ISCTE Business School
Instituto Universitário de Lisboa

COUNTERFACTUAL IMPACT EVALUATION OF VOCATIONAL EDUCATION IN PORTUGAL

DANIELA FILIPA FONSECA CRUZ

Thesis submitted as a partial requirement for the conferral of MSc in Economics

Supervisor:
Prof. Dr. Ricardo Mamede, Assistant professor,
School of Social Sciences, Department of Political Economy

October 2015

Abstract

The Reform of the Portuguese education system, which took place in 2004, led to the introduction and rapid expansion of vocational education in public high schools. This thesis aims to assess the impacts of vocational education on students' academic and labor market performance using a Counterfactual Impact Evaluation (CIE) approach. The dataset used in the analysis includes data for the academic years of 2008/2009, 2009/2010 and 2010/2011. For both methodological and data availability reasons, we restrict our analysis to Portuguese students without special educational needs, aged between 15 and 18 years old, who are enrolled in public high schools with tutelage of the Ministry of Education and Science, and that were enrolled in general education in the previous year (basic education). For this group of students, we find that being enrolled in vocational education has positive impacts in school performance (transition, graduation, and dropout rates) and in labor market performance (employment rate, average salary and average worked months per year). We also find that vocational education students have lower chances of proceeding to higher education. These results are robust to variations in the estimation method.

JEL classification:
C14 – Non parametric methods
I26 – Returns to education

Keywords:
Vocational education; Secondary education;
Counterfactual Impact Evaluation; Coarsened Exact Matching;

Resumo

A reforma do sistema de ensino Português, que ocorreu em 2004, originou a introdução e rápida expansão do ensino profissional nas escolas secundárias públicas. Esta tese tem como objetivo avaliar os impactos da educação profissional no desempenho escolar e laboral dos alunos, utilizando a lógica de Avaliação Contrafactual de Impactos (CIE). A base de dados inclui informação dos anos letivos de 2008/2009, 2009/2010 e 2010/2011. Devido a questões metodológicas e de disponibilidade de dados restringimos a nossa análise a estudantes portugueses, sem necessidades educativas especiais, com idades entre os 15 e 18, que estão matriculados em escolas públicas sob tutela do Ministério da Educação e Ciência, e que frequentaram educação geral no ano anterior (ensino básico). Para este grupo de alunos concluímos que o ensino profissional apresenta impactos positivos no desempenho escolar (taxas de transição, de conclusão e desistência) e no desempenho laboral (taxa de emprego, salário médio e média de meses trabalhos por ano). Os nossos resultados também mostram que os estudantes do ensino profissional têm menor probabilidade de prosseguir para o ensino superior. Estes resultados são robustos face a variações do método de estimação.

Classificação do JEL:
C14 – Métodos não-paramétricos;
I26 – Retornos da educação;

Palavras-chave:
Ensino profissional; Ensino secundário;
Avaliação contrafactual de impactos; Emparelhamento exato por grupos;

Acronyms

VE – Vocational Education

GE – General Education

HE – Higher education

MES – Ministry of Education and Science

CEM – Coarsened Exact Matching

CIE – Counterfactual Impact Evaluation

Acknowledgments

I would like to think to my family, the main reason why I have always continued to investing in my academic and professional career.

A special acknowledgment to my supervisor, Ricardo Mamede, for giving me the opportunity to participate in an incredible research project and for the patience to teach me everything (about the project and related technical aspects). This project was an amazing journey that allowed me to develop not only my professional but also my personal characteristics. I would also like to thank all the colleagues involved in the research project, with a highlight to Teresa Fernandes, for all the help in technical aspects.

Also a special acknowledgement to Rodrigo Duque, for all the motivation and support over the academic cycle.

Index

Index of Tables

Index of Equations

1. Introduction

Globalization exposed national economies to high competitive pressures, especially those countries that cannot compete with developing countries' low wages. In these cases, competitiveness increasingly depends on the quality of the goods and services they produce (Field, Hoeckel, Kis, Justesen, & Kuczera, 2009). This, in turn, requires a skilled labor force, including both high-level skills based on university education and mid-level skills, of a more technical and professional nature (*idem*). Individuals with low-skills are more exposed to competition from other countries and are thus more vulnerable in the labor market.

Portugal presents structural problems with the education system for a long time. It has one of the lowest levels of qualifications in the European Union (EU) (e.g., in 2014 only 36% of the resident population had secondary or higher education as the highest level of qualifications[1]) and these characteristics directly influence the country's economic performance. According to Pereira & Aubyn (2009), increasing the levels of qualifications below the tertiary level can have positive and significant impacts in the Portuguese economic growth. Moreover, Portugal presents one of the highest dropout rates in the EU (19,2% in 2013[2]). In periods of economic crisis, the youth with low/incomplete qualifications are pushed either to unemployment, additional training or to a status of inactivity (NEET - not in employment, education or training). In Portugal, the percentage of individuals in this situation is especially high, reaching 15,5% of the young population (ages between 15 and 29) in 2012[3].

In such context, Vocational Education (VE) seems to be, as Eichhorst *et al*. (2015) put it, "(…) the silver bullet for the youth joblessness problem", allowing individuals to increase their education while having a much more practical component that directly responds to the needs of the local economic tissue, thereby increasing their chances "of a qualified professional career"/"finding

[1] Resident population with 15 years old or more. Source: Pordata.
[2] Percentage of students with ages between 18 and 24 that dropout from the education system without completing secondary education. Source: Pordata.
[3] According to *Education at a Glance 2014* (OECD). Data of 2012.

employment as skilled workers". The purpose of this thesis is to assess the impact of a policy implemented in Portugal as part of the education system reform that took place in 2004. This reform introduced VE in secondary public schools, being largely funded by the European Social Fund[4]. This thesis aims at estimating the impacts of this policy on students' academic and labor market performance, focusing on students enrolled for the first time in VE high school programs in the academic years of 2008/2009, 2009/2010 and 2010/2011.

In estimating the impact of a public policy one cannot directly compare the output of treated and not treated[5] individuals as a whole, since the two groups tend to display different distributions of the characteristics that determine individuals' performance. Thus, in order to isolate the impact of the public policies under analysis from other cofounding factors we will use a Counterfactual Impact Evaluation (CIE) logic. That is, we attempt to estimate individuals' performance, had they not been subjected to the policy – the counterfactual case. In particular, we use the Coarsened Exact Matching (CEM) procedure, a non-parametric method that is based on the comparison between individuals that are similar according to all observable characteristics to estimate the hypothetical outcome. In sum, this thesis attempts to answer the question "What would have happened to students, had they pursued a general education program instead of a vocational course in secondary education?".

The remainder of the thesis is structured as follows. In part Two we present the topic and the related literature. In part Three we present the main components and characteristics of VE. This consists in a short introduction to the history, purpose and main features, supply determinants, and evolution of VE in Portugal. Part Four presents the data, the methodology, main results and sensitivity analysis. Part V concludes the thesis conclusions and discusses possible future extensions.

[4] Vocational education presents several courses in the education system (adult vocational courses, apprenticeship systems, education and training courses, etc. The objective is to analyze the impact of vocational courses of secondary level, "Cursos Profissionais").

[5] Following the common practice in the counterfactual evaluation literature, we use the expressions "treated" and "non-treated" to refer to individuals who directly benefitted from the policy (i.e., VE students) and those who did not (i.e., general education students), respectively.

This thesis was developed as part of the research project "European Social Fund - Supporting youth in Portugal, Counterfactual Impact Evaluation of Vocational Training and Traineeship", financed by the European Commission - DG Employment, Social Affairs and Inclusion and coordinated by the Portuguese Development and Cohesion Agency.

2. Literature review

In this chapter we discuss the role of VE in different education systems and present a brief review of the literature on the impacts of VE in students' performance.

2.1. Vocational education and its role in national education systems

Different expressions have been used to characterize the type of educational programs under analysis, including the following: "Vocational education and training", "Technical vocational education and training", "Initial vocational education and training", "Continuing VE/adult VE", and "Vocational Education" – as presented in CEDEFOP (2007). In this thesis we will refer to "Vocational Education" (VE), which is most frequently found in the literature and coincides with the expression used by the Portuguese national authorities. VE includes Apprenticeships, Adult vocational education and training, Technological courses and VE in basic or secondary level. The programs' structure and the weight of VE varies widely across national education systems. This increases the difficulty of international comparisons, and an acute sensitivity to national contexts is advised when evaluating each specific program (Field *et al*., 2009).

VE programs typically aim at fostering students' specific-knowledge related to a work-area, including a general component and, usually, practical work. The objective is to provide students with specific skills that are in high demand in the labor market, thereby promoting a smoother transition from school to work. This type of education is an alternative to general education (GE), where the main objective is to transfer general knowledge of different subjects to help developing students' future adaptability skills (mostly orientated to higher education (HE)). Students may prefer VE over GE "(…) as it implies shorter investment of human capital and facilitates earlier entry into the labor market." (Eichhorst *et al*., 2015).

