A Work Project, presented as part of the requirements for the Award of a Master Degree in Economics from the NOVA – School of Business and Economics.

Vocational education: coursetaking choice and impact on dropout and college enrollment rates

Rafaela Martinho Henriques

Student Number 3100

A Project carried out in the Master in Economics Program, under the supervision of:

Ana Balcão Reis, Luís Catela Nunes and Maria do Carmo Seabra

January 3rd, 2018

Vocational education: coursetaking choice and impact on dropout and college enrollment rates

Abstract

This work investigates which factors underlie the students' decision on which secondary track program to enroll in and its impact on dropout and college enrollment rates. By separating the analysis between low and high-ability students, we find a heterogeneous effect of vocational coursetaking on dropout probabilities, which increases for high-achievers and decreases for low-achievers. Hence, whereas vocational education appears to be successful in engaging students "at risk", it appears to prejudice the academic success of the high-achievers. A special attention is given to the confounding variables used to estimate this effect. Apart from past school performance, previous retentions, college expectations and parents' education seem to be good predictors of both educational decisions and outcomes.

Keywords: Vocational Education; Tracking; Academic Success; High School Dropout

Acknowledgements

I would like to thank Professors Ana Balcão Reis, Luís Catela Nunes and Maria do Carmo Seabra for their orientation and DGEEC for making the data used in this work available.

I dedicate this Work Project to my family.

1. Introduction

Sound public policy in the education system of a country is key to achieve "*smart, sustainable and inclusive growth"* (Cedefop, 2011). In Portugal, there are visible efforts in education policy to tackle the low qualifications of its citizens. In 2009, secondary education was established as the minimum compulsory schooling level and, since then much attention is being directed towards strengthening *vocational education and training* (VET) in public schools.

According to the OECD, "VET includes education and training programs designed for, and typically leading to, a particular job or type of job." (OECD, 2009). These programs differ from general "regular" ones, which are in essence more academic and aimed at preparing students for college[1]. The distinct curricula in these two trajectories and the different opportunities associated with such coursetaking is nowadays a topic of debate worldwide.

Traditionally, vocational programs were designed to retain at school students that lacked motivation to learn academic material, and that had no future college plans, which scholars have found to correspond to the same subgroup of students that perform lower academically (Jencks, 1972; Oakes, 1985). Hence, by serving as an alternative to general paths, VET would be particularly helpful at engaging students at risk of dropping out (Rasinski and Pedlow, 1998). Additionally, some scholars support that, by teaching specific job-skills, vocational programs may increase students' employment chances (Kulik, 1998; Neuman and Ziederman, 1999).

Classic research in this field has, however, suggested that this educational track is subject to stigma and that these programs are typically perceived as a second option, aimed for those with lower academic abilities (Oakes, 1985). Hence, in an effort to boost the attractiveness of this track, as a means of improving the overall education of its citizens, the Portuguese government

[1] In Portugal, academic programs are designated as "*cursos científico-humanísticos*" corresponding to the "via académica", while vocational educational programs are included in the "*via profissionalizante*".

has, in the last decade, increased both the number and diversity of vocational programs, and made modifications in these programs so that vocational students have a larger set of opportunities upon high school graduation[2].

Despite large research in the field of vocational education, the impact of vocational coursetaking on educational outcomes is not consensual concerning dropout propensities. This is however an important question as research as proved that students who have not completed high school are more frequently unemployed (Agodini and Deke, 2004). The negative effect on access to higher education is, on the other hand, typically found across the literature. These findings are robust for Portugal, despite the government's efforts to adjust vocational education programs to prepare students with a "dual certification".

The current study attempts to answer two central questions: i) who are the students currently enrolling in vocational programs and ii) how does this decision impact on future educational outcomes? By extending previous literature's analysis to subgroups of students, this study sheds new light on the heterogeneous effect of such coursetaking across different ability groups.

The study is organized as follows. Section 2 surveys the relevant literature. Section 3 provides a brief descriptive analysis of the dataset. Section 4 explains the econometric models and presents the main findings and section 5 concludes the report, suggesting avenues for further research.

[2] The signature of Memorandum of Understanding in May, 2011, compels the Portuguese government to "address early school leaving and improve the quality of secondary education and vocational education and training, with a view to raise the quality of human capital and facilitate labour market matching".

2. Literature Review

Literature on vocational education focuses on two strands of investigation: individuals' returns to vocational education and the factors influencing students' decision to enroll in different curricular tracks (known as tracking research). Within the literature on the consequences of vocational coursetaking, scholars have paid special attention to its effects on later educational outcomes, such as dropout and college enrollment rates. In this line of investigation, a common concern across all studies regards the extent to which the researcher is able to properly isolate the effect that vocational education has on such educational outcomes. For instance, the decision to enroll in vocational programs is not random (and in some countries it may not even be students' choice[3]). Determinants like past school performance, parental education, family's financial conditions and student's ambitions may influence both the student's decision to enroll in vocational programs and the propensities to dropout and attending college. Therefore, not adjusting for such selection-effects may deceive the internal validity of the study. A quite popular approach in this literature is the addition of an extended set of confounding variables in the model. Tracking research emerges in this context, where scholars have been especially interested in determining whether academic ability or socioeconomic status plays a more important role in track placement.

2.1 Effect on college enrollment and dropout rates

In the US, several studies regarding these topics have been performed. Whereas the effect on college enrollment probabilities seems to converge to similar findings across the literature, the impact on dropout rates is not so clear. Some studies that have evaluated the effect on dropout propensities have found associations with reductions in the dropout propensities (Rasinski and

[3] In countries that adopt a formal means of *tracking* education, this decision is imposed by teachers or the school system, depending on students' past school performance.

