks are relatively small, with the only significant difference emerging between academic- and trades-based students.[8]

Regression Results. In Table A.4, I present evidence from an OLS regression to examine whether there are significant differences in hourly wages across educational paths conditional on students' academic preparation and background characteristics. The first column shows that students in the two vocational tracks do not earn lower wages than their peers in academic studies. In the last three columns, I document significant wage differences by higher education graduation status for students completing different tracks in upper-secondary school: the estimated wage return to higher education completion exceeds 20% among students who pursued trades-focused vocational tracks. These results do not constitute causal evidence of the returns to schooling choices in Switzerland, as an extensive literature has shown that test scores and non-cognitive skill constructs measure latent ability with significant error (Carneiro et al., 2003; Heckman et al., 2006).[9] Controlling for these variables may thus fail to correctly account for endogenous selection into educational choices. In fact, these results do not capture the importance of dynamic complementarities in secondary and tertiary schooling, and do not provide evidence on the extent of heterogeneous returns to educational choices across the skill distribution.

[8]Table A.3 examines outcomes across students by their tertiary education completion status. Across the three secondary tracks, higher education graduates had significantly higher reading test scores vis-á-vis their non-completer peers. Across all tracks, higher education graduates outearn their peers without a degree.

[9]Table A.5 shows the estimated returns to reading skills in the full sample are small in magnitude (2.5%), fitting in with Korber (2019). The returns to non-cognitive and reading skills for students across each upper-secondary track as small; the last column shows similar evidence for math skills. These results may differ from the existing literature (Heckman et al., 2006, 2018) due to cross-country differences in the returns to skills driven by heterogeneous labor market structures (Hanushek et al., 2015) and/or to the fact that observed test scores measure skills with error. Section 3 considers the importance of the latter issue.

3 Model of Educational Attainment

In this section, I introduce a discrete choice model to capture the dynamics of schooling choices in secondary and tertiary education in Switzerland and its impacts on labor market outcomes. In the model, a vector of low-dimensional latent abilities, unobserved to the econometrician, affect both educational decisions and associated outcomes. Given the available measures in the TREE data set, I consider two dimensions of ability, including cognitive and non-cognitive components, which are identified through a measurement system of age fifteen test scores. The model is thus in the vein of Carneiro et al. (2003), Heckman et al. (2006), and Heckman et al. (2018), allowing for essential heterogeneity as the returns to secondary and tertiary education may vary across the latent ability distribution.

3.1 Model Structure

Upper Secondary Education. Upon completing lower secondary education, Swiss students first decide whether to pursue upper secondary education in an academic or vocational track. The decision of student i to pursue the academic track ($D_i = 1$) or the vocational track ($D_i = 0$) is determined by:

$$D_i = \mathbb{1}\left[\beta^D X_i^D + \alpha^D \boldsymbol{\theta_i} + \varepsilon_i^D > 0\right] \tag{1}$$

where $\mathbb{1}$ is an indicator function, X_i^D includes characteristics measured at baseline which affect the tracking decision, $\boldsymbol{\theta_i}$ represents the vector of unobserved ability and ε_i^D is an error term which is independent of observed and unobserved characteristics ($\varepsilon_i^D \perp X_i^D, \boldsymbol{\theta_i}$). The set of variables in X_i^D includes family and student characteristics, including their age, gender and pre-gymnasial lower secondary track — which may affect the likelihood of pursuing an upper secondary academic education, as well as exclusions affecting the tracking decision.

Students who pursue a vocational education subsequently decide whether to train in a trades- or services-related occupation. Students' decisions to pursue training in trades

($v_{it} \equiv \{V_i = 1\}$) or services ($v_{is} \equiv \{V_i = 0\}$) is given by:

$$V_i = \mathbb{1}\left[\beta^V X_i^V + \alpha^V \boldsymbol{\theta_i} + \varepsilon_i^V > 0\right] \tag{2}$$

where X_i^V measures observed characteristics affecting the tracking decision and $\boldsymbol{\theta_i}$ contains the vector of latent ability. The error term ε_i^V is also uncorrelated with X_i^V and $\boldsymbol{\theta_i}$.

Tertiary Education. Upon completing secondary education, academic and vocational track students may complete a tertiary education degree. For student i in upper-secondary track k, her decision to complete a higher education degree ($T_i = 1$) is given by:

$$T_{i,k} = \mathbb{1}\left[\beta_k^T X_{i,k}^T + \alpha_k^T \boldsymbol{\theta_i} + \varepsilon_{i,k}^T > 0\right] \tag{3}$$

where $\varepsilon_{i,k}^T$ is independent of observed and unobserved characteristics ($\varepsilon_{i,k}^T \perp X_{i,k}^T, \boldsymbol{\theta_i}$). $X_{i,k}^T$ encompasses observed characteristics affecting tertiary education choices. All in all, the combination of educational choices, given by $[D_i, V_i, T_{i,k}]$, leads to a final level of attainment $l \in \mathcal{L}$ captured by the dummy variable $D_{i,l}$.

Outcome Variables. In this framework, potential wages ($Y_{i,l}$) vary across students' final educational attainment and latent skills as follows:

$$Y_{i,l} = \beta_l^Y X_{i,l}^Y + \alpha_l^Y \boldsymbol{\theta_i} + v_{i,l}^Y \tag{4}$$

where $X_{i,l}^Y$ includes the same observed characteristics previously included in the choice equations, as these variables may directly affect labor market outcomes. $v_{i,l}^Y$ captures an idiosyncratic shock to fourth round outcomes, which is independent of observed and unobserved characteristics ($v_{i,l}^Y \perp X_{i,l}^Y, \boldsymbol{\theta_i}$). Across equations (1)-(4), I also assume the error terms are independent across decisions and potential outcomes.[10] Following the Quandt (1958) switching regression framework, I can use wages across final educational outcomes (equation (4))

[10] Specifically, $v_{i,l}^Y \perp v_{i,l'}^Y (\forall\, l, l' \in \mathcal{L}) \perp \varepsilon_{i,k}^T \perp \varepsilon_i^V \perp \varepsilon_i^D$.

to recover potential outcomes across academic (d), trades-based (v_t) and services-based (v_s) upper-secondary tracks:

$$Y_{i,k} = Y_{i,k,T_1}T_{i,k} + Y_{i,k,T_0}(1 - T_{i,k}) \quad for \;\; k \in \{d, v_t, v_s\} \tag{5}$$

where Y_{i,k,T_0} represents wages for students in upper-secondary track k who do not complete a college degree. I can similarly define potential wages in the vocational track $Y_{i,v}$, to recover the returns to academic vis-a-vis vocational tracks.[11] The structure of the model implies that $\boldsymbol{\theta_i}$ drives the cross-correlations of educational choices and outcomes, implying that OLS will recover biased estimates of the returns to secondary and tertiary education. As a result, identifying the distribution of $\boldsymbol{\theta}$ is of paramount importance in this context.

3.2 Measurement System

The latent ability vector is unobserved to the econometrician, as there are no direct measures of ability available. Furthermore, observed test scores cannot be used as measures of ability, as they measure it with error. I follow an extensive literature (Carneiro et al., 2003; Hansen et al., 2004; Heckman et al., 2006), and allow for $\boldsymbol{\theta}$ to be proxied by multiple age 15 test scores. Given the available measures, I consider two components of latent ability $\boldsymbol{\theta}$, including cognitive (θ_C) and non-cognitive ability (θ_{NC}), and allow for these components to be correlated (Heckman et al., 2018). I posit the following linear model for PISA test scores C_i^j:

$$\boldsymbol{C}_i = \beta^C \boldsymbol{X}_i^M + \alpha^C \theta_{i,C} + \alpha^{NC} \theta_{i,NC} + \boldsymbol{e}_i^C \tag{6}$$

where $\boldsymbol{C}_i$ is the vector of test scores, $\boldsymbol{X}_i^M$ captures baseline control variables and $\boldsymbol{e}_i^C$ represents the error term, which is independent across test scores j, observed characteristics and

[11] In particular, $Y_{i,v} = Y_{i,v_t}V_i + Y_{i,v_s}(1 - V_i)$, where Y_{i,v_t} and Y_{i,v_s} denote potential wages for students in trades- and services-based vocational tracks, respectively.

latent ability. As such, the measurement errors are assumed to be classical in nature. Since the PISA test scores considered in the measurement system represent achievement, rather than intelligence tests, I allow for test scores to load on both latent cognitive and non-cognitive abilities. While the structure of the PISA assessment implied that just one-fourth of the sample took all three exams, Williams (2019) shows that the distribution of latent factors can be identified as long as the variance-covariance matrix of observed measures can be consistently estimated. The likelihood of taking all three exams is not correlated with students' characteristics (Table A.1), lending credence to the missing-at-random assumption required for such identification assumptions to hold in this context.

I additionally incorporate information on non-cognitive skills through the six available measures in the TREE data, which are observed for the full sample. I posit a dedicated measurement system of latent ability where non-cognitive measures (NC_i^n) linearly depend on observed characteristics and on θ_{NC} as follows:

$$\boldsymbol{NC}_i = \beta^{NC}\boldsymbol{X}_i^M + \gamma^{NC}\theta_{i,NC} + \boldsymbol{e}_i^{NC} \tag{7}$$

I assume the error terms are independent across test scores, non-cognitive skill measures, as well as from decisions and outcomes.[12]

3.3 Identification and Estimation

Model Identification. Carneiro et al. (2003), Hansen et al. (2004) and Heckman et al. (2016) and Heckman et al. (2018) present the formal argument for identification of a discrete choice model, akin to the one presented in this paper. The distribution of latent ability $\boldsymbol{\theta}$ is identified following the arguments discussed in Appendix B. Furthermore, Hansen et al. (2004); Heckman et al. (2016) show that in the absence exclusion restrictions, the joint

[12]For observed test scores $j, j', e_i^{C,j} \perp e_i^{C,j'}$, for non-cognitive skill measures $n, n', e_i^{NC,n} \perp e_i^{NC,n'}$ and across test score j and non-cognitive measure n, $e_i^{C,j} \perp e_i^{NC,n}$. For test score j and non-cognitive measure n, $e_i^{C,j} \perp e_i^{NC,n} \perp v_{i,l}^{Y} \perp \varepsilon_{i,k}^{T} \perp \varepsilon_i^{V} \perp \varepsilon_i^{D}$.

distribution of choices and potential outcomes can be non-parametrically identified as long as the support on the covariates in the choice equations (for instance, $\beta^D X_i^D$ in equation (1) in the paper) matches the support of the corresponding error terms ($\eta^D = \alpha^D \boldsymbol{\theta_i} + \varepsilon_i^D$). In this framework, a conditional independence assumption — implying that educational choices and outcome variables are conditionally independent of observed characteristics and latent ability components (a 'matching-on-unobservables assumption') — secures model identification. At the same time, Heckman and Navarro (2007); Heckman et al. (2016) alternatively show that these models can be identified with sufficient variation across the choice equation, where the variation could arise from exclusion restrictions or functional form assumptions, without relying on the factor structure. The implementation of the discrete choice model outlined above includes a factor structure while also incorporating exclusion restrictions in each choice equation. I assess the empirical importance of exclusion restrictions in driving the estimated results by alternatively estimating the model without instruments.