The program is usually designed to engage youth who lack the skills or motivation to pursue HE or are at the risk of dropout from the education system. These characteristics often lead to a discriminatory image of this type of education (as evidenced in the Portuguese case[6]). The program includes a short learning component in a workplace environment, usually referred to as

[6] For more on this see, for example, Grácio (1989) and Silva (1998)

apprenticeship. This workplace training experience is a crucial component of the course as it allows students to acquire work-relevant technical and general skills that cannot be learnt in the classroom, while providing opportunities for students to contact with employers (Field *et al.*, 2009). The employers have the advantage of meeting the programs' curriculum (which may help in future recruitment needs) and, by maintaining a close relation with graduates, the problem of mismatch of skills can be reduced.

Depending on the curricula and the linkage with the workplace training component, VE's programs can be classified into three models (Eichhorst *et al*., 2015), which vary according to the level of the workplace training component:

- *School-based* or *School-work based*: the majority of the course's duration (minimum of 75%, according to OECD (2014)) is offered in a school environment and, in most countries, this model includes a short workplace learning component at an institution or company;
- *Dual-system*: presents some small derivations due to the length of training and the place where formation occurs (more than 25% of the program), thus being designated as dual-system, dual-apprenticeship or work-based VET.
- Apprenticeships: corresponds to the higher level of work integration, where the program is essentially workplace-based and provided in a company or at qualification centers[7].

The level of interaction between schools and companies varies across models, posing different challenges to their operation. National education systems differ not only with respect to the overall role of VE, but also to the predominant VE model (see Table I, in the next page).

[7] In the portuguese case this system corresponds to the "Sistema de aprendizagem".

Table I – Distribution of GE or VE over countries

<table>
<tr><th>Countries with predominance of GE (more than 50% of the education system)</th><th colspan="3">Countries with predominance of VE (more than 50% of the education system)</th></tr>
<tr><td rowspan="2">Canada, Chile, Estonia, Greece, Hungary, Iceland, Israel, Ireland, Japan, Portugal, Korea, Mexico, New Zealand, Turkey, United Kingdom, United States of America</td><td>School-based VE</td><td>Dual-system (over 30% of VE's presence)</td><td>Dual-system/ Apprenticeship (over 90% of VE's presence)</td></tr>
<tr><td>Australia, Belgium, Finland, Italy, Luxemburg, Netherlands, Norway, Slovenia,</td><td>Austria, Czech Republic, Slovak Republic</td><td>Denmark, Germany, Hungary, Switzerland</td></tr>
</table>

Source: Own elaboration. Data from *Education at a Glance 2014: OECD indicators* (data of 2012)

2.2. Evidence of impacts

Most of research on VE focuses on the transition from school to work and early labor market impacts, while studies focused on students' performance in the secondary cycle are scarce. The literature converges in the conclusion that this type of education facilitates students' transition to the labor market, reducing the probability of unemployment due to lack of skills and of being employed in low-skilled jobs.

Most studies on the effectiveness of VE in facilitating students' transition to labor market focus on the comparison of different VE's models or the comparison of VE to GE. For example, Bishop & Mane (2004) analyze the impacts of VE in secondary conclusion rates and in labor market's earnings. The authors apply an Ordinary Least Squares regression and, to control for self-selection bias, include as many observable characteristics as possible. Using 12 years of United Sates' panel data, they found that VE students spent more time in employment (in the immediate year after graduation and 8 years later) and earned more in these two periods, compared to students of GE (for students who did and did not enrolled into HE). Cappellari (2004) analyzes the impact of high school types in labor market outcomes in Italy. The author carefully accounts for endogenous selection issues by employing an instrumental variables approach and found that enrolling in GE

increases the probability of attending HE and reduces the participation in labor market, while enrolling in VE improves the transition into labor market and the probability of employment.

Lee & Coelli (2010) analyze the effects of obtaining a VE certificate level of qualifications in Australia, for distinct groups of students. Accounting for selection bias, the authors use a propensity score matching and estimate the differential impacts in the employment rate and weekly earnings for completers and non-completers of the secondary cycle. For graduate students they found that there is no advantage in enrolling in a low level VE course (after the conclusion of the secondary education), only the conclusion of a VE certificate level or above will confer positive impact. For students that do not complete secondary education there is significant impact, regardless of the level of VE.

The comparison of School-based VE and a small derivation of this model can be found, for example, in Polidano & Tabasso (2014)[8]. The authors analyze the impacts of incorporating a short structured workplace learning component, within School-based VE, in students' school performance and the effects in post-school education. Using the counterfactual approach (and propensity score matching), they found that enrolling in a VE's work-based subject increases the chances of graduation, the transition from school to work and reduces the chances of being employed in part-time one year after out of school. Although it does not affect significantly the chances of enrollment in post-secondary education, it does affects the post-school education pathway followed.

Despite the recognized advantages of VE, the available literature has not found "(…) a universal advantage of vocational over academic education for youth's labor market outcomes, although the analysis has been problematic." (Hanushek *et al.*, 2011). Critics to this type of education point that it does not teach enough general skills to students and this influences their future capacity of adaptability, as "(…) specific skills become obsolete too quickly and it is necessary to give people

[8] Ryan (2001) and Wolter & Ryan (2011) compare GE with the Apprenticeship model of VE. Alet & Bonnal (2011); Parey (2009) and Eichhorst *et al*., (2015), in turn, compare VE School-based model with the Apprenticeship Model. We do not present the detailed explanation of these studies as the Apprenticeship model is not our focus of analysis in VE

the ability to adapt to new technologies" (Hanushek *et al*., 2011). This system can also be seen as a discriminatory mechanism or one that increases the risk of social segregation, because it deviates students from higher education and managerial positions (e.g, Arum & Shavit (1995)).

The last aspect was carefully discussed by Shavit & Muller (2000). The authors analyzed the relation between the concepts 'safety net' (expected lower unemployment rate and lower employment in low skilled jobs) and 'diversion' (students' employment in jobs that do not represent the main field of graduation or a deviation from HE). They show that the effect of VE depends on institutions factors, such as if it is occupation-specific rather than general, the possibility of mobility between courses and if there is history of stratification in the education system. The authors found advantage of specific VE (employers are involved in the programs' curriculum) over School based VE and these impacts are significant higher when occupation-specific knowledge is aligned with the future job position, an idea also defended by Boesel *et al*., (1994: 137): "(…) the strongest, most consistent finding throughout the literature is that improved earnings do accrue in situations where vocational training is directly related to job tasks". They also found that VE students got jobs with lower occupational prestige than those who attained academic tertiary education and this difference intensifies in countries where VE serves as an effective safety net. The authors conclude that the concepts are "the flip sides from the same coin" rather than mutually exclusive, as generally seen in the literature.

Using the same data set as Bishop & Mane (2004), Meer (2007) analyzes VE in a period in which this type of education is being highly discriminated in the American education system, by the claims that this VE is, as he quotes, a "(…) dead-end path for students" (Meer, 2007: 1). The author divides VE in tracks (technical and business), considers GE, and carefully accounts for selection bias by applying a multinomial logit to estimate the counterfactual. The aim is to analyze the impacts of self-selection in the program on students' performance. Based on the counterfactual, the author finds comparative advantage of tracking: those on the technical track would not earn more had they choose other path; those on GE would benefit from some technical education and students in the business track would observe some additional impacts if increased their share GE. He concludes by saying that each of them are better where they are and "In any case, whether [VE]

serves as a ‘‘safety net’’ or an alternate path for those with different skills and interest, vocational education should not be stigmatized as a lesser or demeaning option.” (p.572).

To analyze the long-term effects of VE and its implications in future capacity of adaptability, Hanushek *et al*., (2011) use a difference-in-differences approach (to account for selection bias) and analyze the impacts of this type of education in long-term employment trends, for 18 countries. They find evidence that the relative advantage of VE students (higher employment rate in the short-run) diminishes with age, due to the fact that that VE leads to slower adaptability to technological and structural changes in the economy. There is heterogeneity of employment effects: the pattern is more pronounced in countries with apprenticeship programs (Germany, Denmark and Switzerland) and it is not clear for countries with School-based VE. Their results are robust when using propensity score matching.

In Portugal, studies on the impact of education are focused, essentially, on the impact of different levels of education and not on the comparison of different types of education (with a special focus of studies on HE, as Portugal still presents considerable returns to this level). Recently, Oliveira (2015) estimated the returns of VE and GE on wages. The author analyses a panel data of 16 years, following individuals since their first job. The author finds that VE has a return of additional 2% over GE’s wages. However, the rate of growth of wages seems to be smaller for VE students and is surpassed after eight years of experience.

It is clear that VE successfully helps students in the transition from school to work and may improve jobs’ quality (at least in the short run). The pointed disadvantages are related to, essentially, its deviation from HE, the possibility of deviation in the future pathway and, finally, a mechanism of social discrimination. However, it must be noted that the focus group of VE may differ for each education system. So, the impacts are not feasible of generalization and should carefully consider this aspect.

3. Vocational education in the Portuguese secondary education

The purpose if this chapter is to present the main components and characteristics of VE. This corresponds to a small introduction to the history, purpose, main features, supply determinants, and evolution of VE in Portugal.

3.1 Historical origins of vocational education

In Portugal, the origins of technical and professional education date back to the 19th century, in 1836, when two conservatories of arts and crafts were created (one in Lisbon and another in Oporto). These were equipped with several factory machines to practice and analyse the evolution of production techniques. The first industrial school was created in 1852, and thirty years later there were 28 "technical schools" in the country. In 1929 the national school network was reorganized, establishing 19 industrial schools, seven commercial schools and 20 industrial/commercial schools, with technical education being offered only in these institutions and immediately after the conclusion of primary education (Jardim, 2000).