Pedlow, 1998; Plank, 2001), while others have found insignificant effects (Agodini and Deke, 2004; Mamede et al., 2015) and even increases in this probability (Ainsworth and Roscigno, 2005; Ramos, 2017). Using data from NELS 1988[4], Rasinski and Pedlow (1998) predict an overall negative effect of vocational coursetaking on the propensity to dropout. The authors use logistic regression models and a variety of covariates of students' characteristics and preferences, socioeconomic environment, past school performance and data on school and class settings to estimate this effect. They find, however, that when class rank is included as a control, this effect becomes insignificant. As a result, the authors conclude that vocational coursetaking may have an indirect effect on dropout propensity, but that this effect is mediated by performance[5]. Using the same database, Agodini and Deke (2004) extend the latter analysis to study whether this effect persists when specific subgroups of students are analysed, mainly the educationally and economically disadvantaged individuals. Nevertheless, they find an insignificant effect of vocational coursetaking on this outcome. Contrasting with the previous literature, Ainsworth and Roscigno (2005) find evidence of positive effects of vocational coursetaking on dropout. These authors use also the NELS 1988 database, and logistic regression models to estimate this effect. The difference is in the way they measure the participation in vocational programs. While the two previous papers used a ratio of CTE[6]-to-academic credits, Ainsworth et al. (2005) used binary variables for each type of vocational program (Agricultural, Blue Collar and Low-Wage Service), where the academic program is the omitted category, and found that participating in a Blue Collar vocational course increases the probability of dropping out[7].

[4] National Education Longitudinal Survey (NELS) is a commonly used database across the prevalent literature on educational outcomes in the US, as it is usually seen as a nationally representative sample of students.
[5] The findings in this literature motivate the separate analysis on low and high-ability students employed in the current work. The results on this analysis corroborate the findings in Rasinski and Pedlow (1998).
[6] Career and Technical Education (CTE).
[7] The remaining vocational programs' effect on dropout rates are not statistically significant.

Concerning the effect on college enrollment propensities, literature typically finds a negative effect of participating in vocational training programs (Vanfossen et al., 1987; Arum and Shavit, 1995; Ainsworth and Roscigno, 2005; DeLuca et al., 2006; Ramos, 2017). On this matter, Ainsworth and Roscigno, (2005) finds that being allocated to either a Blue Collar or Agricultural vocational program reduces the probability of attending college. DeLuca et al. (2006) uses a ratio of CTE-to-academic credits and find that, for students who took more than half of their credits in vocational courses the odds of attending college are about 80% lower.

2.2 Previous investigation in Portugal

Only recently literature on this field of study is starting to emerge in Portugal, and while they seem to agree that vocational education has a positive effect on students' grade transition, the effects on dropout rates appear to be poorly significant. Mamede et al. (2015) performed a counterfactual analysis to study the impact of vocational programs in several academic progression measures, as well as on the subsequent access to higher education and on employability, by tracking the students' trajectory along high school. They find a positive effect on grade transition, a negative impact on access to higher education and a null impact on dropout rates. Ramos (2017) extends the previous work by including students' college enrollment expectations as an important covariate. The author finds that vocational courses are not only unsuccessful at preventing early-school leavings as they also inhibit college enrollments. He uses both logistic regression models and propensity score matching approaches to estimate these effects. The main difference between both studies is in the variety of covariates used. Whereas Mamede et al. (2015) uses a large set of controls, Ramos, (2017) excludes some important confounders from his study such as parents' education and previous grade retentions. However, the inclusion of an extended set of confounding variables is crucial to avoid omitted variable endogeneity, and as we will see further in this study, previous retentions and parent's education are two of the most important predictors of both models predicting vocational coursetaking and educational outcomes. Hence, including them as controls in the model estimating the treatment effect of vocational coursetaking is essential[8].

Furthermore, there is a lack of literature concerning the factors that influence students' enrollment decisions in Portugal. Without a clear understanding of whom are the students participating at higher rates in each curriculum track, it becomes especially difficult to realise where educational policies regarding vocational programs are most needed. Hence, an analysis concerning the most important predictors of vocational coursetaking is employed in this study with two main goals: i) to find the main determinants of vocational enrollments, and use them to isolate the treatment effect on educational outcomes, and ii) to shed some light on the students' educational decision factors.

2.3 Social determinants of education decisions

Literature on the determinants of educational choices has highlighted preferences, ability, parental education and financial conditions as the main critical factors of educational choices. Personal preference appears to be the most important determinant of track choice (Jencks, 1972) and some scholars include college plans as a proxy to control for such preferences (Ainsworth et al., 2005; Ramos, 2017). The next most important determinant of curriculum placement is academic ability. Jencks (1972) finds a correlation between test scores and curriculum assignment of around 0.50. On this matter, Spence (1973) suggests that low-ability students have a higher perceived psychological cost when opting for an academic program since it is typically more difficult. Socioeconomic background seems to be the most controversial aspect across Tracking Research. Heyns (1974) found that a student whose father's education is 1 standard deviation above the average is more likely to enroll in academic programs, where

[8] This work includes all covariates used by Ramos (2017), except a dummy indicating whether the school attended by the student is private or a public institution, and includes other important exogenous variables, such as parents' education and previous grade retentions.

Ainsworth et al. (2005) and Deluca et al. (2006) find significant class disparities, even after controlling for past school performances.

This work contributes to complement the prior studies performed in Portugal. Not only do we analyse the factors that influence a student's decision to enroll in different curricular tracks, as we also extend the analyses to subgroups of students. Hence, by analysing low and high-achievers' groups separately, we expect that high-ability vocational students would face higher propensities to enroll in college than low-ability ones, and that the probability of dropping out from school would be low, particularly for the low-performance group.

3. Data and Methodology

To answer the previous questions, a cohort of students from academic and vocational tracks, which have enrolled in the 10th grade on the school year of 2010/11, is followed up until 2014/2015. In regular conditions, a student is expected to have enrolled in college or being employed by the school year of 2013/2014. The follow-up period is however extended for two additional academic years in order to allow to evaluate the educational outcomes of both students that finished high school in the regular scheduling, and of those that have failed a school year, reporting weaker aptitudes towards school.