Model Implementation. Table A.6 shows the variables used in the implementation of the model. In the measurement system, educational decisions and wage equations, I include students' observable characteristics. Moreover, I incorporate choice-specific covariates in each decision node. I take advantage of information from the 1970, 1980, 1990 and 2000 Swiss Federal Population Censuses to calculate canton-level variables. I follow Heckman et al. (2018) in controlling for long-run local economic conditions along with temporal deviations which are relevant when students make their educational decisions, thus acting as exclusion restrictions. In equation (1), I control for long-run (averaged across the 1970-1990 Censuses) and contemporaneous (2000) local unemployment rates for academic-track graduates who did not complete a tertiary degree as well as for the share of prime-age adults who completed an academic-based upper-secondary degree. In equation (2), I include the long-run and current share of prime-age vocational-track graduates who work in service-based occupations. Lastly, the higher education choice equations include relevant unemployment rates for students who pursued different upper-secondary tracks.[13]

I assume that the vector of unobserved ability is a random variable with mean zero. While the distribution of $\boldsymbol{\theta}$ is identified non-parametrically (Freyberger, 2018), I estimate the density of each unobserved ability component by using a mixture of two normal distributions for computational convenience. I thus approximate the distribution of each ability component k using a mixture of two normal distributions with means $(\mu_{1,k}, \mu_{2,k})$, probabilities $(p_{1,k}, p_{2,k})$, with $p_{1,k} + p_{2,k} = 1$, and variances $((\sigma_{1,k})^2, (\sigma_{2,k})^2)$ as follows:

$$\theta_k \sim p_{1,k}\ N(\mu_{1,k}, (\sigma_{1,k})^2) + p_{2,k}\ N(\mu_{2,k}, (\sigma_{2,k})^2)$$

To define the sample likelihood, let T_{im} represent observed cognitive and non-cognitive measures $m \in \mathcal{M}$. Let Ψ be the vector of model parameters. I assume the error terms across educational decisions, outcome variables and in the measurement system follow normal distributions.[14] Given the independence across the error terms invoked above, the likelihood for a set of $\mathcal{I}$ individuals is given by:

$$\mathcal{L}(\Psi \mid \mathbf{X}, \mathbf{Y}) = \prod_{i \in \mathcal{I}} \left[\int_{\boldsymbol{\theta}} \prod_{m \in \mathcal{M}} f(T_{im} \mid X_i^M, \boldsymbol{\theta}) \prod_{l \in \mathcal{L}} \Big\{ Pr(D_{il} = 1 \mid \boldsymbol{X_i}, \boldsymbol{\theta}) f(Y_{il} \mid X_{il}^Y, \boldsymbol{\theta}) \Big\}^{D_{il}} dF_{\boldsymbol{\theta}}(.) \right]$$

I estimate the model using a Gibbs sampler as the Markov Chain Monte Carlo (MCMC) algorithm (Hansen et al., 2004; Heckman et al., 2006). I generate 500 draws from the estimated posterior distribution of the model parameters and note that inference follows standard Bayesian arguments, as the standard errors are calculated from the standard deviations across these 500 draws. I simulate 100 samples where each simulated sample draws from the posterior of the estimated model parameters, yielding 101,000 simulated observations.[15]

[13] The relevant unemployment rate in the tertiary education choice corresponds to canton-level Tertiary B unemployment rates for vocational track students and Tertiary A unemployment for academic-track students.

[14] Specifically $\varepsilon_i^D \sim N(0,1)$; $\varepsilon_i^V \sim N(0,1)$; $\varepsilon_{i,k}^T \sim N(0,1)\ \forall k \in \{d, v_t, v_s\}$; $v_{i,l}^Y \sim N(0, \sigma_{v,l}^2)$; $e_i^{C,j} \sim N(0, \sigma_{C^j}^2)$; $e_i^{NC,n} \sim N(0, \sigma_{NC^n}^2)$.

[15] The vector of initial parameters follows from the transition kernel, and the Markov Chain is thus generated according to the Gibbs sampler: as n $\rightarrow \infty$, the limiting distribution is the posterior. Once convergence is achieved, I make 500 draws from the posterior distribution of estimated model parameters to

4 Model Results

4.1 Model Fit

Goodness of Fit. I first examine the accuracy of the model in predicting educational attainment and outcome variables. In Figure A.1, I contrast students' observed and simulated education decisions. The first panel shows the model predicts both upper-secondary and vocational track choices accurately. The second panel shows predicted tertiary completion rates across upper-secondary tracks: 80.1% of academic track students complete a higher education degree, and the model predicts that 80% would do so, as well. Table A.7 contrasts individuals' observed and simulated hourly wages across secondary and tertiary educational choices. For students who complete a tertiary degree and for their peers who remain upper-secondary graduates, the model predicts hourly wages well, too. Lastly, Figure A.2 shows the model accurately matches predicted shares of academic tracks and hourly wages for students in different lower secondary tracks and for students whose parents completed (did not complete) a tertiary degree.

Measurement System. In Figure A.3, I present a variance decomposition of the measurement system.[16] Family and individual characteristics explain close to 20% of the variance of math, reading and science test scores. Latent cognitive ability explains a sizable share of the variance in PISA performance, ranging from 30% in the math assessment to 50% in the reading exam, whereas non-cognitive ability accounts for a small share of the variance in test score performance. On the other hand, θ_{NC} explains an important share across all six non-cognitive skill measures (30-70%), yet background characteristics account for a negligible share of the variance in these measures. Lastly, the correlation between cognitive and non-cognitive ability is small, equaling 0.10, yet not too dissimilar from the estimated correlation (0.21) across these two components in the NLSY79 (Prada and Urzua, 2017).

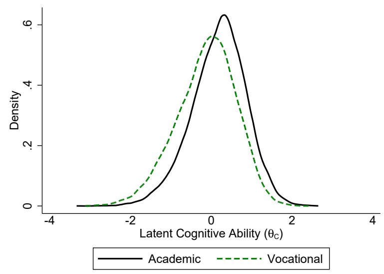
compute the mean and the standard errors of the parameters of interest (Hansen et al., 2004).

[16] Table A.8 presents the estimated loadings from the measurement system. In Figure A.4, I show the marginal distribution of each component of the latent ability distribution.

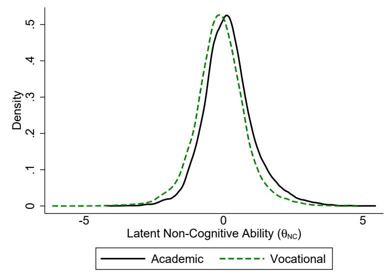
4.2 Determinants of Educational Choices

Figure 2: Density of Latent Ability by Educational Choices

Density of Latent Ability by Secondary School Choice

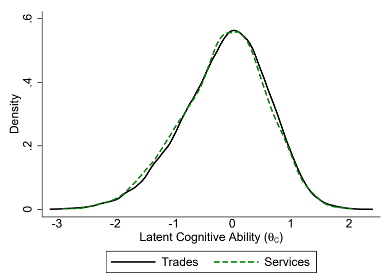


(a) Density of θ_C by Secondary Path

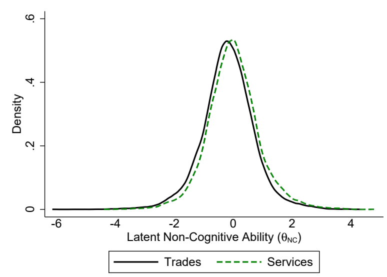


(b) Density of θ_{NC} by Secondary Path

Density of Latent Ability by Vocational Track



(c) Density of θ_C by Vocational Track



(d) Density of θ_{NC} by Vocational Track

Note: Figure 2 shows the marginal densities of the cognitive and non-cognitive latent factor across secondary school choices.

Using the estimated model parameters, I examine how students sort into vocational or academic tracks based on their latent ability. Figure 2 compares the estimated distributions of cognitive (Panel A) and non-cognitive (Panel B) ability across choices. I find significant sorting-into-academic tracks on both skill dimensions, as a one σ increase in θ_C increases the likelihood of completing an academic degree in upper-secondary schooling by 10.0 percentage points, whereas a corresponding increase in the θ_{NC} dimension yields a 8.4 percentage point increase. The last two panels indicate important differences in the sorting patterns within vocational tracks. While there are no sizable differences in terms of students' cognitive ability across trades- or service-based tracks, important differences emerge in the non-cognitive dimension, as a one standard deviation increase in θ_{NC} is associated with a 4.6 percentage point increased likelihood of pursuing training in service-based occupations.

In Figure A.5, I show the importance of cognitive ability as a driver of higher education completion for students in academic as well as in trades- and service-based tracks. On the other hand, sorting into higher education on θ_{NC} is muted for students in the different upper-secondary tracks. Given the extent of sorting on latent abilities, raw wage differences across educational paths may thus overstate the wage returns to studying these tracks.

5 Estimated Returns to Educational Choices

I first take advantage of the estimated model parameters to recover the estimated returns to pursuing different upper-secondary options. Letting E[.] denote the expected value taken with respect to the distribution of (X, $\boldsymbol{\theta}$, ε), the average treatment effect (ATE) of academic tracks (d) relative to vocational tracks (v) is given by:

$$ATE_{d,v} \equiv \int \int E[Y_{i,d} - Y_{i,v} | \boldsymbol{X} = x, \boldsymbol{\theta} = \underline{\boldsymbol{\theta}}] dF_{X,\boldsymbol{\theta}}(x, \underline{\boldsymbol{\theta}}) \tag{8}$$

where $Y_{i,d} - Y_{i,v}$ captures the wage return to pursuing an academic track relative to a vocational one in upper-secondary school. I further follow equation (8) to recover the returns to academic studies relative to specific vocational tracks.

5.1 Returns to Upper-Secondary Tracks

In the first column of Table 2, I present the estimated returns to academic studies vis-à-vis vocational tracks in upper-secondary schooling in Switzerland. While the raw wage gaps presented in Table 1 showed that students in academic tracks outearn their peers in vocational tracks by 3.6%, the estimated ATE of academic schooling is in fact negative, equaling -2.2%. As shown in Figure 2, higher-skilled students are far more likely to have completed an academic track. Since the discrete choice model accounts for differences in sorting-into-tracks based on observed characteristics and latent skills, the estimated average treatment effect of academic schooling becomes negative once such sorting is accounted for.[17]

Table 2: Estimated Returns to Vocational Tracks in Secondary School

Estimate	Vocational	Trades	Services
	(1)	(2)	(3)
ATE	-0.022	-0.034	-0.012
	(0.002)***	(0.002)***	(0.002)***
TT	-0.042	-0.065	-0.010
	(0.004)***	(0.003)***	(0.004)**
TUT	-0.008	-0.003	-0.019
	(0.003)***	(0.004)	(0.006)***

Note: Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table 2 presents the estimated age-30 wage returns to academic studies relative to vocational tracks (column 1) and relative to specific vocational tracks (in columns 2-3).

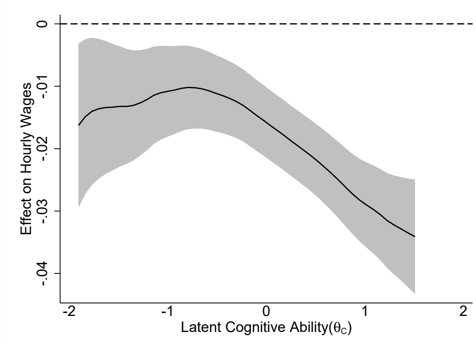
I further take advantage of the estimated model parameters to recover the returns to academic studies for both students who pursued such tracks (TT) and for those who did not (TUT). While the latent factor loadings in the wage equations are not substantially different across tracks (Table A.10), the estimated coefficients on observed characteristics which are more prevalent among academic-track students — such as having completed a pre-gymnasial lower-secondary track and having parents with a tertiary degree — are larger in the wage equations in vocational tracks.[18] As a result, the estimated return to studying an academic degree for students who completed such tracks (TT) would have been lower, equaling -4.2%, compared to the TUT returns indicating smaller (negative) returns for students who completed a vocational track, at -0.8%. While these results stand in contrast with an extensive literature finding larger returns to schooling investments for treated individuals relative to the average population (i.e. TT > ATE) (Carneiro et al., 2011; Heckman et al., 2018), Nybom (2017) and Rodríguez et al. (2018) have recently found limited heterogeneity

[17]The estimated parameters from the educational choice equations (Table A.9) show the instruments are insignificant across all choice equations, implying the model is identified through a conditional independence assumption. In Appendix C, I estimate an alternative version of the model without instruments and largely find similar estimated returns to upper-secondary and tertiary schooling choices in Switzerland.