In 1947/48, technical and general[9] education was promulgated. General education was composed by three elements: the first cycle with two years ("preparatory course"), the second cycle with three years ("general course") and the third cycle with two years ("complementary course"). Technical education was also composed by the "preparatory course" and the "general course" (with three/four years), fixing in five/six the total number of years required to complete a vocational-technical course. The technical course's curriculum was composed by a low workload of general subjects and a strong technical component, in contrast to the general course, composed mainly by general subjects. Graduation from a technical course did not allowed access to higher education, which contributed to produce a discriminatory image of this type of education and limited its expansion.

The Reform in the education system of 1967[10] increased compulsory education from four to six years and the technical programs' curriculum was reorganized with the aim of approaching general programs' one. Technical education was now composed by the "preparatory course"

[9] "Ensino liceal", in Portuguese. This is a specific Portuguese with no immediate English correspondence. Given its content and place in the education system, general education constitute the best approximate translation
[10] Decree-Law Nº. 47/430 of 2 of January 1967

(corresponding to the fusion of the first cycle of general and technical education), the "general course" (now with three years of duration) and the "complementary course" with two years[11]. Technical courses covered several areas of education, including commercial, industrial, visual arts, applied visual arts and agriculture.

In 1975, following the so-called Veiga Simão's Reform[12], compulsory education increased from six to eight years. The "general course" of both programs was merged and technical education was introduced in high-schools[13]. This merge, along with the discrimination associated with VE programs – which were attended mainly by disadvantaged people (Grácio, 1989; Silva, 1998) – led to the extinction of technical-professional education in the following years.

In 1983, in an attempt to renew technical education, *Technical-professional* and *Professional* programs[14] were created at the secondary level, initially designed as a pedagogical experience involving those educational establishments that volunteered to participate (including private establishments). The *Technical-professional* programs had duration of three years, providing students with a certificate of conclusion of secondary studies, therefore allowing access to higher education. The *Professional* programs, in contrast, had duration of one and a half years, with the first year being dedicated to classroom education and the following six months to workplace training. These programs did not allow a direct access to higher education, having as their crucial objective facilitating students' transition to labor market.

Due to the difficulty of organizing apprenticeships and the lack of demand, the *Professional* education petered out. In the academic year of 1985/86, the unfilled vacancies reached 59% of the total – comparing to 35% in the case of *Technical-professional* programs (Azevedo, 2014).

[11] Both types of education offer now the same duration in general and complementary course
[12] Veiga Simão was the Minister of Education from 1970 to 1974. His name is associated with a number of important milestones, including the creation of the universities Nova de Lisboa, Aveiro and Minho, and the Law N.º5/73 of 25 of July (an important antecedent of the Basic Law of the Education System)
[13] High schools would now offer three types of courses: artistical, technical and general courses
[14] Legislative Order N.º 194-A/1983 of 21 of October

The lack of demand of both types of programs can be explained, among other factors, by the dominant tradition of general courses in the education system, which placed technical-professional education as a second choice, and by the fact that virtually all students who attended secondary education at the time intended to proceed to higher education. In the same year, the Basic Law of Education System (still in force today) was approved, extending from two to three years the duration of the secondary school programs (this entered into force in 1989[15]).

A new model of professional schools (public and private)[16] started in 1989, as a result of the joint work of the Ministry of Education and the Ministry of Employment and Social Security. The creation of professional schools implied the celebration of a contract-program between the government and local entities (several agents and partners)[17]. The entities/partners adhered rapidly to this model, leading to the creation of 168 new schools in the first four years[18]. In the same year, the *Technical-professional* programs were replaced by the *Technological* programs.

By that time, secondary education was characterized as follows:

- The education establishments were composed by (i) high-schools (teaching general and technological programs), (ii) professional schools (teaching professional programs) and (iii) formation centers (with alternance training);
- All types of programs had a duration of three years and were composed by three components of formation (general, specific and technical);
- Upon conclusion, all types of programs provided students with certificates that were equivalent with the respect to the level of education attainment, therefore allowing access to higher education.

[15] Decree-Law N.º 286/89 of 26 of August
[16] Decree-Law N.º 26/89 of 21 of January
[17] This innovative model altered the traditional perspective of the education system, in which the government had the monopoly of education's supply. Entities offering this new model had increased autonomy in its management (pedagogical, administrative and financial autonomy)
[18] Data from Department of Secondary Education (MES), see Azevedo (2014: 433)

After 1994, a decrease in the growth of professional schools became evident, which is explained by their dependency on European funds, that limited the expansion of this type of education (Jardim, 2000: 53).

A new reform, adopted in 2004[19], led to the modification of the operating and evaluation models of each type of program, the clarification of their purpose, and the increase and diversification of the education offer at the secondary level, leading to the introduction of VE in public high-schools. The 2004 reform had as main objectives: (i) reducing the high rates of school failure and dropout, (ii) increasing and diversifying alternative educational pathways and (iii) improving the match between labor markets' demand and supply of technical skills.

Upon the reform of 2004, the legislative framework for VE was revised and in October 2007 the Agreement for VE's Reform[20] was approved. This new regulatory framework created the System of National Qualifications (SNQ) responsible for the management, categorization, certification and skills recognition of all the types of VE and training (initial, adult, apprenticeships, etc.). The SNQ makes a diagnosis of the framework and defines a set of guidelines for public investment in vocational training. The categorization of VE is present in the National Qualifications Catalog (CNQ), in which all sectors of activity are covered. This catalog is in constant update to ensure that the courses satisfy the needs and dynamics of labor market's demand of technical skills. The National Qualifications Framework (QNQ) matches national qualifications with the European Qualifications Framework (EQF).

[19] Decree-Law N.º 74/2004 of 26 of March
[20] "Acordo para a Reforma da Formação Profissional". Consult the document in: http://www.catalogo.anqep.gov.pt/.

3.2 Purpose and main features of vocational courses

3.2.1 Education components and evaluation system

VE is one of the alternative types of programs offered in the current configuration of the Portuguese upper secondary education system (ISCED 3). VE was introduced in public secondary schools as a result of a reform in the secondary education system, introduced by Decree-Law nº.74/2004 of 26 of March[21]. This reform had as main objectives to reduce school failure and dropout, increase and diversify alternative educational pathways and to improve the match between labor markets' demand and supply of technical skills. The operation and evaluation standards of the courses were modified, as well as the clarification of their purposes:

1. General education/Science and Humanities courses ("Cursos Científico-Humanísticos"), designed with the aim of subsequent access to HE;
2. Technological courses ("Cursos Tecnológicos"), with a double perspective of insertion into labor market or access to HE (into a vocational course of superior level or undergraduate course);
3. Artistic courses/Specialized art courses ("Cursos Artísticos"), designed to offer qualification in different artistic areas with the aim to access labor market and, depending on the specialization, there is possibility of access to HE;
4. Vocational education ("Cursos Profissionais"), designed with the aim of providing initial qualification to student, prioritizing its insertion into labor market and allowing access to HE[22];
5. Recurrent education ("Ensino Recorrente"), allowing a second opportunity of education to the previous courses (except for VE);

VE aims to provide the student with general knowledge, occupation-specific knowledge of a training area and practical skills (developed in a workplace environment). VE targets students who completed basic education or equivalent and have the maximum age (at entry) of 20 years[23]. The

[21] And with some changes by: Amendment N.º 44/2004 of 25 of May; Decree-Law N.º 24/2006 of 6 of February; Amendment N.º 23/2006 of 7 of April and Decree-Law N.º 272/2007 of 26 of July

[22] Although not mentioned in the decree-law, there are two other courses of VE (excluding adult training), Apprenticeship training ("Sistema de aprendizagem") and Education and training courses ("Cursos de educação e formação"). However, the tutelage of these courses is not exclusive of the Ministry of Education and Science and, therefore, not mentioned

[23] Except for students with disabilities or students with repetitive situations of school failure or dropout, to which the age limit does not apply

course has a duration of three academic years and is organized into modules of different length, allowing for a more flexible learning process.

Students can choose from 39 areas of training[24] and the programs' curriculum includes the following education and training components[25]:

a) Sociocultural component: aims to contribute to the growth of students' personal identity, social and cultural aspects. It is composed of five general subjects;
b) Scientific component: corresponds to the acquisition and development of a set of basic knowledge of the course. It consists of two to three subjects;
c) Technical component: acquisition and development of a set of occupation-specific knowledge and its application in a short workplace environment.