The dataset used to conduct this study results from the combination of 5 databases: *MISI*, *INQ-PRIV*, *ENEB*, *OTES* and *CNAES*. The first two databases, *MISI* and *INQ-PRIV* contain an extensive set of information of the population of students in the Portuguese Education System from both public and private institutions, respectively, making available data on student's characteristics and socioeconomic context, since 2007. *ENEB*[9] records student's performance

[9] *Exames Nacionais do Ensino Básico* (Basic Education National Exams).

on the 9th grade national exams. This study uses both scores from the exams of Portuguese and Mathematics as a proxy of student's previous knowledge and ability: along with the particularity of being equal and compulsory for all students, researchers have found that the results of typical IQ tests and standardized tests are often highly correlated (Rinderman, 2007). Hence, national exams' scores seem to establish an effective measure of students' cognitive ability. The college enrollment variable is built using *CNAES*[10], a database that aggregates information on students' college application choices. The richness of the dataset used in this work is largely due to the data collected from the survey "*OTES*"[11], which was conducted at national level, on 67043 students from 748 private and public schools of Portugal[12]. During the process of collecting information, each of the respondents is surveyed at three key moments of their school journey: The first, when the students were on the 10th grade, the second when they were concluding the 12th grade and a final one applied via email 14 months after the expected date for completion of high school. The current study uses predominantly the answers collected from the first stage, where the respondents are questioned about several aspects that are not obtainable from the previous databases, such as parental education[13], the household they live in and their educational and job expectations upon high school graduation[14].

3.1 Methodology

The analysis proceed in four steps. First, I present a descriptive analysis of the sample of students, in order to identify the main differences between the academic and vocational track groups' composition, regarding: individual characteristics, socioeconomic context, schooling

---

[10] *Concurso Nacional de Acesso ao Ensino Superior* (National Competition for Access to Higher Education).

[11] *Observatório de Trajetos dos Estudantes do Ensino Secundário (*Observatory of the Secondary Education Student's Trajectories). This survey was conducted by the DGEEC, with the objective to collect and disseminate statistical information about the school and professional paths of secondary school students.

[12] With exception from Azores and Madeira.

[13] MISI has information on parents' education but only for public schools' students.

[14] As we will see later in this study, the addition of these variables is essential to infer causality in model 2, as they seem to influence both the dependent and independent variables (vocational coursetaking) in this model.

experience and student's educational prospects. The differences found in these statistics further motivate the empirical study, which employs both Linear Probability and Logistic Regression Models to measure whether and by how much the above-mentioned factors contribute to the student's educational decisions. In a third step, I reproduce the prior analysis to the bottom and top performance's students, aiming at investigating whether the previous findings change for these two groups of students. Finally, I estimate the implications of such participation in vocational programs on ultimate college enrollment and dropout rates. To check the obtained results, a Propensity Score Matching approach is employed after the logistic regression.

3.2 Descriptive analysis

Table 1 (in appendix) reports summary statistics of individual characteristics of the cohort of students previously defined. The dataset is composed by 44919 observations[15], where 75% of the students were enrolled in an academic program while the remaining 25% were in a vocational one. Going through the statistics, one can observe large disparities between the two groups. First, the portion of females is substantially larger in the academic track, an outcome commonly identified in literature (DeLuca et al, 2006). In the academic track we find 58% of female students, whereas in the vocational track this percentage is nearly 47%. The average student age is higher in the latter group, which most likely reflect the large observed percentage of students that have ever failed a school year, (which is the majority of the vocational students, 52%, in contrast to 16% in the general track). Parental education is measured as the highest educational degree achieved by both parents[16]. It is gathered into 3 subcategories: primary education[17], high school and college. The portion of vocational students whose parents have

[15] The number of students in this sample differs from the number of respondents of the survey, due to lack of reported answers for some of the questions covered by the survey, or due to mismatches with the above mentioned databases.
[16]These data is independent of whether the parent is or not biological. Tutors and step-parents educational grades are considered when best suitable in regard of student's household composition.
[17] This category also covers the cases where both parents/tutors have not attended school, since the number of students in this situation was not significant to conduct analysis.

primary education[18] as the highest educational attainment is nearly 70%, an extremely large portion when compared to 45% among academic students. As expected, the former group also detains the lowest percentage of parents holding a superior education degree (8% in contrast to 26% for academic track youths). The academic track reports the highest percentage of students living in nuclear families, whereas vocational youths report the highest portions of students living in a single-parent family, stepfamily and other arrangements. Regarding previous school performance, academic students report higher average scores on both national exams of Portuguese and Mathematics. Perhaps one of the most surprising numbers, is the high percentage of vocational students seeking to attend college. This sums up to nearly 44%, which is almost half of the students in the sample enrolled in vocational programs. This percentage is, as predictable, much higher in academic tracks (85%).

4. Empirical model and results

Throughout the paper, two models are estimated: Model 1 analysing student's program decisions and a Model 2 evaluating the impact of vocational programs' choice on college enrollment and dropout rates.[19]

4.1 Which factors influence students to enroll in different curricular tracks?

The econometric model predicting the likelihood that a student participates in a secondary vocational program is the following:

$$\underline{Model\ 1}$$

$$\begin{aligned} voctrack_i = \ & \beta_0 + \beta_1 female_i + \beta_2 portuguese_i + \beta_3 subsidy_i + \beta_4 internet_i + \delta_1 parenteduc_i \\ & + \delta_2 household + \beta_{10} examscorePT_i + \beta_{11} examscoreMATH_i + \beta_{12} retention_i \\ & + \beta_{13} expectations_i + \varepsilon_i \end{aligned}$$

[18]In Portugal, elementary education is divided in three sequential cycles: first cycle stands from 1st to the 4th grade, the second cycle covers the 5th and 6th, and the 3rd cycle (or lower secondary education) extends from the 7th to the 9th grade.

[19] All models were estimated with region dummies (NUTS III), to control for possible regional effects, namely, the offer and diversity of vocational education programs in local high schools.

The dependent variable, $voctrack$ is a binary variable equal to 1 if the student is enrolled in a vocational program in his 10th grade, and 0 if he is enrolled in an academic one. Table 2 reports the full set of covariates included in models 1 and 2, which were chosen based on previous literature's findings, theoretical relevance and statistical significance.

Table 2. Definition of the full set of explanatory variables included in models 1 and 2.