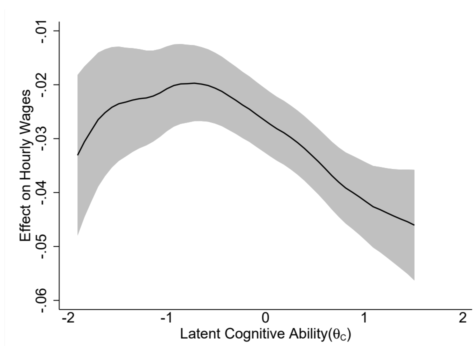
[18]The loadings on the latent factors in the wage equations are largely insignificant in both academic and vocational tracks, fitting in with the reduced form evidence presented in Table A.5. Rodríguez et al. (2018) similarly find small returns to latent abilities in the context of job training participation.

in the returns to higher education in Sweden and job training in Chile, respectively. Moreover, Mountjoy (2019) presents evidence consistent with reverse-Roy selection patterns in the returns to community college in the United States.

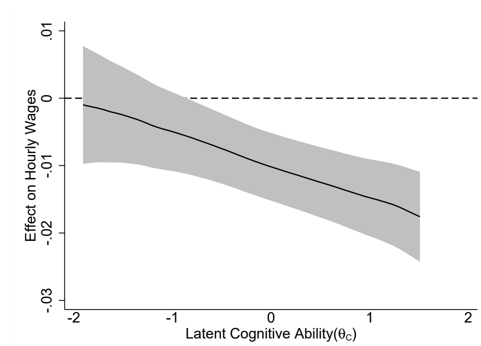
Figure 3: Heterogeneous Returns to Academic Studies Relative to Vocational Tracks



(a) Academic v. Vocational



(b) Academic v. Vocational-Trades



(c) Academic v. Vocational-Services

Note: Figure 3 presents the heterogeneous average treatment effect of academic studies relative to vocational tracks on hourly wages across the latent cognitive ability distribution.

I further analyze the extent of heterogeneous returns to upper-secondary tracks by estimating the returns to academic tracks across the latent ability distribution. I remark the importance of estimating such returns, in light of the posited impacts of vocational studies in helping less-skilled students to successful labor market transitions. In the first panel of Figure 3, I show that the estimated returns to academic studies are largely flat across the latent cognitive ability distribution. At the same time, the returns to academic tracks are slightly increasing across students' non-cognitive ability, as shown in the first panel of Figure A.6. These results thus highlight the importance of considering multiple dimensions of students' abilities when estimating the returns to diverse schooling investments.

While the existing literature on vocational education has largely focused on analyzing differences in employment levels across tracks (Hanushek et al., 2017; Hampf and Woessmann, 2017), a number of papers have recently examined the wage returns to tracking choices. Brunello and Rocco (2017) Golsteyn and Stenberg (2017) find negative wage returns to vocational education in the United Kingdom and Sweden, respectively. Humphries et al. (2018) extend this literature by presenting a Roy model of high school and college educational decisions. They document positive returns to academic and STEM-academic tracks vis-á-vis vocational tracks in Sweden, in the range of 5% and 7%, respectively. Korber (2019) takes advantage of TREE data and presents evidence from a matching estimator, finding negative wage returns to vocational education in Switzerland, in the range of 2.3-2.8%. Her results differ from those presented in this section, as she directly controls for tertiary education rather than modeling as an endogenous choice (equation (3)). As such, the positive returns to vocational education presented so far stand in contrast with the existing literature, remarking the importance of modeling sequential educational choices as well as estimating such returns in contexts with different educational systems.

Returns to Specific Upper-Secondary Tracks. The evidence presented in Figure 2 indicates that students in trades- and service-focused vocational tracks differ in their latent non-cognitive ability. Nonetheless, average wage gaps for students in academic studies relative to their peers in trades- and service-based occupations are largely similar, equaling 3.9% and 3.3% respectively. In the second and third columns of Table 2, I take advantage of the estimated model parameters and present the average returns to academic studies relative to each track. While across both vocational tracks, the estimated ATE is lower than the raw wage difference compared to academic tracks, important differences emerge in the estimated returns across such tracks. The returns to academic studies relative to trades-based vocational tracks are negative and statistically significant, exceeding -3%, whereas the corresponding returns relative to service-based tracks are in the range of -1%. In fact, for students who completed academic tracks in lieu of trades-based tracks, the estimated TT parameter denotes even larger (negative) returns, reaching -6.5%.[19] On the other hand, I find limited heterogeneity in the returns to academic studies relative to service-based tracks.[20]

In the last two panels of Figure 3, I further analyze heterogeneous returns to vocational tracks across the latent skill distribution. I find limited evidence of heterogeneous returns to academic studies vis-à-vis both vocational tracks across the cognitive ability distribution. At the same time, Figure A.6 documents increasing returns to academic studies relative to service-based tracks across the non-cognitive skill distribution. Altogether, these results indicate that academic tracks offer negative returns, yet the extent of these returns varies depending on the counterfactual vocational track under consideration, remarking the importance of incorporating specific tracking choices to the analysis. To the best of my knowledge, Eckardt (2019) is the only other paper to consider heterogeneous returns to vocational tracks. She examines returns to apprenticeship training in Germany, finding larger returns for individuals working in occupations related to their apprenticeships.

5.2 Returns to Higher Education

In Table 1, I had shown that students in academic tracks were more likely to have completed a higher education degree by age 30. Yet an important share of vocational track students had completed tertiary degrees, as well. While the raw wage differences across tertiary degree completion status shown in Table A.3 indicate that college completion may yield sizable wage returns, the extent of sorting-into-college on unobservables shown in Figure A.5 indicate these differences cannot be interpreted as causal. I thus use the estimated model parameters to recover the treatment effects associated with tertiary education for students

[19] The estimated TT parameter follows from a comparison of hourly wages for students who actually completed an academic track relative to their counterfactual wages in the corresponding vocational track.

[20] To understand the extent to which selection bias and sorting gains — which arise if students who stand to have the largest gains from pursuing specific tracks are more likely to have completed those tracks — account for the difference in observed wage gaps across tracks vis-a-vis the estimated ATE, I present the results from a decomposition introduced by Heckman et al. (2018) in Table A.11. Across the three estimated treatment effect parameters, selection bias accounts for these differences. In fact, since the estimated ATE is generally larger than the ATE of academic studies, I find evidence of 'negative' sorting bias into such tracks.

in different upper-secondary tracks k in:

$$ATE_{k,T}|\{H=k\} \equiv \int\int E[Y_{i,k,T_1} - Y_{i,k,T_0}|\boldsymbol{X}=x, \boldsymbol{\theta}=\underline{\boldsymbol{\theta}}, H=k]dF_{X,\boldsymbol{\theta}|H=k}(x,\underline{\boldsymbol{\theta}}) \; \forall \; k \in H \tag{9}$$

where $ATE_{k,T}|\{H=k\}$ denotes the returns to tertiary education for students who completed upper-secondary track $k \in H = \{d, v_t, v_s\}$. In Table 3, I present the returns to tertiary schooling by upper-secondary track in Switzerland. There are large returns to tertiary education for students in all tracks. For students who pursued academic studies, the returns to tertiary degree completion exceed 14%, with larger returns for students who in fact completed a tertiary degree. At the same time, while students in vocational tracks tend to complete higher education degrees in vocational programs, there are sizable returns to such degrees. Completing a higher education degree for students in trades-based training yields a wage return of 20%, with larger returns for those who actually completed a tertiary degree (TT = 23%). For their peers in service-based occupations, the estimated returns to higher education are lower, at 10.4%, yet statistically and economically significant.

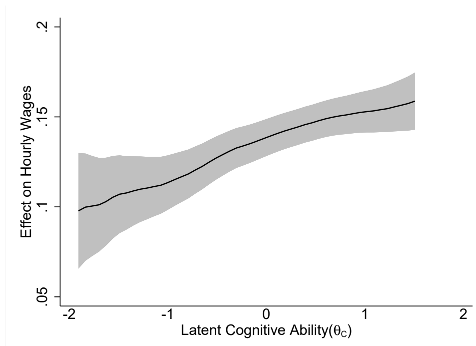
Table 3: Estimated Returns to Tertiary Schooling Across Secondary School Tracks

Estimate	Academic (1)	Trades (2)	Services (3)
ATE	0.142	0.200	0.104
	(0.004)***	(0.003)***	(0.005)***
TT	0.152	0.230	0.115
	(0.004)***	(0.005)***	(0.007)***
TUT	0.102	0.172	0.091
	(0.009)***	(0.004)***	(0.008)***

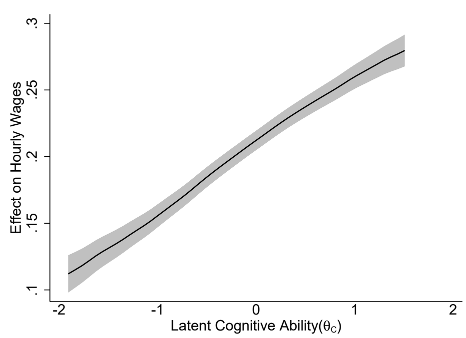
Note: Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table 3 presents the estimated age-30 wage returns to tertiary education across secondary schooling tracks.

In Figure 4, I examine whether the returns to tertiary education are heterogeneous across the cognitive ability distribution. In the first panel, I present evidence for students who completed upper-secondary academic studies. I find larger returns to higher education for students with high θ_C ability, as the estimated ATE of tertiary education reaches 15% for those in the top quintile of the cognitive ability distribution. The second panel shows that the extent of heterogeneous returns is stronger for students who pursued a trades-based vocational track, and slightly weaker (yet positive) for their peers in service-focused tracks.[21] All in all, these results show that completing a tertiary degree would largely yield positive labor market outcomes for students across both academic and vocational tracks.

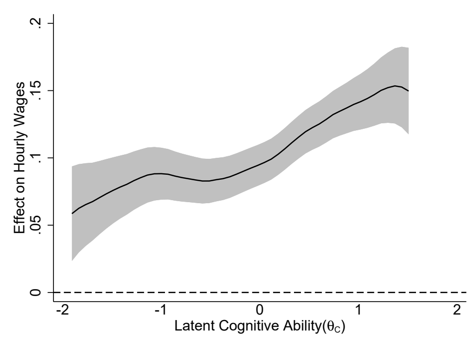
Figure 4: Heterogeneous Returns to Tertiary Education Across Secondary School Paths



(a) Academic Tracks



(b) Vocational-Trades Track



(c) Vocational-Services Track

Note: Figure 4 presents the heterogeneous average treatment effect to tertiary education across secondary school paths on hourly wages across the latent cognitive ability distribution.