The evaluation of knowledge is performed by the diagnosis, formative and summative evaluation[26]. The *diagnosis* evaluation is carried out in the beginning of each academic year and intends to define (if necessary) strategies of pedagogical differentiation, to overcome difficulties and students' learning and scholar integration. The *formative* evaluation is a continuous and systematic process and allows the collection of information on students' learning process. Finally, the *summative* evaluation intends to classify and certify the acquired knowledge. The latter includes the external summative evaluation through National Exams (when applicable – more on this bellow), conducted by the Ministry of Education and Science (MES), and the internal summative evaluation, carried out by the schools' teaching staff, and which is composed by the following elements:

a) Evaluation by subject: the number of subjects depends on the specific programs, varying between from 10 to 12. This part of the evaluation is done by modules, and the final classification is obtained by simple arithmetic average of all modules. The weight of this evaluation component in the final classification is approximately 67%;

b) Vocational ability test ("Prova de Aptidão Profissional"): students present a project, embodied in a material or intellectual product, to demonstrate the knowledge and

[24] To see all the areas consult Table I, in annex
[25] For more detail on the components' subjects and workload see Table II, in annex.
[26] Decree-law nº.139/2012 of 25 of July

professional skills developed throughout the course. This test has a weight of approximately 23% in the programs' final classification;

c) Workplace training component ("Formação em Contexto de Trabalho"): students put in practice the occupation-specific knowledge developed in the classroom during an internship in local employing entities (public or private). It may occur in the final year of the course, or throughout the duration of the program. Upon its conclusion students must present an evaluation report for the internship. This component corresponds to 14% of the total workload of the course and has a weight of approximately 10% in the courses' final classification.

Based on the study of Neves (Coord.) *et al*., (2009: 50–53), we can identify some aspects of the workplace training component. Is a distinctive element of the vocational track, presenting considerable advantages for the participants of the process. For students represents the first contact with the labor market, allows them to put in practice the knowledge developed in the classroom and stimulates a new posture towards the labor market 'world'. For the cozy entities allows to understand the profile of vocational courses, the quality of training provided by schools and if the courses correspond to the entities' needs of recruitment. By organizing this component, schools create and improve their network of local entities. To select the cozy entities the schools determine some criteria such as localization, previous record of internship with students and students' preference. Upon demonstration of interest from the entity, the students' training plan is defined with the definition of objectives, workload and training elements.

According to data from "Survey on Schools" (Pereira (Coord.) *et al*., 2011), 205 out of 212 had established partnerships with municipalities and companies, while cooperation with civil society actors (like business associations and other training entities) is less representative. In a general framework there is no difficulty for schools to connect with the hosting entities: these are usually willing to offer internship to students and participate in activities promoted by schools, and are directly involved in the evaluation of vocational courses (e.g., integrate the jury for evaluation of the vocational ability test). From the same evaluation study, according to data from "Inquiry to the Hosting Entities of the Workplace Training component of 2011", the host are equally distributed (enterprise size) and the majority is of private nature. In general, they consider students'

performance positive and identify as strengths the technical knowledge of the professional area, computer skills and team work capacity; as weak points they identify the poor initiative, the knowledge of foreign languages, and capacity to identify and solve problems.

3.3.2 Access to higher education

To complete a VE program students are not required to perform national exams[27] and this also applies if students choose to continue their studies through a vocational course of higher level. However, students who intend to proceed to HE are legally required to perform two national exams[28]. Contrary to students from GE, who have to perform four national exams, VE students have to perform only two. This may been seen as an advantage over the former group. However, it must be noted that vocational students' workload is more demanding for two reasons: first, the workplace training component typically occurs in the end of the year, which is concomitant with the delivery of the vocational ability test and the National exams; second, one of the exams is chosen from the general courses' subjects, so it may correspond to a subject not taught in the vocational track (requiring an extra effort from VE students in preparing for national exams).

3.3 Supply-side determinants of vocational courses

This section presents the determinants of vocational programs' supply related with schools' financial support, physical conditions and human resources.

3.3.1 Schools' financial support

Vocational courses are financed by national funds, contributing with 30%, and by European Union funds, through the European Social Fund (ESF), with a contribution of 70%. The funds are calculated in function of the total eligible cost, deducted by private contributions and own revenues. The total eligible cost includes costs with students (study materials, food, transportation, etc.); teaching staff; rents, leases and amortizations; overheads and thematic seminars and others.

The access to funds is obtained by multiannual application, per academic year. The application is based on a three-year training plan as a part of the annual public network of programs in order to

[27] Decree-Law nº74/2004 of 26 of March
[28] Ordinance nº.165-B/2015 of 3 of July

obtain educational advice and consent by the regional education authorities. POPH selects applications based on 11 criteria such as: the quality of human resources responsible for direct and administer training; the capacity and adequacy of infrastructures to the proposed offer; the relevance of the proposal for the local economic fabric; coordination with the local offer network; etc. The applications can be presented by the following educational institutions, whose operation was authorized previously by the MES and working under its tutelage[29]:

a) Private vocational schools: teaching establishments focused on offering vocational education and dual education at a non-superior level, created by single or collective private nature, with or without lucrative purposes;
b) Public vocational schools: teaching establishments focused on offering vocational education and dual education at a non-superior level, work under the Ministry of Education and Science and include the public network of schools;
c) Vocational schools of business/entrepreneurial reference: private vocational schools created by companies or business entities to offer courses directly related to the economic activity.

In conclusion, the funding of vocational programs depends on: first, the previous teaching permission by regional education authorities and the MES; second, the adequacy of infrastructures and competences by the school; and, third, on the approval of the costs' plan by POPH.

Beyond the access to funds, schools can also create vocational courses if they believe and present evidence that the proposal goes in line with the needs of the region. For that purpose they have to present the courses' objectives and curriculum to the MES and to ANQEP[30]. Upon approval, the MES determines the program for the sociocultural and scientific component of formation and the school is responsible for the organization of the technical component.

[29] Decree-Law N.º 92/2014 of 20 of June
[30] National Agency for Qualification and Vocational Education and Training ("Agência Nacional para a Qualificação e o Ensino Profissional"), one of the responsible for the System of National Qualifications (SNQ)

3.3.2 Infrastructures and human resources

The infrastructures, equipment and human resources are features that determine the opening and expansion of programs, and influence the quality of the teaching process. Trough application to POPH, schools obtain funds for the acquisition of consumables and expenses of maintenance of equipment and installations[31]. However, as argued by Neves (Coord.), *et al*. (2009), the funds are insufficient to insure vocational courses' conditions of good functioning: "The suitability of equipment to professional courses is a situation far from resolved". Some schools establish agreements with local entities to benefit from technical support, sharing of spaces and equipment. Faced with this situation, the government has made a number of investments. The Renovation Program of Schools and the Education Technological Plan[32] are examples of interventions that aim to improve the physical structure of schools. These interventions allowed the construction of workshops spaces, providing computers, projectors and internet access in classrooms. Despite the government's effort, according to Neves (Coord.), *et al*., (2009), schools still lack updated computer software, which is especially relevant to run more technologically-oriented programs (e.g., computer sciences; multimedia; electricity; etc.).

Regarding human resources, schools prefer resort to internal teaching staff, although there is the possibility of external recruitment. However, according to the legislation, this can only occur through fixed-term contracts[33]. According to Neves (Coord.) *et al*., (2009: 43) it is difficult to recruit teaching staff in specific technical areas. Since the expected remuneration is unappealing, schools tend to attract teachers with little professional experience, especially recent graduates who accumulate the provision of services in more than one school. The numbers of years of experience in the area, as well as the experience in the programs' sector of activity are the most important determinants in the recruitment process.

The same study identifies other difficulties, such as the more demanding organization of classes, the additional collaborative work between teachers and increased monitoring of students' learning

[31] The acquisition of equipments, vehicles, infrastructure, real estate and furniture are not eligible costs for application, so not funded - Legislative Order N.º 74-A of 2008
[32] For more information see Decree-Law N.º 41/2007 of 21 of February, Council of Ministers Resolution N.º 137/2007 of 18 of September and "Parque escolar 2007-2011. Intervenção em 106 escolas" (PE-EPE, 2011)
[33] Decree-Law N.º 35/2007 of 15 of February

process, all of which make the permanent coordination between the teaching staff highly valued. To facilitate this situation schools try to maintain the same teaching team over the training cycle, something that is often jeopardized by fixed-term contracting.

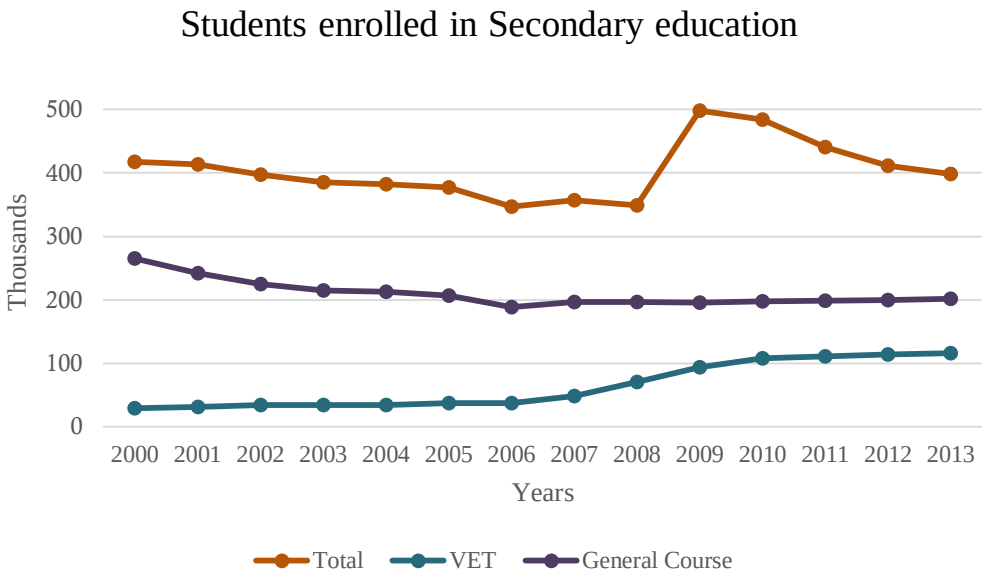
3.4 The policy under analysis

The main source of funds for the new Vocational Education policy, introduced after the 2004 reform, was the Third Community Support Framework ("Terceiro Quadro Comunitário de Apoio", QCA III 2000-2006), under the Operational Program for Education ("Programa Operacional Educação", PRODEP III). This program was composed by four axes and vocational education was integrated in "Axis one: Initial qualification of youth - Measure one". The axis had as main objectives: improve the quality of basic education; to expand and diversify initial training of young people focusing on high quality and employability; to promote lifelong learning and to improve the employability of the workforce; to guide and promote the development of the Information and Knowledge Society in Portugal. The ESF budget for this axis amounted to 390.520.000 euros.