Exogenous variables	Definition
female	Dummy variable equal to 1 if the student is female and 0 if male.
portuguese	Dummy variable equal to 1 if the student has Portuguese nationality and 0 if not.
subsidy	Dummy variable equal to 1 if student receives school subsidy and 0 otherwise.
internet	Dummy variable equal to 1 if the student has access to internet and 0 if he does not.
parentaleduc: (basiceduc), second_educ, higher_educ	Composed by 3 dummy variables: (*basic_educ)*, *second_educ* and *higher_educ* if the parent's maximum education completed is, respectively, the primary education (including the cases where parents have not frequented school at all), secondary education or higher education degrees. *basiceduc* is the omitted category.
household: (hh1), hh2, hh3, hh4	Composed by 4 dummy variables: (*hh1)*, *hh2*, *hh3* and *hh4*, depending on whether the student lives in a nuclear family, single-parent family, stepfamily or another situation, respectively. The omitted category is living in a nuclear family, hh1.
examscorePT	Continuous variable (ranging from 0 to 100), reporting student's score on the 9th grade national exam of Portuguese.
examscoreMATH	Continuous variable (ranging from 0 to 100), reporting student's score on the 9th grade national exam of Mathematics.
retentions	Dummy variable equal to 1 if the student has ever repeated a school year and 0 otherwise
expectations	Dummy variable equal to 1 if the student has expectations of pursuing college studies, at the beginning of his secondary school program, and 0 otherwise.

4.1.1 Model estimation

The model is first estimated by OLS. Despite the commonly attributed weaknesses of the OLS estimator in predicting probabilities, (for instance, it does not restrict the probabilities to the interval [0,1] and the variance is always heteroskedastic, by construction), it is often used by econometricians as a preliminary exploratory tool, since "it provides a good guide to which variables are statistically significant" (Cameron and Trivedi, 2005). To further investigate by *how much* each of these determinants affect the probability of observing a student participating in vocational studies, the same model is estimated by logistic regression[20]. These estimations

[20] Cameron & Trivedi, (2005) suggest that the advantages brought up by this estimation methodology, (not only probabilities are constrained the interval [0,1] but also estimates resultant from this method can be used for

are presented as the average marginal effects (AME) of each explanatory variable on the probability that a student will attend a vocational program[21]. Table 3 exhibits these results.

Consistent with the expectation that vocational education choice presents dissimilarities between genders (Brown et al., 1997; Grebennikov, 2009; Lubinski, 1992), the negative coefficient of $female$ in both estimation methods suggests that female students are, on average, 6.6 percentage points less likely to choose a vocational path than males (column 2). Both model estimation methods show no statistical significance of variables $portuguese$ and $subsidy$, (however, when we add college expectations[22] as a control, $portuguese$ turns out to be significant). $Internet$ has a significant and negative marginal effect. Hence, it appears to exist a positive link between material deprivation (which may point towards lower socioeconomic status) and a higher likelihood to participate in non-academic tracks. Moreover, and in accordance to prior literature, parental education reports a negative sign, suggesting that students whose parents hold degrees in higher or secondary education levels are more likely to follow the academic trajectory[23]. Hence, students whose parents' highest educational degree is secondary or higher education are, respectively, 5.7 p.p. and 11.4 p.p. less likely to enroll in vocational tracks, compared to students whose parents' highest educational attainments are primary education.

An interesting result emerges when we observe the relationship between family structure and participation on vocational education. For instance, students living in single-parent,

---

predictions about the dependent variable) make this a better method (compared to a standard OLS regression) for final data analysis.

[21] Differently from the marginal effect at the mean (MEM) which assumes all the other exogenous variables in their mean values (therefore assuming there is a typical "average student to whom the effects are shown in comparison), in the AME method, the marginal effect is first calculated for each individual with their observed covariates' values. These values are then averaged across all individuals in the sample.

[22] The effect of college expectations on the remaining covariates is further analysed in section 4.1.2.

[23] The amount of students graduating from secondary school was of nearly one third of what it is today (PORDATA). Thus, one may think of such a graduation as a determinant for medium/higher socioeconomic status. Therefore, it is likely that this variable carries a negative coefficient, thus predicting a higher likelihood that a student will follow the academic path.

stepfamilies or any other non-nuclear arrangement, seem to predict higher propensities to participate in vocational studies. This supports the literature findings regarding the influence of family structures on youths' educational choices, which have shown that non-nuclear families shape educational outcomes due to turmoil or changes in family structure and diminished resources (Sandefur, McLanahan and Wojtkiewicz, 1992; Downey, 1995b). Amongst these settings, leaving in a stepfamily seems to report the highest marginal effect (3.3 p.p., in contrast to 1 p.p. for single-parent families and 3.1 p.p for other non-nuclear households).

Concerning school behaviours and prior achievements, the model predicts that scoring higher in the national exams is associated with reductions in the odds that a student will follow a vocational trajectory. Plus, students that have already repeated a school year are, on average, 15 p.p. more likely to enroll in a vocational program. These two findings are in accordance to expectations. Indeed, vocational programs were originally designed to help engaging students reporting risk of dropout. Thus, one would expect students reporting prior retentions and lower grades to be more likely to enroll in these programs.

4.1.2 Including college expectations as a control[24]

Regressions in columns (1a) and (2a) include students' expectations regarding college studies as a control. The marginal effect of this covariate is very high and negative on the probability of participating in vocational education. On average, a student that expects to pursue college studies is 16.6 p.p. less likely to enroll in a vocational program. In this respect, DGEEC (2016) explains that "given the differences in focus between the two types of education, it is natural for most of the young people who, by the end of the 9th grade choose to enroll in an academic

[24] When estimating Model 2 a similar approach will be used, where college expectations are added in parallel way, to evaluate the change of the impact of the treatment effect (participating in vocational education) on college enrollment and school dropouts. This approach is used in a similar way as in Ramos (2017).

course, already have a firm intention of enrolling in higher education in the humanistic scientific (academic) programs".

Table 3. Results for model 1 predicting participation in vocational track. Regressions in columns (1) and (1a) were estimated with robust standard errors.