While the returns to tertiary education are sizable for students across all tracks, the evidence presented in Table 1 indicated that students in academic studies were 29 percentage points more likely to have completed a tertiary degree by age 30 vis-a-vis their peers in vo-

[21]Figure A.7 presents corresponding evidence in the non-cognitive skill dimension. The extent of heterogeneous returns for students in academic and service-based tracks is largely muted, whereas the returns to higher education completion for trades-focused students are decreasing across the θ_{NC} distribution.

cational tracks. Using the estimated model parameters, I find that completing an academic track increases the likelihood of tertiary degree completion by 21.5 percentage points. The estimated return is consistent with a tracking story in which following an academic track in upper-secondary school leads to sizable increases in the likelihood of further schooling (Altonji et al., 2012). The positive wage returns to academic tracks coupled with the increased likelihood of tertiary degree completion for these students implies that higher education may explain a sizable share of the estimated returns.

5.3 Continuation Values and Dynamic Complementarity

Conceptual Framework. An extensive literature has examined the importance of option values in educational investments, which capture the extent to which the returns to schooling may be explained by the option of completing additional schooling (Stange, 2012; Stinebrickner and Stinebrickner, 2012; Eisenhauer et al., 2015; Arcidiacono et al., 2016). Since the model presented Section 3 does not directly specify students' preferences or information sets, I cannot recover the option values associated with high school track choices. Yet the model allows me to recover the continuation value of educational investments (Heckman et al., 2016), which captures the extent to which the returns to academic studies are explained by an increased likelihood of tertiary degree completion. Individual-level returns to academic tracks can be thus decomposed as follows:

$$Y_{i,d} - Y_{i,v} = \underbrace{Y_{i,d,T_0} - Y_{i,v,T_0}}_{\text{Direct Effect (DE)}} + \underbrace{T_{i,d}(Y_{i,d,T_1} - Y_{i,d,T_0}) - T_{i,v}(Y_{i,v,T_1} - Y_{i,v,T_0})}_{\text{Continuation Value (CV)}} \tag{10}$$

where $T_{i,k}$ indicates whether student i in track k completed a higher education degree (equation (3)). In equation (10), the direct effect captures the returns to academic studies relative to vocational tracks conditional on not pursuing higher education, while the continuation value depends on both the effect of an academic education on the likelihood of college completion, and on the additional wage return associated with college completion for academic graduates relative to their counterparts in vocational degrees. As argued by Rodríguez et al. (2018), the econometrician may be interested in understanding whether academic schooling leads to higher returns to college graduation, irrespective of its impacts on the likelihood of such completion. This parameter corresponds to the dynamic complementarities of educational decisions, which measure whether early investments increase the returns to subsequent investments (Cunha and Heckman, 2007; Cunha et al., 2010). To this end, the continuation value parameter can be further decomposed as follows:

$$\underbrace{T_{i,d}(Y_{i,d,T_1} - Y_{i,d,T_0}) - T_{i,v}(Y_{i,v,T_1} - Y_{i,v,T_0})}_{\text{Continuation Value (CV)}} = \underbrace{(Y_{i,d,T_1} - Y_{i,d,T_0}) - (Y_{i,v,T_1} - Y_{i,v,T_0})}_{\text{Dynamic Complementarity (DC)}}$$
$$+ \underbrace{(1 - T_{i,v})(Y_{i,v,T_1} - Y_{i,v,T_0}) - (1 - T_{i,d})(Y_{i,d,T_1} - Y_{i,d,T_0})}_{\text{Dynamic Sorting Gains}} \tag{11}$$

The continuation value of pursuing an academic track can be thus decomposed into the dynamic complementarity, measuring whether academic tracks increase the returns to tertiary education, along with a sorting term labeled "dynamic sorting gains." As such, when the continuation value of academic tracks exceeds the dynamic complementarity associated with such choices, students in academic tracks positively sort into tertiary education.

Empirical Evidence. In the first column of Table 4, I present evidence following equation (11). The estimated direct effect and continuation values of academic studies are both negative, albeit in different magnitudes: the direct effect equals -1.7%, whereas the negative continuation value associated with academic studies is fairly small, reaching -0.5%. These results stand in contrast with Heckman et al. (2018), who had found the returns to schooling in the U.S. were explained through large continuation values.

To examine whether the difference is explained by the large returns to higher education for vocational track students, I present evidence from the decomposition presented in equation (11) in the last two rows of Table 4. The returns to academic studies in Switzerland exhibit dynamic substitutability, as pursuing such tracks instead of vocational education lowers the estimated returns to higher education completion, fitting in with evidence shown in Section 5.2.[22] These results closely fit in with evidence by Rodríguez et al. (2018), who show that job training participation in Chile exhibits dynamic substitutability along with small continuation values. In fact, the first panel of Figure A.8 shows that the direct effect of academic tracks is increasing across the cognitive ability distribution, along with a decreasing continuation value, which track the results documented in the Chilean context. As a result, the returns to specific investment tracking choices in upper-secondary school may resemble the returns to job training more closely than to formal schooling in the United States.

Table 4: Estimated Returns to Academic Education: Decomposition

Estimate	Vocational	Trades	Services
Direct Effect	-0.017	-0.004	-0.051
	(0.003)***	(0.002)*	(0.003)***
Continuation Value	-0.005	-0.030	0.040
	(0.002)**	(0.003)***	(0.003)***
Dynamic Comp	-0.034	-0.070	0.033
	(0.003)***	(0.003)***	(0.003)***
Dyn. Sorting Gains	0.029	0.040	0.007
	(0.002)***	(0.002)***	(0.002)***
ATE	-0.022	-0.034	-0.012
	(0.002)***	(0.002)***	(0.002)***

Note: Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table 4 presents a decomposition of estimated wage returns academic education relative to vocational education into the direct effect and continuation value. The continuation value is further decomposed into dynamic complementarities of human capital investment and a dynamic sorting gains parameter.

In the last two columns of Table 4, I examine whether these findings hold across the two vocational tracks considered in the analysis. The estimated parameters exhibit striking differences across tracks: whereas the direct effect of academic studies relative to trades-based vocational tracks is small (-0.4%), the estimated magnitude is far larger vis-à-vis service-focused tracks (-5.1%). As such, the CV parameters exhibit important differences across tracks, as well, since the continuation value relative to trades is negative (-3%), while positive and significant (4%) when compared to services-based occupations. The difference in the estimated continuation values across vocational tracks largely stems from differential returns to tertiary education documented in Section 5.2. As a result, while the returns to

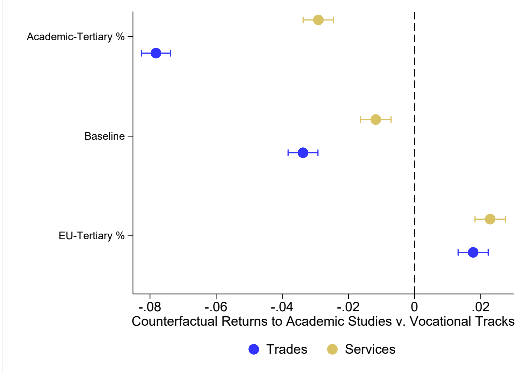
[22] The dynamic sorting gains parameter is positive and significant (2.9%), thus indicating a sizable share of the returns to academic studies is explained through the increased likelihood of tertiary degree completion.

academic studies relative to trades-based tracks exhibit dynamic substitutability, equaling -7%, I find evidence of dynamic complementarity when compared to service-based tracks (3.3%). All in all, these results further highlight important differences in the underlying economic parameters governing the returns to *specific* vocational tracks.

Counterfactual Exercise: Tertiary Education in Vocational Tracks. The evidence presented above shows sizable returns to tertiary education for vocational-track students. In fact, as shown in Figure 1, Switzerland has one of the highest tertiary-education completion rates for these students, yet their graduation rates are far below their peers in academic tracks. To examine the importance of higher education in driving the estimated returns to tracks in Switzerland, I conduct two counterfactual exercises and present the results in Figure 5. I first estimate the returns to academic tracks under an alternative scenario in which tertiary-education participation rates for vocational students in Switzerland would equal the European Union average. The estimated returns to academic tracks relative to services and trades would become positive and statistically significant, reaching 2.3% and 1.7% respectively. A counterfactual scenario with equivalent higher-education participation rates across all tracks (80%) would result in far larger returns to vocational tracks relative to academic studies, exceeding 8% for vocational tracks and equaling 2.9% for service-based tracks. This exercise further shows that the positive returns to vocational education in Switzerland are driven by the strong linkage between vocational and tertiary education.

While Switzerland's firm-based approach to vocational schooling has received extensive attention, these results show that the returns to distinct types of vocational training are not significantly different from each other. On the other hand, across both types of tracks, the strong linkage of Switzerland's vocational education to tertiary schooling is the critical driver of positive outcomes for students in these tracks. The findings presented in this paper thus offer insights on the returns to vocational education not just in countries with firm-based training, but rather in contexts with high tertiary-enrollment rates among its vocational graduates, such as the Netherlands, Belgium, Denmark and Spain (Eichhorst et al., 2015).

Figure 5: Counterfactual Simulation: Higher Education Completion in Vocational Tracks



Note: Figure 5 shows the estimated wage returns to academic studies relative to trades- and service-focused vocational tracks across counterfactual scenarios with different graduation rates in tertiary education.

6 Conclusion

The posited impact of vocational education on labor market outcomes has received growing attention as a potential pathway for improving the school-to-work transition. Yet receiving specific training in the early life may limit students' capacity to adapt to a changing work environment, thus worsening their wage outcomes. In this paper, I have estimated the returns to secondary school tracking choices in Switzerland, relying on longitudinal data covering students' skills, their schooling progression and age-30 labor market outcomes.

The results indicate that the returns to academic studies vis-à-vis vocational tracks are negative. In fact, by incorporating tracking choices within vocational education, I have further shown heterogeneous returns across the specific track under consideration. Moreover, while students in secondary schooling progress through different tertiary education programs at differential rates, the returns to higher education degrees are positive for students in academic and vocational tracks. Counterfactual policy simulations thus show that the positive returns to vocational schooling are driven by its strong linkage with the higher education system. Future work on vocational education should further examine the relationship between specific investments in secondary and tertiary schooling.

References

Altonji, J. G. (1995). The effects of high school curriculum on education and labor market outcomes. *Journal of Human Resources 30*(3), 409–439.

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

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2016). College attrition and the dynamics of information revelation. Technical report, National Bureau of Economic Research.

Bertrand, M., M. Mogstad, and J. Mountjoy (2019). Improving educational pathways to social mobility: Evidence from Norway's "Reform 94". Technical report, National Bureau of Economic Research.

Brunello, G. and L. Rocco (2017). The labor market effects of academic and vocational education over the life cycle: Evidence based on a British cohort. *Journal of Human Capital 11*(1), 106–166.

Carneiro, P., K. T. Hansen, and J. J. Heckman (2003). 2001 Lawrence R. Klein Lecture Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review 44*(2), 361–422.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011). Estimating marginal returns to education. *American Economic Review 101*(6), 2754–81.

Cortes, K. E., J. S. Goodman, and T. Nomi (2015). Intensive math instruction and educational attainment long-run impacts of double-dose algebra. *Journal of Human Resources 50*(1), 108–158.

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

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Eckardt, D. (2019). Are chemists good bankers? Returns to the match between training and occupation. Technical report, London School of Economics.

Eichhorst, W., N. Rodríguez-Planas, R. Schmidl, and K. F. Zimmermann (2015). A road map to vocational education and training in industrialized countries. *ILR Review 68*(2), 314–337.

Eisenhauer, P., J. J. Heckman, and E. Vytlacil (2015). The generalized Roy model and the cost-benefit analysis of social programs. *Journal of Political Economy 123*(2), 413–443.

Freyberger, J. (2018). Non-parametric panel data models with interactive fixed effects. *The Review of Economic Studies 85*(3), 1824–1851.