After 2006 the VE policy was funded by the National Strategic Reference Framework ("Quadro de Refrência Estratégico Nacional", QREN 2007-2013) under the Operational Program of Human Potential ("Programa Operacional Potencial Humano", POPH). In general, POPH aims to increase the level of qualification of the Portuguese population and set secondary education as the minimum level; to transform the productive fabric and to enhance economic sectors that contribute with added value; to stimulate the creation of jobs and encourage entrepreneurship. POPH's activity translates into 10 priority axis, and vocational education is englobed in "Axis one: Initial qualification – Typology 1.2". This has as main objectives: increasing the qualification of young; promoting vocational education of secondary and non-tertiary level; promoting employability; reducing the high levels of dropout and school failure; and to increase the equality of gender in access to opportunities. For this period, the ESF contributed with a total of 2.019 million euros for Axis one.

3.5 Recent dynamics of vocational education

The number of students enrolled in secondary education in Portugal had been decreasing since 2000 (from 417.705 students in that year to 347,400 students in 2006, the lowest level in that decade), reversing the trend in 2007 (reaching a total of 498.477 students in 2009). The increase since 2007 was due, mainly, to the increase in the number of students enrolled in vocational courses and recurrent education (with increases of 56.673 and 104.303, respectively, between 2006 and 2009[34]).



Source: School Census – DGEEC/MEC
Note: the spite of 2009 is due to the increase in recurrent education

The increase in the number of students enrolled in VE does not appear, however, to occur at the expense of a decrease in other educational pathways or at a decrease in the private sector: in the academic year of 2004/2005, the number of students enrolled in private vocational schools was just over 32.000, a number that remained nearly constant until 2007/2008 and increased to 36.000 in 2008/2009[35]. Only technological courses had a decrease in the number of their students, a trend present since 2000/2001. The increase in the number of students enrolled in VE programs is associated with the opening of 40.000 vacancies in 2008/2009 and nearly 50.000 in 2009/2010 (Neves (Coord.), *et al*., 2009: 21).

[34] Includes public and private education. See Table III, in annex, with the evolution of all types of education
[35] Idem

The following data translates the recent dynamics of expansion of VE (Pereira (Coord.) *et al.*, 2011: 172):

- Students enrolled by institutions' nature

  2005/2006 – 4.302 students in public institutions (11,6%) and 32.641 in private (88,4%)

  2009/10 – 63.623 students in public institutions (58,6%) and 45.020 in private (41,4%)
- Evolution of number of schools offering VE:

  1990 – 95 vocational schools and promoters[36]

  2009 – 664 vocational schools and secondary schools offering vocational courses[37]
- Evolution of number of courses:

  2004/2005 – 3 courses

  2010/2011 – 130 courses
- Ratio of students[38] of VE to students from GE:

  2000 – 0,40

  2010 – 0,63

[36] Data of Department of Secondary Education, in Azevedo (2014: 433)
[37] Data GEPE – MES in (Pereira (Coord.) *et al.*, 2011: 202)
[38] Students enrolled in technical-professional and vocational courses. Includes public and private institutions. Data from DGEEC/MES – School Census, in Pordata

4. Assessing the returns to vocational education in Portugal

We now turn to the analysis of the impact of vocational courses in students' academic and labor market performance. We start by presenting the data and the estimation method, followed by the results and the sensitive analysis.

4.1 Data Collection

To construct a viable counterfactual we need data related to: the treatment status (in our case, individuals are considered as *treated* if they were enrolled in a vocational course and *non-treated* if they were enrolled in a scientific-humanistic course); characteristics that influence students' performance (sociodemographic characteristics, household and school context, before and after the policy intervention); and, finally, indicators of students' academic and labor market performance (variables related of transition, conclusion, dropout, access to HE and transition to employment).

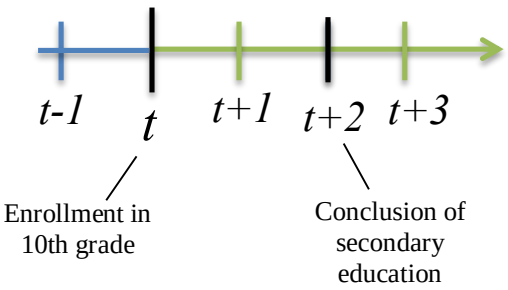
The data used in the research project draws from different sources, namely:

- The Statistics Bureau of Education and Science (*Direção Geral de Estatísticas da Educação e Ciência*, DGEEC-MEC), provided all the information regarding students' performance, schools characteristics and HE data. For this purpose DGEEC merged several administrative datasets, collected from schools and HE institutions;
- The Institute of Computing of Social Security (*Instituto de Informática da Segurança Social* – IISS) provided information about students' employment situation and details of the job.

For the purpose of the project a special key-code was created in order to allow crossing the two databases while assuring anonymity. After merging the data, the information was anonymized and delivered to the research team by the data providing institutions. The researchers involved in the project, including the author of this thesis, did not have access to the corresponding key, in order to comply with statistical secrecy laws.

The final database contains information from the academic years of 2008/2009, 2009/2010 and 2010/2011. Each academic year contains information related to students' sociodemographic characteristics, household's characteristics and schools' context from the year prior to enrolment

in secondary education (t-1) until the year after expected conclusion of the 12th grade (t+3), as illustrated in the figure below:



4.1.1 Group of interest

In assessing the impact of the policy, several factors led us to restrict the universe of students under analysis, namely: the availability of data, the focus of the policy and the proper existence of a counterfactual.

The policy focused on the expansion of VE in public high-schools, so students enrolled in private schools or special public schools with tutelage different from the MES were left out of the analysis. The evaluation of impact is, therefore, focused on regular public schools and not an all schools offering VE programs.

Students with special educational needs often have an adapted curriculum and it is difficult to find a match to these students (that is, a reliable counterfactual). The same applies to students with nationality different from Portuguese; those with age above 18 or students enrolled in other courses in t-1 (e.g., apprenticeship courses, recurrent education of basic level, technological courses, etc.). Thus, observations with these characteristics are also excluded from the analysis.

In sum, the scope of our analysis is composed by observations of students who satisfy the following criteria:

- Ages between 15 and 18;
- Portuguese nationality;
- Are enrolled in public schools with tutelage of the MES;
- Attended basic education in t-1;
- Do not have any special educational needs in t or t-1.

Observations satisfying these conditions correspond to 25,5% of total VE students and 68,1% of total GE students (see Table II, below).

Table II – Distribution of the target group by type of education

	Total data (1)		**Target group (2)**		**(2)/(1)**
	N.	**%**	**N.**	**%**	**in %**
Vocational education (VE)	125.154	38,0%	31.940	18,7%	25,5%
General education (GE)	204.616	62,0%	139.268	81,3%	68,1%
Total	329.770	100%	171.208	100%	51,9%

Note: statistics for the three cohorts under analysis
Source: Own elaboration.

4.1.2 Data processing

The first step in this process was the analysis of general descriptive statistics, which aimed to assess the quality and quantity of data. During this phase, we used as much information as possible to fill the gaps in the data (e.g., using information at t-1 to fill in information gaps at t, in the case on invariant variables – e.g. sex).

In the following step, a number of (in)consistency criteria were used to check the quality of both the control and output variables. The criteria were based on comparing the values of related variables (e.g., if student sex was female in t-1, then it should be the same in t[39]). Only observations that complied with all the criteria and had no missing values were kept for further analysis[40].

[39] To consult the complete list of inconsistency's criteria see Table IV, in annex
[40] To see the distribution of observations that were excluded from the analysis - invalid data, see Table V, in annex

Table III – Distribution of valid data by type of education

	Target group (1)		Valid data (2)		(2) / (1)
	N.	%	N.	%	in %
Vocational education (VE)	31.940	18,7%	25.792	17,9%	80,8%
General education (GE)	139.268	81,3%	118.694	82,1%	85,2%
Total	171.208	100%	144.486	100%	84,4%

Note 1: includes the three cohorts under analysis
Note 2: the number of valid observations may change according to the outcome variable (because of inconsistency criteria). The present values refer to the variable "Dropout in t or t+1"
Source: Own elaboration.