	(1) OLS	(1a) OLS	(2) Logit	(2a) Logit
female	-0.0654***	-0.0378***	-0.0660***	-0.0401***
	(0.00379)	(0.00370)	(0.00362)	(0.00356)
portuguese	-0.0149	-0.0235**	-0.00797	-0.0158*
	(0.0111)	(0.0107)	(0.00881)	(0.00847)
sase	0.000288	-0.00373	0.00394	-0.000406
	(0.00517)	(0.00497)	(0.00447)	(0.00434)
internet	-0.0534***	-0.0258**	-0.0362***	-0.0167**
	(0.0113)	(0.0109)	(0.00828)	(0.00802)
second_Educ	-0.0640***	-0.0346***	-0.0565***	-0.0313***
	(0.00450)	(0.00439)	(0.00426)	(0.00416)
higher_Educ	-0.0797***	-0.0445***	-0.114***	-0.0750***
	(0.00460)	(0.00450)	(0.00598)	(0.00580)
household_2	0.00730	0.00846	0.00999*	0.0103**
	(0.00556)	(0.00539)	(0.00519)	(0.00501)
household_3	0.0277***	0.0223**	0.0334***	0.0269***
	(0.00907)	(0.00887)	(0.00765)	(0.00735)
household_4	0.0282*	0.0224	0.0308***	0.0247**
	(0.0146)	(0.0141)	(0.0113)	(0.0110)
examscore_PT	-0.00436***	-0.00336***	-0.00485***	-0.00390***
	(0.000158)	(0.000154)	(0.000149)	(0.000146)
examscore_Math	-0.00307***	-0.00232***	-0.00327***	-0.00257***
	(0.000110)	(0.000107)	(0.000100)	(9.85e-05)
retentions	0.228***	0.193***	0.149***	0.127***
	(0.00565)	(0.00559)	(0.00358)	(0.00354)
expectations	-	-0.253***	-	-0.166***
	-	(0.00550)	-	(0.00330)
Observations	44919	44919	44,919	44,919
R^2 / Pseudo-R^2	0.234	0.286	0.227	0.268

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Moreover, when *expectations* is included as a control, interesting dynamics emerge across the remaining covariates' marginal effects: the variable *portuguese* becomes significant, which suggests that among the students pursuing postsecondary studies, the non-Portuguese are more likely to pursue vocational tracks. For instance, if the non-natives face difficulties in fluently speaking Portuguese, they may try to avoid the (more difficult) academic programs[25]. Another suggestive effect of adding college expectations is the variation in the magnitude of parental education's coefficients. The resulting lower statistic suggests that student's expectations regarding college enrollment were formerly reflected in parent's education variable. In this respect, Munk (2011) suggests that parents with higher degrees of education may encourage his child to follow a similar trajectory. This relationship has also been addressed by Wentzel, (1998), which concludes that perceived social and emotional support from parents is linked to higher academic effort and engagement in school.

4.1.3 How do the previous findings change when we analyse the bottom and top performance's students?

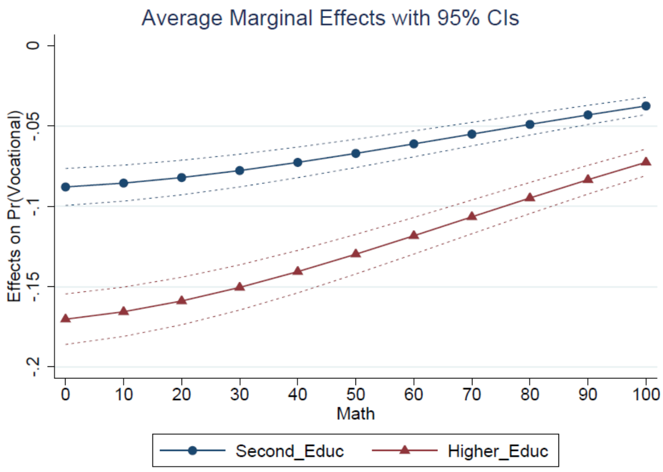
An interesting outline arises when we attempt to analyse the difference in the impact of the previous covariates *for different students' exams scores*, on the probability of choosing vocational programs. What we observe is a heterogeneous effect of socioeconomic context across different ability levels. This means two things: 1) the factors that influence a high-ability student's decision to enroll in a vocational program probably differ from the ones that influence a low-ability student's decision, and 2) the same exogenous factors that influence such coursetaking decision might as well predict students' later educational outcomes, differently for both ability groups.

Figure 1 exhibits the AME of parental education across different score levels in the national exam of Mathematics. We verify that for lower scores there is a higher effect of parental education (measured by both $second\_educ$ and $higher\_educ$, in contrast to the omitted category $basic\_educ$), and a lower effect for higher scores[26]. Hence, for students scoring below

[25] In the US, Natriello et al., (1990) finds that students with limited English proficiency are more likely to be among the group of lower-achieving students and more likely to dropout.
[26] For the sake of graphical representation simplicity, parental education is the only covariate whose effect is presented in figure 1, which is also the one whose heterogeneous effect is the largest across students' exam

50, the marginal effect of a parent holding college studies ($higher\_educ$) on the probability of pursuing vocational programs has the greatest magnitude (-0.17), whereas for higher scores, this effect is reduced to -0.07.[27] A possible explanation is that students whose parents hold higher degrees of education may be encouraged to follow an academic trajectory, as their parents perceive this as the most prosperous track, disregarding their lower aptitudes towards a more study-intensive program. The purpose of this graphical analysis is to aware of the presence of a heterogeneous influence of the student's socioeconomic context on one's educational decisions that disregards the student's ability.

Figure 1. AME of parental education across different score levels of Mathematics National Exam.



Following this analysis, a new approach to estimate Model 1 is employed. The previous sample is narrowed into 2 subgroups: the lower and the higher-performance groups, where the first contains the students who have *failed*[28] in both national exams, and the second contains the ones that scored *above 60* in both of them[29]. Both OLS and logistic regression methods were conducted to estimate the marginal effects of each of the covariates.

score levels. Regarding the remaining variables, female, internet and college expectations report a similar pattern to the one represented by parental education, whereas household composition and school retentions exhibit an opposite pattern. Moreover, this graphical analysis serves merely as a motivation for the analysis carried after.

[27] Although producing lower estimates, a similar pattern is found when observing the AME of $secondeduc$.

[28] Failing an exam is equivalent to scoring strictly below 50.

[29] Table 1a) in appendix summarizes the statistics for each of these subgroups.

Table 4 reports the results for this analysis. Amongst the low-performance group, the former analysis is corroborated: previous school retentions and the lack of expectations regarding future college studies are the best predictors of participation in vocational programs. Nevertheless, parental education also predicts a large effect on the probability of participating in vocational education. Thus, disregarding the cognitive ability of the student, there seems to be an influence by parents holding higher degrees of education for the student to follow an academic program.