Golsteyn, B. H. and A. Stenberg (2017). Earnings over the life course: General versus vocational education. *Journal of Human Capital 11*(2), 167–212.

Gomensoro, A. and T. Meyer (2017). Tree (transitions from education to employment): A Swiss multi-cohort survey. *Longitudinal and Life Course Studies 8*(2), 209–224.

Goodman, J. (2019). The labor of division: Returns to compulsory high school math coursework. *Journal of Labor Economics 37*(4), 1141–1182.

Hall, C. (2012). The effects of reducing tracking in upper secondary school evidence from a large-scale pilot scheme. *Journal of Human Resources 47*(1), 237–269.

Hall, C. (2016). Does more general education reduce the risk of future unemployment? evidence from an expansion of vocational upper secondary education. *Economics of Education Review 52*, 251–271.

Hampf, F. and L. Woessmann (2017). Vocational vs. general education and employment over the life cycle: New evidence from PIAAC. *CESifo Economic Studies 63*(3), 255–269.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121*(1-2), 39–98.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review 73*, 103–130.

Hanushek, E. A., G. Schwerdt, L. Woessmann, and L. Zhang (2017). General education, vocational education, and labor-market outcomes over the lifecycle. *Journal of Human Resources 52*(1), 48–87.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2016). Dynamic treatment effects. *Journal of Econometrics 191*(2), 276–292.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

Heckman, J. J. and S. Navarro (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics 136*(2), 341–396.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Hoeckel, K., S. Field, and W. N. Grubb (2009). Learning for Jobs: OECD Reviews of Vocational Education and Training: Switzerland. *Organisation for Economic Co-Operation and Development. April, Paris.*

Hoeckel, K. and R. Schwartz (2010). Learning for Jobs OECD Reviews of Vocational Education and Training. *Austria: Organisation for Economic Co-operation and Development (OECD).*

Humphries, J. E., J. S. Joensen, and G. F. Veramendi (2018). Complementarities in high school and college investments. *mimeo*.

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

Kautz, T., J. J. Heckman, R. Diris, B. Ter Weel, and L. Borghans (2014). Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. Technical report, National Bureau of Economic Research.

Korber, M. (2019). *The labour market returns to vocational education over the life course.* Ph. D. thesis, Université de Lausanne, Faculté des sciences sociales et politiques.

Kotlarski, I. (1967). On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics 20*(1), 69–76.

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

Malamud, O. and C. Pop-Eleches (2010). General education versus vocational training: Evidence from an economy in transition. *The review of economics and statistics 92*(1), 43–60.

Mihály, F. and F. Simon (2013). *OECD Reviews of Vocational Education and Training A Skills beyond School Review of Switzerland*, Volume 2011. OECD Publishing.

Mountjoy, J. (2019). Community colleges and upward mobility. *Available at SSRN 3373801*.

Neyt, B., D. Verhaest, and S. Baert (2020). The impact of dual apprenticeship programmes on early labour market outcomes: A dynamic approach. *Economics of Education Review 78*, 102022.

Nybom, M. (2017). The distribution of lifetime earnings returns to college. *Journal of Labor Economics 35*(4), 903–952.

Oosterbeek, H. and D. Webbink (2007). Wage effects of an extra year of basic vocational education. *Economics of Education Review 26*(4), 408–419.

Prada, M. F. and S. Urzua (2017). One size does not fit all: Multiple dimensions of ability, college attendance, and earnings. *Journal of Labor Economics 35*(4), 953–991.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the american statistical association 53*(284), 873–880.

Ray, A., W. Margaret, et al. (2003). *PISA Programme for international student assessment (PISA) PISA 2000 technical report: PISA 2000 technical report.* oecd Publishing.

Rodríguez, J., F. Saltiel, and S. S. Urzúa (2018). Dynamic treatment effects of job training. Technical report, National Bureau of Economic Research.

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

Silliman, M. and H. Virtanen (2019). Labor market returns to vocational secondary education. Technical report, ETLA Working Papers.

Stange, K. M. (2012). An empirical investigation of the option value of college enrollment. *American Economic Journal: Applied Economics 4*(1), 49–84.

Stinebrickner, T. and R. Stinebrickner (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics 30*(4), 707–748.

Williams, B. (2019). Identification of the linear factor model. *Econometric Reviews*, 1–18.

SUPPLEMENTARY APPENDICES

Appendix

A Appendix Tables and Figures

Table A.1: Determinants of Taking Math, Reading and Science Tests

	(1)
Age	-0.007
	(0.025)
Pre-Gymnasial Track	-0.030
	(0.030)
German Region	0.004
	(0.048)
French Region	-0.007
	(0.048)
Both Parents	0.027
	(0.037)
Parents' Ed: Tertiary	-0.036
	(0.027)
Constant	0.342
	(0.375)
Observations	1,001
R^2	0.004

Source: TREE (Transitions from Education to Employment). Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Table A.1 presents evidence on the determinants of having taken all three (math, reading and science) assessments.

Table A.2: Non-Cognitive Skill Measures: Underlying Questions

NC Measure	Underlying Questions
Perseverance	"I work as hard as possible,"
	"I keep working even if the material is difficult,"
	"I try to do my best to acquire the knowledge and skills taught"
	"I put forth my best effort"
Memorization	"I try to memorize everything that might be covered,"
	"I memorize as much as possible,"
	"I memorize all new material so that I can recite it,"
	"I practice by saying the material to myself over and over"
Control Strategies	I start by figuring out exactly what I need to learn,"
	I force myself to check to see if I remember what I have learned,"
	"I try to figure out which concepts I still haven't really understood,"
	"I make sure that I remember the most important things"
	"I don't understand something I look for additional information."
Self-efficacy	I'm certain I can understand the most difficult material presented in texts,"
	"I'm confident I can do an excellent job on assignments and tests"
	"I'm certain I can master the skills being taught."
Control Expectation	"When I sit myself down to learn something really difficult, I can learn it,"
	"If I decide not to get any bad grades, I can really do it,"
	"If I decide not to get any problems wrong, I can really do it,"
	"If I want to learn something well, I can."
Elaboration Strategies	"I try to relate new material to things I have learned in other subjects,"
	"I figure out how the information might be useful in the real world,"
	"I try to understand the material better by relating it to things I already know,"
	"I figure out how the material fits in with what I have already learned."

Note: Table A.2 presents an inventory of the questions used to construct the six non-cognitive skill measures available in the 2000 PISA exam.

Table A.3: Summary Statistics by HS Track and Tertiary Educational Attainment

	Academic		Trades		Services	
	No HEI	HEI Grad.	No HEI	HEI Grad.	No HEI	HEI Grad.
	(1)	(2)	(3)	(4)	(5)	(6)
Background Characteristics						
Age	15.444	15.354	15.630	15.563	15.659	15.689
Swiss	0.916	0.922	0.877	0.970***	0.835	0.958**
Pre-Gymnasial Track	0.627	0.737*	0.171	0.284**	0.203	0.354*
German Region	0.301	0.406	0.493	0.599*	0.557	0.646
French Region	0.518	0.451	0.379	0.310	0.304	0.323
Both Parents	0.795	0.848	0.839	0.878	0.823	0.844
Parents' Ed: Tertiary	0.458	0.579*	0.275	0.391*	0.266	0.438*
Test Scores						
PISA Reading Score	-0.077	0.388***	-0.787	-0.163***	-0.672	-0.174**
Control Strategies	0.061	0.233	-0.447	-0.226*	-0.112	-0.103
Perseverance	0.038	0.135	-0.280	-0.142	-0.008	-0.005
Memorization	0.156	0.121	-0.194	-0.233	-0.180	-0.087
Self-Efficacy	0.232	0.468	-0.165	0.151***	0.036	0.071
Control Expectation	0.043	0.356***	-0.215	-0.068	0.002	-0.077
Elaboration Strategies	0.227	0.279	-0.201	0.084**	-0.023	-0.070
Hourly Wages	27.38	33.01***	26.12	33.14***	26.42	30.90***
Observations	83	335	211	197	79	96

Source: TREE (Transitions from Education to Employment). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Table A.3 displays summary statistics by high school track, including vocational tracks (Trades or Services) and tertiary education completion through age 30. The stars in even-numbered columns capture the statistical significance of a t-test comparing the means of the variables of tertiary education completers vis-a-vis their counterparts who did not complete a degree by age 30 within each educational group.

Table A.4: OLS Estimates: Wage Returns to HS Tracks

		Tertiary Education by HS Track		
	HS Track	Academic	Trades	Services
	(1)	(2)	(3)	(4)
Trades	0.022			
	(0.028)			
Services	0.016			
	(0.032)			
Tertiary Graduate		0.123***	0.207***	0.119***
		(0.045)	(0.031)	(0.046)
Background Characteristics	✓	✓	✓	✓
Baseline Test Scores	✓	✓	✓	✓
Baseline Non-Cognitive	✓	✓	✓	✓
Observations	1,001	418	408	175
R^2	0.042	0.057	0.198	0.144

Source: TREE (Transitions from Education to Employment). Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: The first column of Table A.4 presents the estimated OLS age-30 wage returns to vocational tracks in secondary school relative to academic studies. The last three columns present estimated returns to tertiary degree completion for students who completed different secondary school tracks. All regressions control for the background characteristics, test scores and non-cognitive skills presented in Table 1.

Table A.5: OLS Estimates: Wage Returns to Baseline Skill Measures

	All (1)	Academic (2)	Trades (3)	Services (4)	All (5)
Reading Test Score	0.025**	0.030	0.027*	0.015	
	(0.011)	(0.021)	(0.016)	(0.023)	
Non-Cognitive Skills	0.008	0.008	0.013	0.001	0.014
	(0.010)	(0.016)	(0.018)	(0.024)	(0.015)
Math Test Score					0.021
					(0.016)
Background Characteristics	✓	✓	✓	✓	
Observations	1001	418	408	175	563
R^2	0.035	0.018	0.086	0.050	0.032

Source: TREE (Transitions from Education to Employment). Robust standard errors errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: The first column of Table A.5 presents the estimated OLS age-30 wage returns to reading test scores and non-cognitive skills (constructed from principal components analysis of the six underlying measures available in TREE). The second, third and fourth columns present conditional returns to skills for students who completed academic, trades-based and services-based tracks, respectively. The last column replicates the results from the first column using the sample of students who took the PISA math assessment.

Table A.6: Empirical Specification

	Measurement (1)	Secondary Track (2)	Vocational Track (3)	Tertiary (4)	Wages (5)
Observables					
Age	Yes	Yes	Yes	Yes	Yes
Swiss	Yes	Yes	Yes	Yes	Yes
Region	Yes	Yes	Yes	Yes	Yes
Both Parents	Yes	Yes	Yes	Yes	Yes
Parents' Ed.	Yes	Yes	Yes	Yes	Yes
Lower Secondary Track	Yes	Yes	Yes	Yes	Yes
Instruments					
UN Rate in Track		Yes			
% of Canton in Track		Yes	Yes		
UN Rate in Tertiary				Yes	
Latent Ability					
Cognitive Ability	Yes	Yes	Yes	Yes	Yes
Non-Cognitive Ability	Yes	Yes	Yes	Yes	Yes

Note: Table A.6 presents the variables included in model implementation. In the measurement system, I use PISA test scores for reading, math and science along with six non-cognitive skill measures.

Table A.7: Goodness of Fit: Hourly Wage Outcomes

Estimate	Academic	Trades	Services
Tertiary Education Completers			
Actual	3.39	3.44	3.39
	(0.003)	(0.003)	(0.004)
Model	3.38	3.43	3.39
	(0.003)	(0.003)	(0.005)
Secondary Education Completers			
Actual	3.24	3.21	3.24
	(0.006)	(0.003)	(0.004)
Model	3.25	3.21	3.26
	(0.007)	(0.003)	(0.005)

Source: TREE (Transitions from Education to Employment). Note: Standard errors in parenthesis. Table A.7 compares observed and simulated hourly wages for individuals across secondary and tertiary educational choices.