On the basis of the identification and analysis of valid data, the final control and output variables used for the estimation were chosen. Some of the control variables were subjected to a change in their categories in order to improve the rate of matched observations. Table IV presents these variables with the original and final categories.

Table IV – Control variables: original and final categories

Variable	Original categories	Final categories
Student sex	Female, Male	Female, Male
Age	[14; 45]	15, 16, 17, 18
Scholar social support (t-1)	None, level A, B or C	None, level A, level B/C
Guardians' academic qualifications[1]	None, basic (1st, 2nd or 3rd cycle), secondary, bachelor, masters, doctorate, etc.	Up to 1st cycle, 2nd or 3rd cycle, secondary, bachelor, postgraduate, others
Computer at home (t)	Yes/No	Yes/No
Grade of portuguese exam in basic education (t-1)	[0-100]%	In %: [0-50], [51-75], [76-100]
Grade of mathematics exam in basic education (t-1)	[0-100]%	In %: [0-50], [51-75], [76-100]
Portuguese students in school	Continuous: [5-2486]	In %: [0-94], [95-97,36], [97,37-98], [99-100]
Number of students per staff	Continuous: [1,58-286]	[0-22], [23-28], [29-33], [33-268]

[1]Guardian is the person responsible for the student's academic context, until he reaches the age of 18.
Note: the original categories do not include the application of the group of interest or the criteria for valid data.

The output variables were created by crossing information from different variables:

- **Transition in t:** Is a dichotomous variable. If student was in the 10th grade in t and in t+1 is in the 11th grade, this variable equals one. If student is in the same academic year in t+1, this variable equals zero (student did not transit);

- **Transition in t and t+1:** Is a dichotomous variable. Follows the same logic as the previous variable. It results from the merge of the previous variable with information of t+1. If student transited in both years this variable equals one and if student did not transited in one of the years this variable equals zero;

- **Conclusion in t+2:** Is a dichotomous variable. If student is in the 12th grade in t+2 and there is no information in t+3, he concluded the secondary cycle. If student is in the 12th grade in t+2 and t+3 is considered failure (no conclusion);

- **Dropout in t or t+1:** Is a dichotomous variable. If student was enrolled in t and there is no more information in the following years, we consider that he dropout in the end of t; the same logic applies to t+1. This variable is the merge of information from the previous sources;

- **Enrollment in higher education after t+2**: Is a dichotomous variable. If student is enrolled in HE, in one of the next three years after graduation (t+3, t+4 or t+5), we consider that student is enrolled in HE;

- **Employment after t+2:** Is a dichotomous variable. Corresponds to employment in the immediate year after graduation (during t+3). This variable is constructed using only students who graduated and did not enroll into HE;

- **Average daily salary**: Is a continuous variable. Corresponds to students' average daily salary. This variable is constructed using only students employed after t+2 (who graduated and did not enroll into HE);

- **Average number of months worked per year:** Is a continuous variable. Corresponds to the average number of months worked in the year after graduation (t+3). This variable is constructed using only students who were employed after t+2 (who graduated and did not enroll into HE);

- **Average number of days worked per month:** Is a continuous variable. Corresponds to the average number of days worked per month. This variable is constructed using only students who were employed after t+2 (who graduated and did not enroll into HE).

4.2 Methodology

Our goal is to quantify the impact of a policy on individuals' performance. To measure this impact we cannot directly compare the overall performance of those individuals that were subjected to the policy with that of other individuals, since the pre-intervention distribution of relevant characteristics tend to be uneven between the two groups.

Ideally, the analysis of the causal impact of a policy intervention would be based on a controlled experimental approach, in which individuals would be randomly chosen as beneficiaries. In principle, this random selection into treatment would assure that individuals differ only in the treatment status, and thus any differences in their performance could be attributed to the policy/program. However, this logic requires that the evaluation of the policy is included in its planning phase, which is rather infrequent. In the impossibility of following this approach, researchers/evaluators tend to use quasi-experimental methods to assess the impact of public policies, the aim of which is to simulate the situation of explicit randomization.

4.2.1. The counterfactual approach

In abstract terms, we want to observe the individuals' performance both in the presence and the absence of the policy. However, because this is impossible in practice (the individual is only in one of the two situations in a given time period, never in both), we have to look for alternative ways to estimate the hypothetical outcome. One common solution is to take as the counterfactual the performance of individuals who have not benefitted from the policy but who share the same relevant characteristics of the beneficiaries. To apply this logic it is necessary to apply a matching method, in which individuals T and NT are matched based on their observable characteristics (more on this ahead) and there are several data requirements, namely: indicators of the treatment status – Treated (**T**) if individual was supported by the policy and Non-treated (**NT**) if individual was not supported by the policy; indicators of other determinants of performance (the controlling or cofounding factors); and variables of the final output.

One of the main concerns of these methodologies is the self-selection aspect: in our case, students choose one type of education over another, which is influenced by their pre-intervention

characteristics. Therefore, there is an ongoing correlation (self-selection bias) that has to be dealt with in the estimation.

Standard approaches try to solve self-selection bias by introducing as much observable characteristics as possible in the model. However, as Schlotter, *et al*. (2011) notes "(…) controlling for observable factors does not solve the endogeneity problem when it is due to plain reverse causality in that the outcome causes the treatment.". Even if we account for observable factors that influence self-selection bias, there may be other possible important sources that cannot be measured, such as students' innate abilities, parents' preferences for a type of education, professors' aptitude to teach and motivate, etc. These omitted factors may generate unobserved heterogeneity between T and NT individuals and researchers must control, as much as possible, for there aspects.

4.2.2. Quantity of interest

Consider the potential performance of an individual supported by the policy $Y_i(1)$, and the potential performance of the same individual not supported by the policy $Y_i(0)$. Let T_i be a dummy variable indicator of the treatment status for individual i that takes the value of 0 if individual was not treated (NT) and 1 if it was treated (T) by the policy. The observed outcome variable is:

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0) \qquad \text{(I)}$$

For each individual i, $Y_i(1)$ is unobserved if it was not subject to the policy and $Y_i(0)$ is unobserved if individual was subject to the policy. The treatment effect is defined as $\mathrm{TE}_i = Y_i(1) - Y_i(0)$, this is, the difference between the output in the presence and in the absence of the policy. Let $\tau = \{i : T_i = 1\}$ be the set of indexes for the treated units and $n_T = \#\tau$ be a count of elements of this set. The quantity of interest will be, therefore, the average treatment effect over the treated (ATT):

$$\mathrm{ATT} = \frac{1}{n_T} \sum_{i \in \tau} \mathrm{TE}_i \qquad \text{(II)}$$

This logic implies the following simplifying assumptions:

1. Conditional on the covariates (**X**), the treatment variable is independent of potential outcomes (ignorability assumption):

$$T_i \perp \{Y_i(0), Y_i(1)\} | \mathbf{X} \quad \text{(III)}$$

2. "Matching methods focus on ATT so that if they choose to retain all T units, and prune the NT, the quantity of interest remains the same. So, for each observation, $Y_i(1)$ is always observed, whereas $Y_i(0)$ is always estimated (by choosing values from the control units via some matching algorithm or applying some model)" (Iacus, *et al*. 2011: 4).

4.2.3. Matching methods

The objective of the matching procedure is to eliminate observable differences between observations by comparing each of them with the most similar in the opposite group. By improving the empirical distribution of covariates we can control "(…) for the confounding influence of pretreatment control variables in observational data." (Iacus *et al*., 2011). If the perfect balance of data is achieved there is no necessity to apply other statistical methods and a simple difference of the individuals' output is sufficient to capture the causal effect. Thus, in the many commonly approaches, matching methods are used only as a mechanism to balance the data, so, after a positive result, any other statistical estimator can be applied to assess the impact. "Preprocessing via matching can greatly reduce the degree of modeling necessary and thus also the degree of model dependence." (Ho, Imai, King, & Stuart, 2007).

There are different types of matching methods. First, the *one-to-one exact matching*, that matches each NT unit with a T unit. When there are no sufficient NT for all T, an *approximate matching method* is used. This matching method includes two steps. In the first step the common support is determined, a region of T and NT observations with similar characteristics, that excludes observations with characteristics present only in one group (as including these would require a considerable manipulation of data). The second step consists in the matching of the units in the common support, in which if the values of the covariates are not exactly equal, will be close by a defined metric. The final step of both types of matching methods consists in the estimation of

impacts using an estimator or a non-statistical method. One of the main disadvantages of estimators is that balance is not automatically achieved (unless strict assumptions are met). Researchers should check (using, e.g., the absolute difference in the means of both groups, considering the appropriate weights) and, if a negative result is achieved, the matching algorithm should be altered and the process repeated[41].

4.2.4. Coarsened Exact Matching

For the assessment of impacts we used a variation of the Coarsened Exact Matching (CEM) methodology, developed by Iacus, King & Porro (2011). The CEM results from a combination of the *exact matching* and the *approximate matching method*.

We began by coarsening the data, that is, we recoded data into categories in a format that allows keeping crucial information to obtain valid and accurate estimates. In the present case most of the variables are already categorical; this means that the coarsening step consists in aggregating some categories when deemed appropriate (see Table IV above). When the original variables are continuous, by coarsening the data we transform those variables into categorical ones, on the basis of a detailed analysis of their distribution.