Within the high-performance group the effects that stand out are family structure, access to internet and, parental education (although carrying a significantly lower marginal effect than the one among the low-ability group)[30]. Particularly high-performance students living in a stepfamily face, on average, higher probabilities of participating in vocational programs. This effect is constant even after controlling for *college expectations*. Sandefur et al. (1992) explains that, although remarriage may improve a family's economic conditions, the stepparent may have obligations to children outside the household, hence restraining the available income. Even without such obligations, his willingness to share his income with the children he lives with may be reduced. Moreover, if we assume that the access to internet is a good proxy to identify students from lower socioeconomic contexts, we may establish that high performance students might be more likely to enroll in vocational courses if they face some kind of financial restraints (i.e., their families may not be able to bear the college education expenses). As a result, the student may opt to enroll in a vocational program, which he perceives as the best alternative to gain some entry-level occupational skills and enter the job market. Another curious result is the lack of significance of the variable $female$ in this group. It seems that

[30] Although the coefficient of retentions variable also carries a significant coefficient, this is not a surprising effect as it is possible that a low-ability student achieves a high mark on a national exam, if for instance he has already repeated that same exam because he could not succeed to pass at first. Therefore, retentions appear in this regression more as a control for student's "true" ability, than as a variable with explanatory power.

gender only predicts educational decisions when the student is a lower-achiever, which in this case we observe that females are less likely to follow a vocational program.

Table 4. Results for model 1 predicting participation in vocational track, separately for high and low performance students.

	High-performance group				Low-performance group			
	(3) OLS	(3a) OLS	(4) Logit	(4a) Logit	(5) OLS	(5a) OLS	(6) Logit	(6a) Logit
female	-0.00885**	-0.00130	-0.00851**	-0.00119	-0.113***	-0.0710***	-0.112***	-0.0705***
	(0.00394)	(0.00384)	(0.00384)	(0.00379)	(0.0110)	(0.0109)	(0.0107)	(0.0106)
portuguese	-0.0135	-0.0157	-0.0101	-0.0150	-0.00888	-0.0192	-0.00976	-0.0199
	(0.0179)	(0.0175)	(0.0121)	(0.0116)	(0.0228)	(0.0224)	(0.0226)	(0.0218)
subsidy	0.00281	0.00166	0.00282	0.00119	0.0106	0.00389	0.0107	0.00379
	(0.00678)	(0.00655)	(0.00524)	(0.00517)	(0.0137)	(0.0132)	(0.0135)	(0.0131)
internet	-0.0642***	-0.0439**	-0.0386***	-0.0249**	-0.0309	-0.00849	-0.0314	-0.00870
	(0.0229)	(0.0222)	(0.0103)	(0.0103)	(0.0228)	(0.0222)	(0.0227)	(0.0220)
second_educ	-0.0180***	-0.0100*	-0.0151***	-0.00980**	-0.075***	-0.0401***	-0.074***	-0.0391***
	(0.00559)	(0.00540)	(0.00467)	(0.00459)	(0.0133)	(0.0131)	(0.0131)	(0.0129)
higher_educ	-0.0393***	-0.0259***	-0.0423***	-0.0307***	-0.176***	-0.109***	-0.177***	-0.109***
	(0.00508)	(0.00480)	(0.00524)	(0.00509)	(0.0204)	(0.0202)	(0.0211)	(0.0207)
household_2	0.00308	0.00230	0.00261	0.000867	0.000595	0.00340	0.000531	0.00324
	(0.00632)	(0.00619)	(0.00586)	(0.00577)	(0.0162)	(0.0157)	(0.0161)	(0.0155)
household_3	0.0538***	0.0498***	0.0375***	0.0338***	0.0390*	0.0400*	0.0392*	0.0397*
	(0.0148)	(0.0143)	(0.00778)	(0.00759)	(0.0230)	(0.0227)	(0.0230)	(0.0222)
household_4	-0.00153	-0.00228	0.000117	-0.00186	0.00531	-0.00547	0.00585	-0.00490
	(0.0217)	(0.0219)	(0.0153)	(0.0152)	(0.0318)	(0.0308)	(0.0321)	(0.0310)
retentions	0.230***	0.192***	0.0884***	0.0681***	0.224***	0.195***	0.215***	0.187***
	(0.0218)	(0.0211)	(0.0057)	(0.0057)	(0.0112)	(0.0111)	(0.0100)	(0.0099)
expectations	-	-0.173***	-	-0.0750***	-	-0.247***	-	-0.229***
	-	(0.0144)	-	(0.0047)	-	(0.0115)	-	(0.0096)
Observations	12,472	12,472	12,472	12,472	7,584	7,584	7,584	7,584
R^2 / Pseudo-R^2	0.062	0.100	0.108	0.155	0.101	0.156	0.076	0.119

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.2 Impact of vocational education participation on future educational outcomes

With a basic understanding of whom participates in academic and vocational programs, the question now turns to whether (and how) such participation impacts on later educational outcomes, namely: dropout and college enrollment. From the previous analysis we conclude that the participation in vocational education is not random. There are some determinants affecting its likelihood and it is very likely that these same determinants also influence the student's odds of dropping out and of enrolling into college. We have also verified that although the majority of students in vocational programs are low-performance students, there is still a considerable amount of poor performance students enrolling in academic programs, as well as a portion of high performance students enrolling in vocational ones. To adjust for possible selection effects in Model 2, we use the previous model's covariates as controls. Again, a logistic regression model is used to estimate the impact of participating on vocational courses separately for lower and higher-performance subgroups:

$$\underline{Model\ 2}:$$

$$y_i = \beta_0 + \delta voctrack_i + \beta X's + \varepsilon_i$$

The dependent variables, college enrollment and dropout, are represented by y_i in this model. X denotes the full set of covariates used in Model 1 to predict the main determinants of students' program choice. The treatment variable, $voctrack$, is equal to 1 if the student has enrolled in a secondary vocational program and 0 if in an academic one. Table 5 presents the average marginal effects of this econometric model, where columns (1) to (2a) stand for the high performance group and columns (3) to (4a) stand for the low ability group.