Table A.8: Measurement System Loadings

	Reading (1)	Math (2)	Science (3)	Control Strategies (4)	Perseverance (5)	Memorization (6)	Self-Efficacy (7)	Control Expectation (8)	Elaboration Strategies (9)
Constant	2.52	2.41	2.39	-0.37	-0.45	1.17	0.89	0.93	-0.28
	(0.84)	(1.07)	(1.02)	(0.85)	(0.85)	(0.81)	(0.83)	(0.89)	(0.87)
Age	-0.21	-0.20	-0.19	0.00	0.01	-0.07	-0.08	-0.08	0.01
	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)
Pre-Gymnasial Track	0.85	0.78	0.77	0.33	0.14	0.06	0.19	0.11	0.18
	(0.06)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
German Region	0.38	0.51	0.37	-0.01	0.13	-0.62	0.45	0.05	0.03
	(0.11)	(0.13)	(0.13)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)
French Region	-0.19	0.17	-0.10	-0.18	-0.17	0.13	0.24	0.23	-0.08
	(0.10)	(0.13)	(0.13)	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)	(0.11)
Both Parents	0.11	0.24	0.08	0.20	0.10	0.11	0.06	0.16	0.03
	(0.08)	(0.10)	(0.10)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)
Parents' Ed: Tertiary	0.12	0.14	0.11	0.07	0.07	-0.06	0.19	0.11	0.18
	(0.06)	(0.08)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
θ_C	1.00	0.80	0.91	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.07)	(0.08)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
θ_{NC}	0.16	0.00	0.18	1.00	0.89	0.65	0.82	0.83	0.78
	(0.03)	(0.00)	(0.04)	(0.00)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Precision	3.23	1.97	2.43	3.57	2.61	1.96	2.00	2.03	2.02
	(0.48)	(0.16)	(0.24)	(0.25)	(0.15)	(0.10)	(0.10)	(0.10)	(0.10)
Sample Size	1,001	563	567	1,001					

Source: TREE data. Standard errors in parenthesis. Note: Table A.8 presents the estimated parameters from the measurement system presented in equations (6)-(7). I obtain these estimates by simulating 500 values of parameters using the estimated posterior from the MCMC estimator. The 'Sample Size' row denotes the number of individuals included in each node of the data for model estimation.

Table A.9: Loadings on Educational Choices

	Academic Track (1)	Vocational Track (2)	Tertiary: Academic (3)	Tertiary: Trades (4)	Tertiary: Services (5)
Constant	11.04	0.01	7.67	5.64	-0.34
	(2.78)	(1.87)	(2.24)	(2.22)	(2.86)
Age	-0.40	-0.07	-0.48	-0.40	-0.04
	(0.09)	(0.10)	(0.15)	(0.14)	(0.18)
Pre-Gymnasial Track	1.62	-0.33	0.61	0.70	0.55
	(0.12)	(0.14)	(0.20)	(0.19)	(0.30)
German Region	-0.97	-0.18	0.25	0.59	1.03
	(0.17)	(0.27)	(0.27)	(0.47)	(0.68)
French Region	-0.90	-0.01	-0.15	-0.13	0.76
	(0.17)	(0.22)	(0.26)	(0.38)	(0.56)
Both Parents	-0.12	0.16	0.08	0.05	0.08
	(0.14)	(0.16)	(0.20)	(0.21)	(0.30)
Parents' Ed: Tertiary	0.52	-0.05	0.33	0.45	0.36
	(0.10)	(0.11)	(0.16)	(0.15)	(0.23)
% of Canton in Track: Long-Run	-0.72				
	(0.33)				
% of Canton in Track: 2000	0.21				
	(0.10)				
Local UN Rate in Track: Long-Run	-4.47				
	(2.05)				
Local UN Rate in Track: 2000	-0.44				
	(0.36)				
% of Canton in Track: Long-Run		0.51			
		(0.52)			
% of Canton in Track: 2000		-0.28			
		(0.41)			
Local UN Rate in Tertiary: Long-Run			-4.79	3.09	3.82
			(2.36)	(3.06)	(4.71)
Local UN Rate in Tertiary: 2000			1.11	-1.74	-1.75
			(1.20)	(1.52)	(2.23)
θ_C	0.52	0.02	0.33	0.75	0.53
	(0.09)	(0.09)	(0.15)	(0.15)	(0.21)
θ_{NC}	0.36	-0.15	0.10	0.28	0.02
	(0.06)	(0.07)	(0.08)	(0.10)	(0.13)
Precision	1.00	1.00	1.00	1.00	1.00
Sample Size	1,001	518	418	408	175

Source: TREE data. Standard errors in parenthesis. Note: Table A.9 presents the estimated parameters from the educational choice probits presented in Section 3. I obtain these estimates by simulating 500 values of parameters using the estimated posterior from the MCMC estimator. The 'Sample Size' row denotes the number of individuals included in each node of the data for model estimation.

Table A.10: Loadings on Hourly Wage Equations

	Academic Track		Trades		Services	
	No HEI	HEI	No HEI	HEI	No HEI	HEI
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	3.79	4.18	3.36	3.25	3.50	3.91
	(1.23)	(0.63)	(0.58)	(0.64)	(0.90)	(0.84)
Age	-0.04	-0.06	-0.01	0.00	-0.02	-0.05
	(0.08)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)
Pre-Gymnasial Track	-0.06	0.03	0.08	0.12	0.06	0.01
	(0.12)	(0.06)	(0.06)	(0.05)	(0.11)	(0.08)
German Region	-0.07	0.10	0.09	0.15	0.15	0.17
	(0.16)	(0.08)	(0.08)	(0.08)	(0.10)	(0.20)
French Region	0.23	-0.00	0.08	0.11	0.11	0.10
	(0.16)	(0.08)	(0.08)	(0.09)	(0.13)	(0.20)
Both Parents	0.10	-0.01	-0.11	-0.02	-0.14	0.11
	(0.10)	(0.05)	(0.06)	(0.07)	(0.10)	(0.10)
Parents' Ed: Tertiary	-0.02	-0.00	0.05	0.05	0.01	0.01
	(0.09)	(0.04)	(0.05)	(0.04)	(0.09)	(0.07)
θ_C	0.00	0.04	-0.05	0.04	0.01	0.04
	(0.08)	(0.03)	(0.03)	(0.04)	(0.06)	(0.06)
θ_{NC}	0.02	0.01	0.01	-0.01	0.00	0.01
	(0.05)	(0.02)	(0.03)	(0.03)	(0.04)	(0.05)
Precision	7.04	8.03	10.28	12.03	10.87	9.57
	(1.10)	(0.59)	(1.02)	(1.28)	(1.84)	(1.38)
Observations	83	335	211	197	79	96

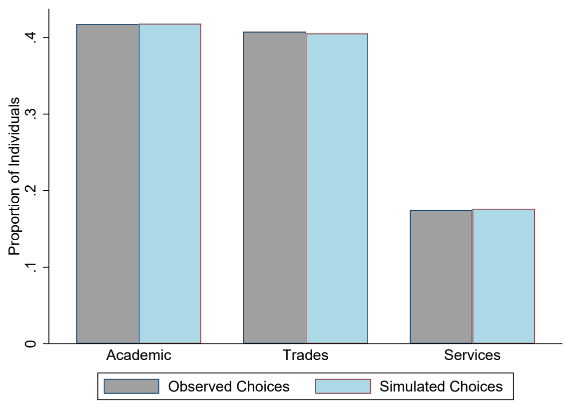
Source: TREE data. Standard errors in parenthesis. Note: Table A.10 presents the estimated parameters from the potential wage outcome equations presented in Section 3. I obtain these estimates by simulating 500 values of parameters using the estimated posterior from the MCMC estimator. The 'Sample Size' row denotes the number of individuals included in each node of the data for model estimation.

Table A.11: Treatment Effects Decomposition by Vocational Track

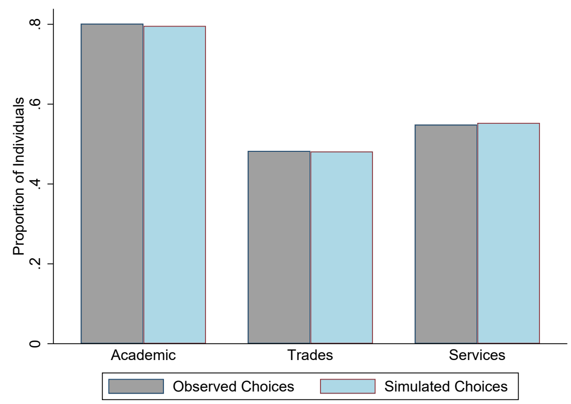
	Decomposition by Tracking Decision		
Estimate	Vocational	Trades	Services
	(1)	(2)	(3)
Observed Differences	0.033	0.037	0.023
	(0.003)	(0.004)	(0.005)
Average Treatment Effect	-0.022	-0.034	-0.012
	(0.002)	(0.003)	(0.003)
Selection Bias	0.076	0.103	0.033
	(0.003)	(0.003)	(0.005)
Sorting Gains	-0.020	-0.031	0.003
	(0.004)	(0.004)	(0.005)

Source: TREE data. Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Table A.11 presents the decomposition introduced in Heckman et al. (2018) of observed differences across academic and vocational tracks into average treatment effects, selection bias and sorting gains as defined in Section 5. All parameters follow from the estimated model. the estimated model.

Figure A.1: Goodness of Fit: Educational Attainment



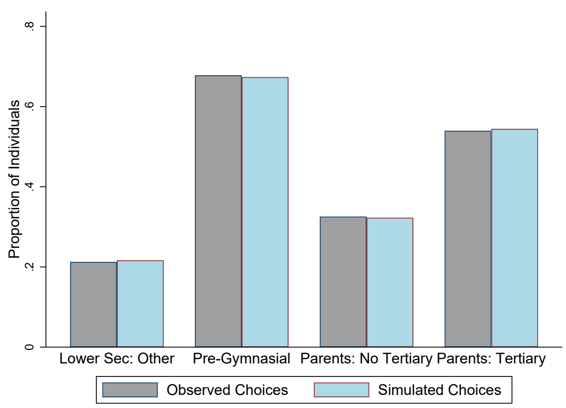
(a) Secondary School Tracks



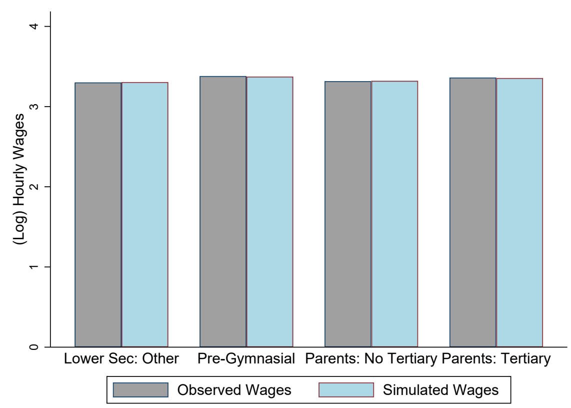
(b) Tertiary Education Completion

Source: TREE data. Figure A.1 compares the observed and simulated share of students who pursue different secondary school tracks (Panel A) and the likelihood of completing a tertiary education degree by secondary school track (Panel B).