In the next step, separately for each academic year[42], we matched all T students with the NT that shared the same (coarsened) observable characteristics, creating a stratum of homogeneous students. After accounting for the different composition of strata we compute the average performance for both groups – T and NT – and its difference. The latter result corresponds to the impact of VE in each homogeneous group. The first estimation of impacts occurs at the stratum level because we assume heterogeneity of impacts, in which they differ according the group of observations under analysis[43].

[41] The measure of imbalance is necessary only if an estimator is applied

[42] We kept the analysis separated for each cohort in order to control for any specific annual factors that could be influencing our impacts

[43] Assuming homogeneity of impacts would not seem appropriate given our context and would produce one impact for the sample. For example, the impact of VE on students from families with lower incomes is expected to be different from the impact on students from more advantaged families

The impact for each cohort corresponds to the sum of impacts of homogeneous stratum (for the year in question). To aggregate the impacts of different cohorts, we sum the impacts with weights given by the proportion of T units in each year related to the total number of T units in the sample. This estimation procedure can be summarized as follows:

i) Students in the database are sorted into T and NT groups. The groups are further divided into successive cohorts (*p)* based on the academic year in which individuals enrolled in secondary education for the first time (2008/2009, 2009/2010 and 2010/2011);

ii) For each cohort, T students are matched with NT students sharing the same exact values of the control variables (X) mentioned above (see table IV), creating strata with homogenous students (*s*);

iii) For each stratum (*s*), within the same cohort (*p*), the local ATT estimates (τ_{ps}) are retrieved as:

$$\tau_{ps} = E[Y_i(1)|\, T_i = 1, X, p] - E[Y_i(0)|\, T_i = 0, X, p] \quad \text{(IV)}$$

iv) The average ATT for each cohort (*p*) is estimated as:

$$\tau_p = \sum_s [\tau_{ps} \times w_{ps}] \text{ and } W_{ps} = \frac{N_s(T_i=1)}{\sum N_{ps}(T_i=1)} \quad \text{(V)}$$

with weights (W_{ps}) computed as the total number of VE students in each stratum over the total number of VE students in each cohort.

v) The global ATT, our quantity of interest, corresponds to the sum of impacts across all cohorts. They are estimated as the weighted average of τ, with the weights corresponding to the proportion of VE students in each cohort compared with the total number of VE students (total database).

4.2.5. Statistical significance of impacts

To test the statistical significance of the impacts we applied a T-statist test. Since we computed the impact for each stratum and we also keep the analysis separated by academic year, our variance of the test has to consider these aspects. We adapt a formula of Becker (2009: 115), that presents the variance for the case of stratification in the impacts, a variant of the matching method *Propensity Score Matching*. In this method, the matching is based on a score, a composite with information of the covariates (it is applied a logistic regression over the covariates and the coefficient will be the score). In our case, the stratification corresponds to the academic years.

Consider the output of a T individual Y_j^T, and the output of a non-treated individual Y_j^C. Letting q index the strata over a defined period (in our case, each academic year) and assuming independence of outcomes across units, the variance for the estimator of the ATT, τ, is calculated as:

$$\mathrm{Var}(\tau) = \frac{1}{m_T}\left\{\mathrm{Var}(Y_i^T) + \sum_{q=1}^{Q} \frac{m_T^s}{m_T}\frac{m_T^s}{m_C^s}\,\mathrm{Var}(Y_j^C)\right\} \qquad \text{(VI)}$$

4.3 Results

The matching procedure aims to eliminate observable differences in the empirical distribution of variables. Table VI, in annex, shows how our target group is distributed before and after the matching. In general, the matching procedure does not significantly alter the distribution of characteristics for VE students, with the exception of age, and it changes slightly the distribution of GE characteristics for age, grades of the national exams in t-1 and for the academic qualifications of the guardian of education. These descriptive also allowed to understand further the specificities of our target group, as VE students present distinctive characteristics from GE students: they tend to be older, have modest grades in national exams (t-1) and their guardians of education have lower levels of academic qualifications.

VE students present 51% of the valid observations older than 15 years, a percentage that drops to 16,7% for GE students. After the match, VE students older than 15 represent 41,9% of the data and 18,9% of GE students, meaning that the weight of older VE students is now lower. Besides age, the distribution of characteristics of VE students is not substantially changed after matching. Therefore, we can safely extend the obtained results to the whole target group of VE.

With regard to GE students, it is interesting to note that, after matching, the percentage of students with grades lower than 50% increases from 33,6% to 50,9%. The same dynamic occurs in the grade of the national exam of Portuguese (t-1). In other words, as could be expected, a relevant proportion of GE students do not have a match among VE students, and this is particularly so among GE students with high academic performances. Similarly, the distribution of academic qualifications of VE students' guardians changes significantly after matching: the proportion of guardians with no more than the basic level of schooling increases from 55,6% to 70,1%

After the application of the matching procedure we produced the results, present in Table V (next page). We observe that, for the group under analysis, in general, VE presents positive and statistically significant impacts associated with transition, conclusion and performance in the labor market, and negative statistically significant impacts in access to HE.

Regarding students' performance in secondary education, the impact in transition presents an upward trend with the year of studies, with 23,5 p.p. in the transition from t to t+1 (from the 10th grade to the 11th), 31,1 p.p. from the second to the final year (from the 11th grade to the 12th) and, finally, with the value of 36,2 p.p. in the conclusion of this cycle within two years. The increasing value of the impacts with the year of studies is not especially surprising, given that goes in line with the percentage of failure, which also tends to increase with the year of studies.

Table V – Results of CEM

Output variable	% of success of T[(a)]	% of success of NT[(b)]	Impact (p.p.)	Std.	T-statistic
Transition in t	87,58	64,06	23,52	0,0014	16945,69
Transition in t and t+1	82,73	51,54	31,19	0,0015	20146,82
Conclusion in t+2	65,34	29,08	36,26	0,0019	19311,33
Dropout in t or t+1	6,74	7,90	-1,16	0,0005	-2357,74
Enrollment in HE after t+2	15,59	27,49	-11,89	0,0016	-7466,70
Employed during t+3	52,59	28,11	24,48	0,0246	535,84
Average daily salary[(c)]	19,8	18,9	0,90[(d)]	0,1021	8,78
Average number of months worked during t+3[(c)]	6,9	6,0	1,01[(d)]	0,0495	20,33
Average number of days worked per month[(c)]	25,7	24,0	1,58[(d)]	0,0570	27,52

(a) Corresponds to the percentage of the variable that equals one for treated observations
(b) Corresponds to the percentage of the variable that equals one for non-treated observations
(c) For students employed during t+3, that concluded secondary education and did not enrolled in higher education
(d) These values correspond to continuous impacts, not p.p.
Source: Own elaboration.

The impact of VE in the dropout on the first and second year of the secondary cycle is much lower than the impacts of other indicators of performance. VE has a learning structure that is more flexible than GE because it uses modules in the teaching system. If students do not complete the modules in the first moment of evaluation there is possibility of an extra evaluation. Because of this

incentive in transition, which in most of times results in a positive outcome, the incentive to dropout is lower. Thus, the low impact of this variable can be explained by the interaction between these two variables.

In the access to HE, we estimate that VE has an impact of, approximately, -12 p.p. The normative implications of these results are not clear. On one hand, they translate the main objective of VE's courses – to help the transition from the secondary cycle to the labor market – so a low impact on this variable would be expected. On the other hand, there are courses of higher level in which students from the vocational and professional track can enroll (e.g., technological specialization courses of level IV, present in universities and polytechnic institutions), and the return to HE in Portugal is still considerable high (see Badescu, *et al*. (2011)), so this situation may be a source of concern, as the policy may be influencing too much students' future access to HE.

Regarding students' performance in the labor market, we found that VE has an impact of 24,5 p.p. in the probability of employment in the year following the conclusion of the secondary cycle (during t+3)[44]. This impact is not especially surprising because, as presented above, it translates the main purpose of VE. The impact in the average daily salary is, approximately, of one euro per day (corresponding to 30 euros in the monthly remuneration). The impact on the average number of months worked during t+3 is, approximately, one month and, finally, the impact in the average number of days worked per month is, approximately, of 1,6 days (corresponding to an annual impact 19 days).

These results suggest that VE has, in general, for the group under analysis, relevant impacts in students' academic performance, increasing the changes of transition and conclusion of this cycle (while having a small impact on the percentage of dropout) and, on the other hand, reducing students' likelihood of proceeding to higher education. In respect to the labor market, there are, in general, positive impacts in the probability of employment and moderate impacts in the variables of job quality. Considering that the main purpose of VE is to help students' transition into labor

[44] Only students that graduated from secondary school and did not proceed to HE were considered in the estimation

market and to decrease the dropout from the secondary cycle, we can confirm that the education policy is reaching its original objectives for this group.

4.4 Sensitivity analysis

To perform sensitivity analysis we used the CEM methodology with a different coarsening, with the objective of keeping observations that in the basic coarsening would be excluded (categories with few observations). The following variables were subjected to a different coarsening, namely:

- Age: the first two years were merged (the new categories are 15 and 16, 17 and 18);
- Guardian's academic qualifications: the lowest and the highest academic levels were merged, resulting in the following categories - none, 1st or 2nd cycle; 3rd cycle; secondary; higher;
- Scholar social support in t-1: the level of social support (A, B/C, none) was transformed into a dichotomous variable (yes/no);
- School's average number of students per employee: this variable was converted into a dichotomous variable (less than 30 students/ 30 or more);
- School's proportion of Portuguese students: this variable was converted into a dichotomous variables (less than 95%, 95% or more).