4.3.1 College enrollment

The negative estimates of $voctrack$ in regressions (1) and (3) predict that, for both subgroups, vocational education reduces the chances of students enrolling in college. Hence, a high performance student and a low performance student that have studied a vocational program are respectively, on average, 20.8 p.p. and 17.5 p.p. less likely to enroll in college, in contrast to the ones that have studied an academic program. When college expectations are added to the model, one would expect that, especially for the high performance group, this effect would be strongly diminished. However, in both groups, the odds are poorly changed by about 2.5 p.p. These findings replicate previous research outcomes (Ramos, 2017; DeLuca et al, 2006; Ainsworth and Roscigno, 2005).

Table 5. Results of model 2 predicting college enrolment and dropout probabilities for high and low performance students. The results shown are the AME estimated by a logistic regression model.

	High-performance group				Low-performance group			
	(1) College	(1a) College	(2) Dropout	(2a) Dropout	(3) College	(3a) College	(4) Dropout	(4a) Dropout
voctrack	**-0.208*** **	**-0.174*** **	**0.0535*** **	**0.0485*** **	**-0.175*** **	**-0.151*** **	**-0.138*** **	**-0.149*** **
	(0.00939)	(0.00965)	(0.00464)	(0.00469)	(0.00935)	(0.00919)	(0.00949)	(0.00976)
female	0.0334***	0.0221***	-0.00664**	-0.00496	0.0372***	0.0228***	-0.0321***	-0.0248***
	(0.00592)	(0.00583)	(0.00330)	(0.00331)	(0.00706)	(0.00706)	(0.00929)	(0.00940)
portuguese	0.0827***	0.0846***	-0.0224**	-0.0236***	0.0358**	0.0373**	-0.0692***	-0.0713***
	(0.0181)	(0.0174)	(0.00881)	(0.00878)	(0.0172)	(0.0168)	(0.0172)	(0.0171)
subsidy	-0.0146*	-0.0129	0.00103	0.000713	-0.00384	-0.00231	-0.0152	-0.0163
	(0.00824)	(0.00815)	(0.00516)	(0.00517)	(0.00879)	(0.00870)	(0.0117)	(0.0116)
internet	0.0499***	0.0271	-0.0118	-0.00803	0.0537***	0.0443**	-0.0415**	-0.0377**
	(0.0188)	(0.0190)	(0.0107)	(0.0108)	(0.0191)	(0.0188)	(0.0178)	(0.0178)
second_educ	0.0494***	0.0398***	0.00550	0.00760*	0.0459***	0.0364***	0.00178	0.00720
	(0.00737)	(0.00725)	(0.00444)	(0.00448)	(0.00799)	(0.00789)	(0.0111)	(0.0112)
higher_educ	0.0992***	0.0804***	0.0114***	0.0148***	0.0843***	0.0661***	-0.0168	-0.00650
	(0.00766)	(0.00752)	(0.00429)	(0.00439)	(0.0108)	(0.0107)	(0.0175)	(0.0176)
household_2	-0.0314***	-0.0304***	0.00319	0.00277	-0.0175	-0.0179*	0.0436***	0.0440***
	(0.00875)	(0.00856)	(0.00496)	(0.00496)	(0.0107)	(0.0106)	(0.0129)	(0.0129)
household_3	-0.0717***	-0.0679***	0.00807	0.00787	-0.0284*	-0.0262*	0.0286	0.0292
	(0.0136)	(0.0133)	(0.00764)	(0.00760)	(0.0160)	(0.0157)	(0.0186)	(0.0186)
household_4	-0.0639***	-0.0640***	0.0365***	0.0371***	-0.0419*	-0.0354	0.0553**	0.0536**
	(0.0222)	(0.0216)	(0.00960)	(0.00960)	(0.0252)	(0.0247)	(0.0248)	(0.0248)
retentions	-0.179***	-0.155***	0.0342***	0.0284***	-0.0730***	-0.0666***	0.0728***	0.0697***
	(0.0117)	(0.0118)	(0.00571)	(0.00582)	(0.00750)	(0.00738)	(0.00962)	(0.00963)
expectations	-	0.164***	-	-0.0262***	-	0.0941***	-	-0.0465***
	-	(0.00846)	-	(0.00495)	-	(0.00833)	-	(0.00992)
Observations	12,521	12,521	12,472	12,472	7,584	7,584	7,584	7,584
Pseudo-R²	0.128	0.160	0.106	0.113	0.181	0.206	0.050	0.053

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.3.2 Dropout rate

An unexpected result arises when we analyse the marginal effects of v*octrack* on the model predicting dropout propensity. Whereas in the low-performance group we find a *negative impact* of participating in this type of education, within the high-ability group we verify a *positive impact*. Hence, while a low-performance vocational student is, on average, 13.8 p.p. less likely to dropout from school vis-à-vis their counterparts in academic programs, high-performance students are found to be, on average, 5.3 p.p. more likely to dropout from high school. Regressions in columns (2a) and (4a) predict that these results persist even after controlling for college expectations. This variable not only is statistically relevant in this model, but also when included, reduces the likelihood of dropping out (from -13.8 p.p. to -14.9 p.p. and from 5.3 p.p. to 4.9 p.p. respectively for low and high-performance students)[31].

4.3.3 Robustness check

As we have already proved in the first part of analysis in the current study, treatment selection in Model 2 lacks randomization. Hence, even after controlling for student's ability, socioeconomic background and student's characteristics, there might still exist a bias due to unobservable factors that, theoretically, may influence both the educational choices and later educational outcomes. Therefore, with the aim of checking the robustness of the previous findings, a Propensity Score Matching approach is further employed. This method has the advantage to match similar individuals and estimate the treatment effect based on the differences found in the outcome observables. Tables 6 and 7 (in appendix) report the resultant average treatment effect of participating in a vocational program on both college enrollment and dropout rates, separately for low and high-performance students. The results show quite similar results to the ones found in logistic regression models, suggesting that the previous results seem to be robust.

5. Discussion and conclusions

Vocational education has been, for the last decade, subject to intense debate in the education policy in Portugal. Recent policies addressing this type of educational programs are focused on increasing the number of high school graduates and in expanding the set of postsecondary

[31] Although the focus of analysis are the subgroups of students in question, an overall analysis of the sample of also employed. Both negative effects on dropout and college enrollment rates were found. Table 5.a) in appendix shows these results.

opportunities for vocational students. For policy-making purposes it is, however, crucial that government's actions are supported by research evidence, since without a clear understanding of the impact of these actions, any following investments may be inefficiently addressed. Previous literature (Mamede et al. 2015, Ramos, 2017) has addressed this subject in Portugal, but there is not an overall consensus on the impact of this type of education, as it is typically found to have negligible effects.