Figure A.2: Goodness of Fit: Conditional on Baseline Characteristics



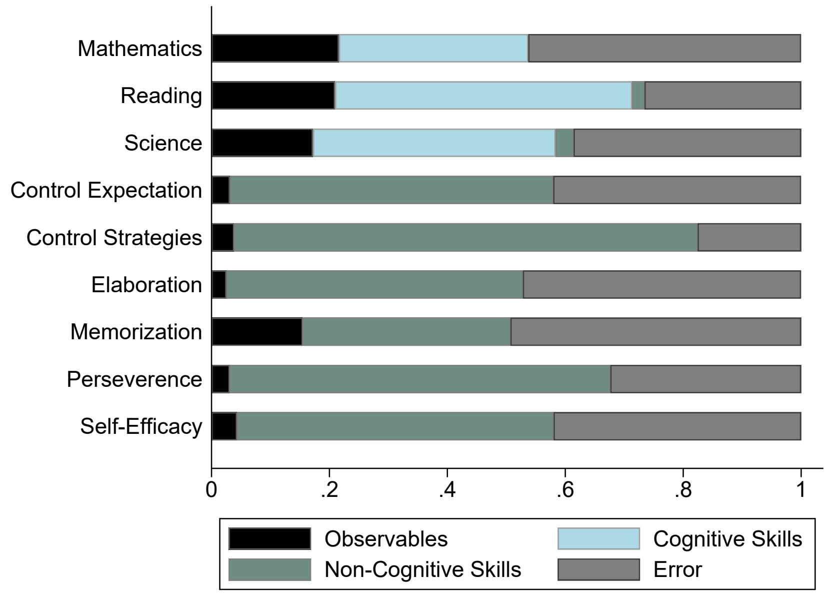
(a) Secondary School Tracks



(b) Hourly Wages

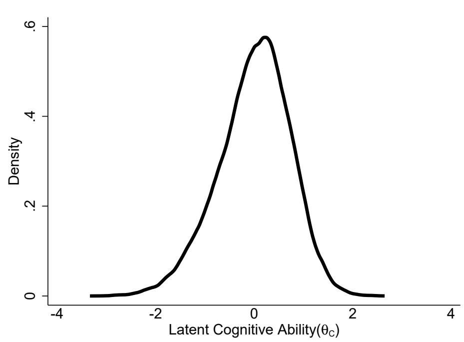
Source: TREE data. Figure A.2 compares the observed and simulated share of students who pursue different secondary school tracks across lower-secondary tracks and parental educational attainment (Panel A) as well as average observed and simulated hourly wages (Panel B) across the same characteristics.

Figure A.3: Variance Decomposition of Observed Measures

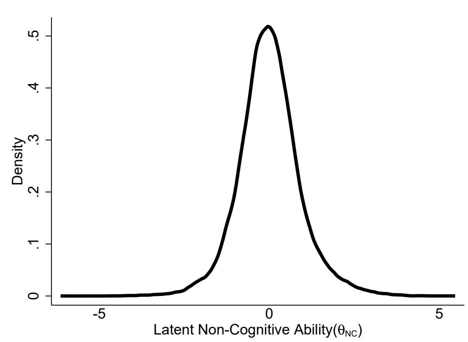


Source: TREE data. Figure A.3 shows the contribution of each variable to the variance of test scores and non-cognitive skill measures using the simulated sample. The row Observables indicates the share of the variance of the measurement variables explained by the observed variables included in the measurement system. Each Factor bar indicates the share of the variance explained by the respective component of Cognitive and Non-Cognitive ability. Finally, the label Error term represents the share of each test score variance explained by the unobserved idiosyncratic error of the measurement system.

Figure A.4: Distribution of Latent Ability



(a) Distribution of Cognitive Ability

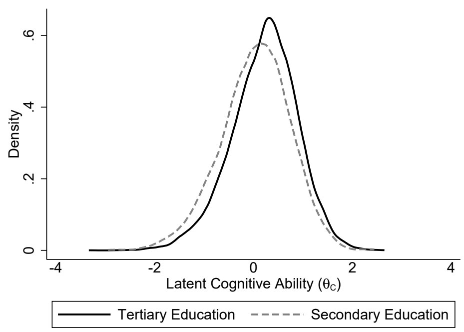


(b) Distribution of Non-Cognitive Ability

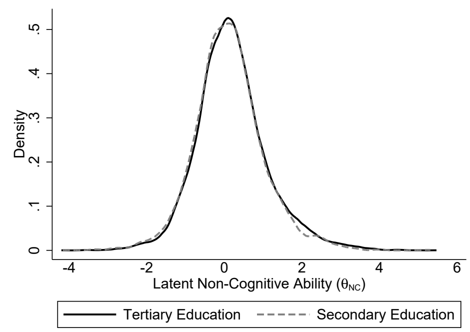
Source: TREE data. Figure A.4 shows the marginal densities of the estimated latent cognitive and non-cognitive factors.

Figure A.5: Density of Latent Ability by Tertiary Education Choice

Density of $\boldsymbol{\theta}$ in Tertiary Education: Academic Track

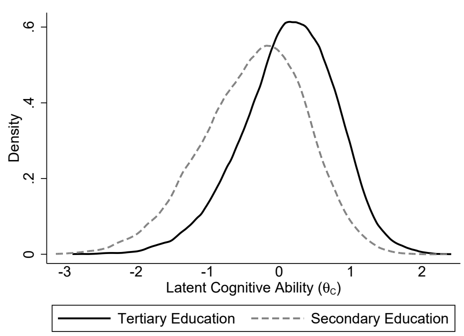


(a) Density of θ_C by Tertiary Education

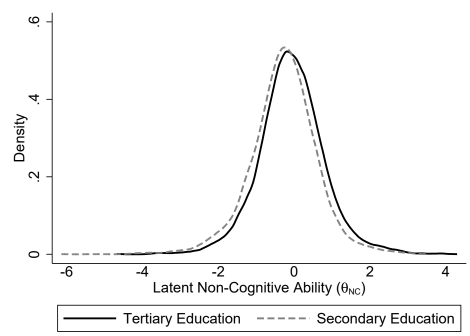


(b) Density of θ_{NC} by Tertiary Education

Density of $\boldsymbol{\theta}$ in Tertiary Education: Vocational-Trades Track

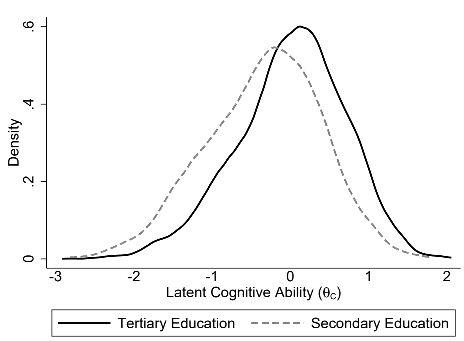


(c) Density of θ_C by Tertiary Education

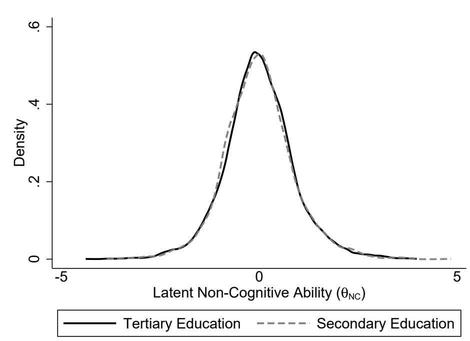


(d) Density of θ_{NC} by Tertiary Education

Density of $\boldsymbol{\theta}$ in Tertiary Education: Vocational-Services Track



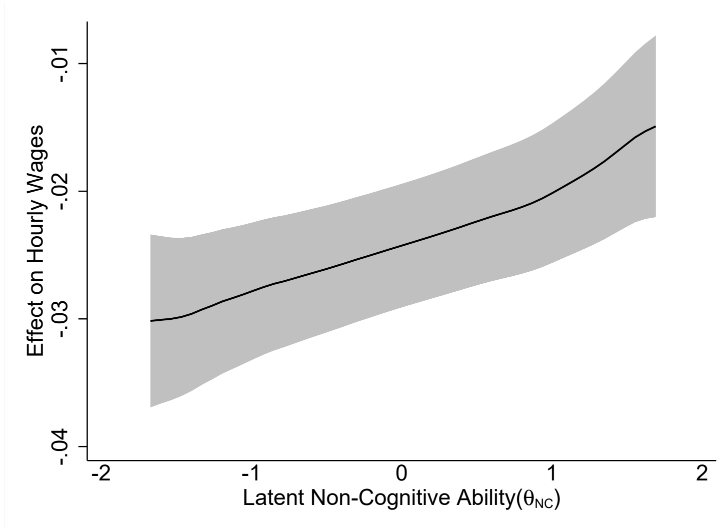
(e) Density of θ_C by Tertiary Education



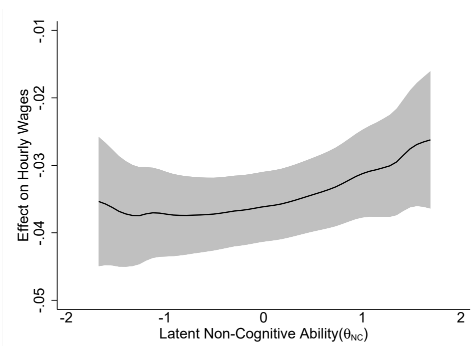
(f) Density of θ_{NC} by Tertiary Education

Source: TREE data. Figure A.5 shows the marginal densities of the cognitive and non-cognitive latent factors across tertiary education completion status for students in academic tracks (panels (a) and (b)), vocational-trades tracks (panels (c) and (d)) and vocational-services tracks (panels (e) and (f)).

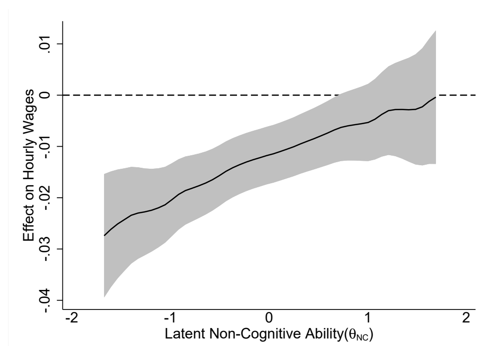
Figure A.6: Heterogeneous Returns to Academic Studies Relative to Vocational Tracks



(a) Academic v. Vocational



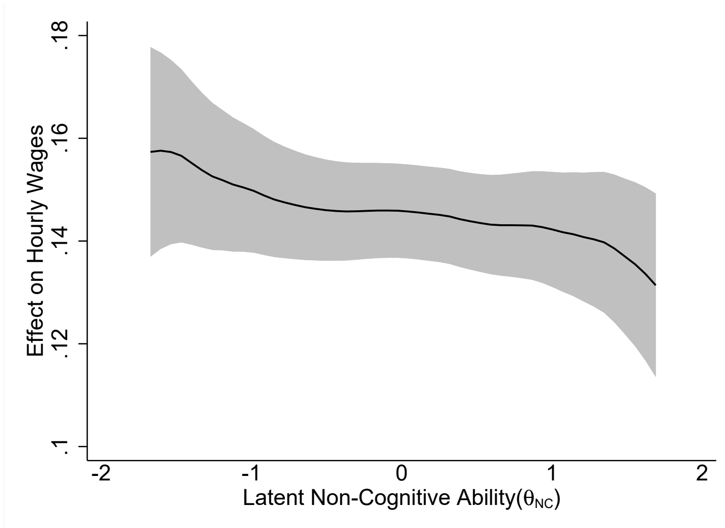
(b) Academic v. Vocational-Trades



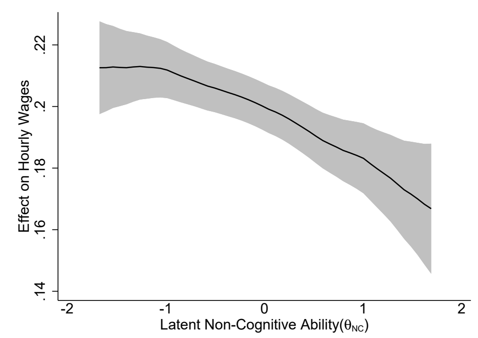
(c) Academic v. Vocational-Services

Source: TREE data. Figure A.6 presents the heterogeneous average treatment effect of academic studies relative to vocational tracks on hourly wages across the latent non-cognitive ability distribution.