The novelty of this study is in the approach used to analyse the impact of the vocational effect, which is employed separately for low and high-achievers. The main finding from this analysis is that the impact of choosing a vocational program is not the same for all students. There is a heterogeneity of the effect of this program choice on dropout probabilities, which increases for high-achievers and decreases substantially for low-achievers. Hence, whereas vocational education appears to be successful in engaging students at risk, policy-makers should be aware of the impact of such programs for the high-performance students.

From the previous analysis is not possible to clearly derive the reasons that underlie this outcome. Additional research is needed to understand why this group of students is found to be particularly more likely of dropping out if they participate in vocational programs. Besides, a limitation of this study is the distortion in the proportion of students allocated to vocational and academic programs within the high-performance group of students. This may indicate that high-achievers in vocational programs do not represent a matching counterpart to high-performance academic students, which if true, may be biasing the results for the high-performance group.

Finally, although the effect of vocational education on college enrollment odds was shown to be negative and higher for the high-ability group, one would expect a smaller effect particularly for this subgroup. Hence, additional research should also focus on the reasons why vocational programs have consistently negative impacts on college enrollments.

References

Ainsworth, James W. and Roscigno, Vincent J. 2005. "Stratification, School-Work Linkages and Vocational Education". Social forces, 84(1): 257-284.

Agodini, Roberto and Deke, John. 2004. The Relationship between high school vocational education and dropping out. Mathematica. Mathematica Policy Research, Inc.

Arum, Richard and Shavit, Yossi. 1995. "Secondary Vocational Education and the Transition from School to Work". Sociology of Education, 68 (3): 187-204.

Brown, Charles and Corcoran, Mary. 1997. "Sex-Based Differences in School Content and the Male-Female Wage Gap". Journal of Labor Economics, 15(3): 431-465.

Cameron, A. Colin and Trivedi, Pravin K. 2005. Microeconometrics: Methods and Applicatons. New York: Cambridge Univesity Press.

Cedefop. 2011. "The economic benefits of VET for individuals". Luxembourg: Publications Office. http://www.cedefop.europa.eu/files/5511_en.pdf.

DeLuca, Stephanie, Plank, Stephen and Estacion, Angela. 2006. "Does career and technical education affect college enrollment?". National Research Center for Career and Technical Education, University of Minnesota.

Downey, Douglas. 1995b. "Understanding Academic Achievement among Children in Stephouseholds: The Role of Parental Resources, Sex of Stepparent, and Sex of Child." Social Forces, 73 (3): 875-894.

Grebennikov, Leonid and Skaines, Ivan. 2009. "Gender and higher education experience: A case study". Higher Education Research & Development, 28(1): 71–84.

Heyns, Barbara. 1974. "Social Selection and Stratification within Schools." American Journal of Sociology, 79: 1434-1451.

Jencks, Christopher, Smith, Marshall, Ackland, Henry, Bane, Mary Joe, Cohen, David, Gintis, Herbert, Heyns, Barbara and Mickelson, Stephen. 1972. Inequality: A Reassessment of the Effect of Family and School in America. New York: Basic Books.

Kulik, James. 1998. "Curricular Tracks and High school Vocational Education". In The Quality of Vocational Education: Background Papers from the 1994 National Assessment of Vocational: 65-131.

Lubinski, David and Benbow, Camilla P. 1992. "Gender differences in abilities and preferences among the gifted: Implications for the math/science pipeline". Current Directions in Psychological Science, 1: 61-66.

Lucas, Samuel. 1999. Tracking inequality. New York: Teachers College Press.

Mamede, Ricardo Paes, Cruz, Daniela. and Fernandes, Teresa. 2015. “CIE of Vocational Training and Traineeships”. European Social Fund.

Munk, Martin D. 2011. "Educational Choice: Which Mechanisms Are at Stake?" Paper presented at RC28, Essex, United Kingdom.

Natriello, Gary, Edward L. McDill, and Aaron M. Pallas. 1990. Schooling Disadvantaged Children: Racing Against Catastrophe. New York: Teachers College Press.

Neuman, Shoshana and Adrian Ziderman. 1999. “Vocational education in Israel: Wage effects of the VocEd-Occupation match.” The Journal of Human Resources, 34(2): 407-420.

Oakes, Jeannie. 1985. Keeping track. New Haven, CT: Yale University Press.

OECD. 2009. “Learning for Jobs: OECD Policy Review of Vocational Education and Training”. OECD Publishing

Plank, Stephen. 2001. “Career and Technical Education in the Balance: An Analysis of High School Persistence, Academic Achievement, and Postsecondary Destinations.” .National Research Center for Career and Technical Education, University of Minnesota.

Ramos, Ângelo. 2017. "The impact of choosing a vocational track on academic performance and college enrollment in Portugal" (Master’s thesis). Nova School of Business and Economics.

Rasinski, Kenneth, and Steven Pedlow. 1994. “Using Transcripts to Study the Effectiveness of Vocational Education.” Journal of Vocational Education Research, 19:23- 43.

Rindermann, Heiner. 2007. “The g-factor of international cognitive ability comparisons: the homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations”. European Journal of Personality, 21(5): 667–706.

Sandefur, Gary, Sara McLanahan and Roger Wojtkiewicz. 1992. "The Effects of Parental Marital Status During Adolescence on High School Graduation." Social Forces, 71: 103-22.

Shavit, Yossi and Müller, Walter. 2000. "Vocational secondary education, tracking, and social stratification". Handbook ofthe Sociology of Education, 19: 437-453.

Spence, Michael. 1973. "Job market signaling". Quarterly Journal of Economics, 87(3): 355-74.

Vanfossen, Beth E., Jones, James D. and Spade, Joan Z. 1987. "Curriculum tracking and status maintenance". Sociology of Education, 60 (2): 104-122.

Wentzel, Kathryn. 1998. "Social Relationships and Motivation in Middle School: The Role of Parents, Teachers and Peers".Journal of Educational Psychology, 90 (2): 202-209.