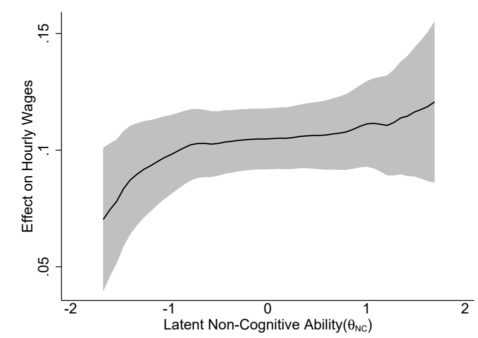
Figure A.7: Heterogeneous Returns to Tertiary Education Across Secondary School Paths



(a) Academic Tracks



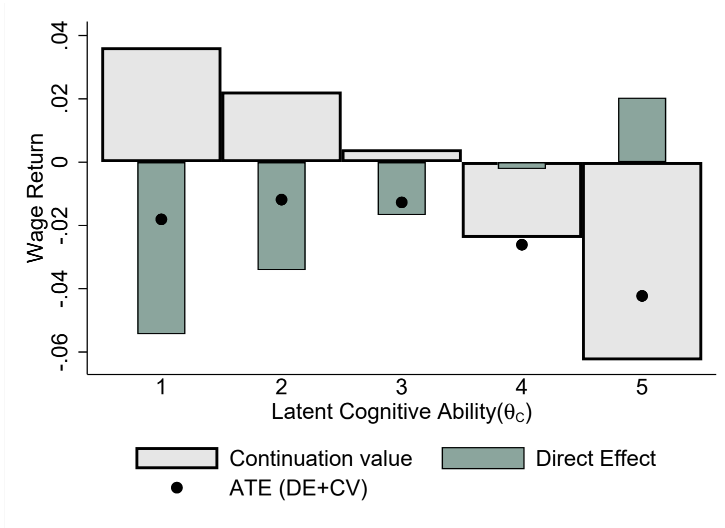
(b) Vocational-Trades Track



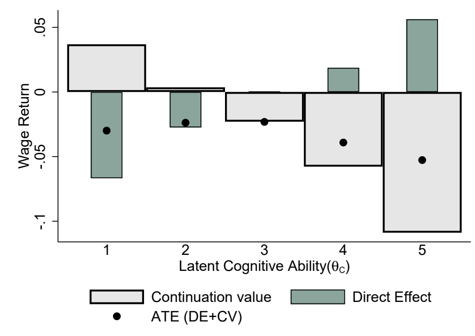
(c) Vocational-Services Track

Source: TREE data. Figure A.7 presents the heterogeneous average treatment effect to tertiary education across secondary school paths on hourly wages across the latent non-cognitive ability distribution.

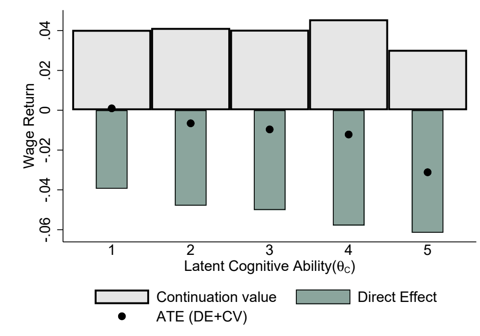
Figure A.8: Heterogeneous Returns to Academic Education: Decomposition



(a) Academic v. Vocational



(b) Academic v. Vocational-Trades



(c) Academic v. Vocational-Services

Source: TREE data. Figure A.8 presents the heterogeneous wage returns to academic education relative to vocational studies the latent cognitive ability (θ_C) distribution, decomposed into the direct effect and continuation value components.

B Identification of the Measurement System

The identification of the distribution of unobserved ability follows the formal arguments presented in Carneiro et al. (2003) and Hansen et al. (2004).[23] I first consider identification of the elements of equation (7) in light of the dedicated measurement system for non-cognitive skills. The diagonal and off-diagonal elements of the matrix $COV(\boldsymbol{NC_i})$ are respectively given by:

$$COV(NC^n, NC^n) = (\gamma^{NC,n})^2\sigma^2_{\theta_{NC}} + \sigma^2_{e^{NC,n}} \tag{12}$$

$$COV(NC^n, NC^{n'}) = \gamma^{NC,n}\gamma^{NC,n'}\sigma^2_{\theta_{NC}} \tag{13}$$

Since latent factors have no scale of their own, Carneiro et al. (2003) note that one of the factor loadings can be normalized to unity to set the scale of each ability component ($\gamma^{NC,1} = 1$). Moreover, as long as the number of observed non-cognitive skill measures N meets the following condition ($\frac{N(N-1)}{2} > N$), the remaining loadings and the variance of the latent factor ($\sigma^2_{\theta_{NC}}$) can be identified from the covariance of the off-diagonal elements. Since this condition is met in the PISA data ($N = 6$), the diagonal elements of $COV(\boldsymbol{NC_i})$ can be then used to identify the distribution of the error terms in the measurement system ($\sigma^2_{e^{NC,n}}$).

A similar argument follows for identifying the elements of the equation (6). First, the off-diagonal elements of the matrix $COV(\boldsymbol{C_i})$ are given by:

$$COV(C^j, C^{j'}) = \alpha^{C,j}\alpha^{C,j'}\sigma^2_{\theta_C} + \alpha^{NC,j}\alpha^{NC,j'}\sigma^2_{\theta_{NC}} + (\alpha^{C,j}\alpha^{NC,j'} + \alpha^{C,j'}\alpha^{NC,j})\sigma_{\theta_C,\theta_{NC}} \tag{14}$$

where equation (14) shows that there are two unknown loadings for each observed cognitive measure (C^j) and the variance of θ_C ($\sigma^2_{\theta_C}$) and covariance of the latent factors ($\sigma_{\theta_C,\theta_{NC}}$) are not yet identified. In fact, the three covariances across cognitive test scores are not sufficient for identifying these eight parameters. Upon normalizing one loading to unity ($\alpha^{C,j}$) and assuming that one cognitive measure is exclusively affected by latent cognitive ability ($\alpha^{NC,j} = 0$), I can rely on the covariance of observed cognitive measures and non-cognitive measure n to identify the remaining parameters.[24] Upon securing identification of the factor loadings, along with the variance and covariance of the latent factors, the diagonal elements of $COV(\boldsymbol{C_i})$ can be used to identify $\sigma^2_{e^{C,j}}$. Having secured the identification of all the loadings and the variance of each component of latent ability, I rely on the identification arguments presented in Freyberger (2018) to non-parametrically identify the distribution of the latent factors and error terms.[25]

[23] In particular, following the identification argument presented in Williams (2019), I can rely on the covariance across math, reading and science PISA test scores, as these measures are observed for one-fourth of the sample. Throughout this section, I keep the conditioning on $\boldsymbol{X}$ implicit.

[24] This covariance is given by: $COV(C^j, NC^n) = \alpha^{NC,j}\gamma^{NC,n}\sigma^2_{\theta_{NC}} + \alpha^{NC,j}\gamma^{NC,n}\sigma_{\theta_C,\theta_{NC}}$. Since latent factors have no scale of their own, the normalization-to-unity assumption is innocuous. Heckman et al. (2018) similarly assume one dedicated measure for identifying a model with correlated latent factors.

[25] I fix the cognitive loading of the reading test score to one and the non-cognitive loading in the control strategies measure to equal one. I assume that math performance is a dedicated measure of cognitive ability. The results are robust to alternative normalizations in the measurement system. Freyberger (2018) extends the identification arguments presented by Kotlarski (1967) to a context with correlated factors.

C Empirical Contribution of Exclusion Restrictions

Appendix C presents the estimated returns to upper-secondary and tertiary education following from the model introduced in Section 3 estimated without exclusion restrictions in the choice equations.

Table C.1: Estimated Returns to Vocational Tracks in Secondary School

Estimate	Vocational	Trades	Services
	(1)	(2)	(3)
ATE	-0.026	-0.041	-0.009
	(0.002)***	(0.002)***	(0.002)***
TT	-0.040	-0.072	0.006
	(0.003)***	(0.003)***	(0.004)
TUT	-0.017	-0.012	-0.026
	(0.003)***	(0.004)***	(0.006)***

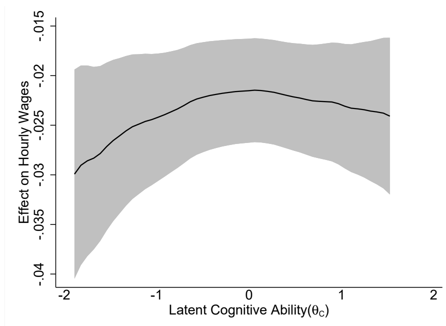
Source: TREE data. Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table C.1 presents the estimated age-30 wage returns to academic studies relative to vocational tracks (column 1) and relative to specific vocational tracks (in columns 2-3). Results follow from a model estimated without exclusion restrictions in the choice equations.

Table C.2: Estimated Returns to Tertiary Schooling Across Secondary School Tracks

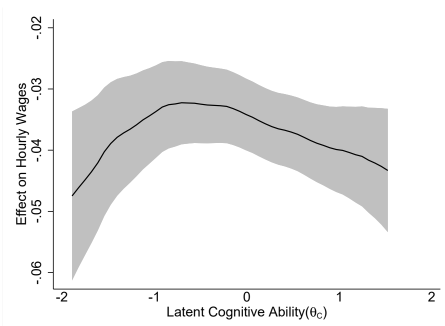
Estimate	Academic	Trades	Services
	(1)	(2)	(3)
ATE	0.158	0.206	0.123
	(0.004)***	(0.003)***	(0.005)***
TT	0.164	0.235	0.125
	(0.004)***	(0.005)***	(0.007)***
TUT	0.132	0.180	0.121
	(0.009)***	(0.004)***	(0.008)***

Source: TREE data. Standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Table C.2 presents the estimated age-30 wage returns to tertiary education across secondary schooling tracks. Results follow from a model estimated without exclusion restrictions in the choice equations.

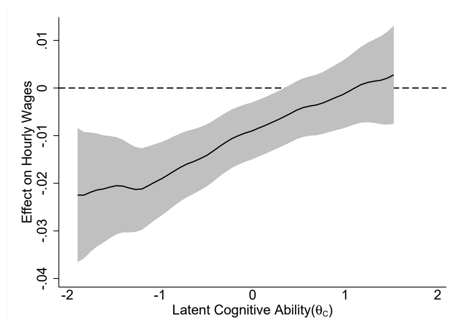
Figure C.1: Heterogeneous Returns to Academic Studies Relative to Vocational Tracks



(a) Academic v. Vocational



(b) Academic v. Vocational-Trades



(c) Academic v. Vocational-Services

Source: TREE data. Figure C.1 presents the heterogeneous average treatment effect of academic studies relative to vocational tracks on hourly wages across the latent cognitive ability distribution. Results follow from a model estimated without exclusion restrictions in the choice equations.