The alternative coarsening allowed increasing the number of valid observations and, therefore, the number of homogenous groups (see Table VI, below):

Table VI – Distribution of matched data: main model and alternative coarsening

	Main model		**Alternative coarsening**	
	N.	**%**	**N.**	**%**
Vocational education (VE)	20.086	26,85	24.570	25,26
General education (GE)	54.728	73,15	72.694	74,74
Total	74.814	100	97.264	100

Note 1: includes the three cohorts under analysis.
Note 2: the number of valid observations may change according to the outcome variable (because of inconsistency criteria). The present values refer to the variable "Dropout in t or t+1"

After the application of the alternative coarsening we applied the same methodology, obtaining very similar results as in the main method (see Table VII, below).

Table VII – Results of sensitivity analysis with CEM

Output variable	% of success of T[(a)]	% of success of NT[(b)]	Impacts (p.p.)	Impacts – Basic CEM	Std.	t-statistic
Transition in t	86,66	63,49	23,16	23,52	0,0011	21638,68
Transition in t and t+1	81,95	50,11	31,84	31,19	0,0012	25814,67
Conclusion in t+2	63,33	29,05	34,28	36,26	0,0015	23559,05
Dropout in t or t+1	7,02	8,75	-1,73	-1,16	0,0006	-2688,02
Enrollment in HE after t+2	14,96	26,60	-11,63	-11,89	0,0013	-9247,49
Employed during t+3	52,90	27,64	25,26	24,48	0,0038	6614,97
Average daily salary[(c)]	19,85	18,95	0,81[(d)]	0,90[(d)]	0,0715	11,29
Average number of months worked during t+3[(c)]	6,81	6,05	0,84[(d)]	1,01[(d)]	0,0364	23,12
Average number of days worked per month[(c)]	25,59	23,95	1,57[(d)]	1,58[(d)]	0,0570	27,52

(a) Corresponds to the percentage of the variable that equals one for treated observations
(b) Corresponds to the percentage of the variable that equals one for non-treated observations
(c) For students employed during t+3, that concluded secondary education and did not enrolled in HE
(d) These values correspond to continuous impacts, not p.p.
Source: Own elaboration.

5. Conclusions

In 2004, as a consequence of a reform in the Portuguese education system, VE programs were introduced in public high schools. Since that reform there was a substantial increase in the number of students enrolled in VE, especially during 2008-2011. This thesis focused on the impacts of such expansion in students' academic and labor market performance. For that purpose we implemented a Counterfactual Impact Evaluation (CIE) logic and a Coarsened Exact Matching (CEM) methodology. Our analysis focuses on a very specific group: Portuguese students, with no special educational needs, ages between 15 and 18, were enrolled in GE in the last year (basic education) and are enrolled in public high schools under tutelage of the MES. Regarding the characteristics of VE students, we found that these students tend to be older, have a lower performance in the National exams at lower secondary education, and their guardians have lower levels of education attainment.

We found that, for the group under analysis, in general, VE presents positive and statistically significant impacts associated with transition, conclusion and performance in the labor market, and negative statistically significant impacts for dropout and access to higher education. In respect to the labor market, there are, in general, positive impacts in the probability of employment and moderate positive impacts in different variables of job quality. He found that our results are in line with the ones of Cappellari (2004), Polidano & Tabasso (2014) and with the Portuguese case, Oliveira (2015)[45].

Considering that the main purpose of VE is to help students' transition into labor market and to decrease the dropout from the secondary cycle, we can confirm that the education policy is reaching the original objectives for this group. Related to our specific focus group, VE also helps to break the effect of intergenerational level of academic qualifications, a strong factor in the Portuguese education system (recall that around 70% of these students have parents with academic qualifications lower than the secondary level). Given that most of the literature focuses on VE

[45] Although, for this study our results are only in line with the author's findings that the wage after graduation is higher for VE than to GE

impacts in the labor market, our study also had the important contribution to the literature of VE's impacts in school performance.

For future research, some aspects could be further analyzed. Given the richness of data, the assessment of impacts could be extended to analysis by gender, specific VE programs, and, finally, by applying a different matching method.

6. References

Alet, E., & Bonnal, L. 2011. ***Vocational schooling and educational success: Comparing apprenticeship to full-time vocational high school***. Mimeo: Toulouse School of Economics.

Arum, R., & Shavit, Y. 1995. Secondary Vocational Education and the Transition from School to Work. ***Sociology of Education***, *68*(3), 187–204.

Azevedo, J. 2014. Ensino Profissional. In M. de L. Rodrigues (Ed.), ***40 anos de políticas de Educação em Portugal - Volume I - A construção do sistema democrático de ensino***. Coimbra: Almedina.

Badescu, M., Villalba, E., & D'Hombres, B. 2011. ***Returns to education in European countries. Evidence from the European Community Statistics on Income and Living Conditions (EU-SILC)***. Italy. doi:10278831157

Becker, S. O. 2009. ***Methods to estimate causal effects. Theory and applications.*** Stirling: University of Stirling. Retrieved from http://www.restore.ac.uk/Longitudinal/surveynetwork/documents/master_class_notes.pdf

Bishop, J. H., & Mane, F. 2004. The impacts of career-technical education on high school labor market success. ***Economics of Education Review***, *23*(4), 381–402. doi:10.1016/j.econedurev.2004.04.001

Boesel, D., Hudson, L., Deich, S., & Masten, C. 1994. ***National Assessment of Vocational Education, Final Report to Congress. Volume II: Participation in and Quality of Vocational Education***. Washington, DC: U.S. Government Printing Office.

Cappellari, L. 2004. **High School Types, Academic Performance and Early Labour Market Outcomes**. IZA Discussion Paper no. 1048.

CEDEFOP. 2007. ***Portugal. Overview of the Vocational Education and Training System in 2007.*** Lisbon.

Eichhorst, W., Rodríguez-planas, N., Schmidl, R., & Zimmermann, K. F. 2015. A Roadmap to Vocational Education and Training Systems Around the World. ***ILR Review***, *68*(2), 314–337. doi:10.1177/0019793914564963

Field, S., Hoeckel, K., Kis, V., Justesen, T., & Kuczera, M. 2009. ***Learning for Jobs. OECD Reviews of Vocational Education and Training: Initial Report. Organisation for Economic and Co-Operation and Development (OECD)***. Paris.

Grácio, S. 1989. ***Política Educativa como Tecnologia Social. As reformas do ensino técnico de 1948 e 1983***. Lisboa: Horizonte.

Hanushek, E. A., Woessmann, L., & Zhang, L. 2011. ***General Education, Vocational Education, and Labor-Market Outcomes over the Life-Cycle***. National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17504

Ho, D. E., Imai, K., King, G., & Stuart, E. A. 2007. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. ***Political***

Analysis, *15*(3), 199–236. doi:10.1093/pan/mpl013

Iacus, S. M., King, G., & Porro, G. 2011. Causal inference without balance checking: Coarsened exact matching. ***Political Analysis***, *20*(1), 1–24. doi:10.1093/pan/mpr013

Jardim, J. C. 2000. O sistema de formação profissional. In J. M. Adams & R. Nobauer (Eds.), ***O Sistema de Formação Profissional em Portugal***. Lisboa: CEDEFOP.

Lee, W.-S., & Coelli, M. B. 2010. The Labour Market Effects of Vocational Education and Training in Australia. ***Australian Economic Review***, *43*(4), 389–408. doi:10.1111/j.1467-8462.2010.00609.x

Meer, J. 2007. Evidence on the returns to secondary vocational education. ***Economics of Education Review***, *26*(5), 559–573. doi:10.1016/j.econedurev.2006.04.002

Neves (Coord.), A., Pereira, C., Santos, F., Godinho, R., & Pereira, T. 2009. ***Avaliação Externa do Impacto da Expansão dos Cursos Profissionais no Sistema Nacional de Qualificações***. Lisboa.

Oliveira, S. 2015. ***Returns to vocational education in Portugal***. Policy Paper, (3), 1–26.

Organisation for Economic Co-operation and Development. 2014. ***Education at a Glance 2014: OECD indicators*.** Paris: Organisation for Economic Co-operation and Development. doi:10.1787/19991487

Parey, M. 2009. ***Vocational Schooling versus Apprenticeship Training — Evidence from Vacancy Data***. University of Essex and Institute for Fiscal Studies, 43.

Pereira (Coord.), C., Godinho, R., Estevão, P., Calado, A., Santos, F., & Pereira, T. 2011. ***Estudo de Avaliação Externa dos Percursos Pós-Formação dos Diplomados de Cursos Profissionais no Contexto da Expansão desta Oferta no Sistema Nacional de Qualificações***. Lisboa.

Polidano, C., & Tabasso, D. 2014. Making it real: The benefits of workplace learning in upper-secondary vocational education and training courses. ***Economics of Education Review***, *42*, 130–146. doi:10.1016/j.econedurev.2014.06.003

Ryan, P. 2001. The School-to-Work Transition: A Cross-National Perspective. ***Journal of Economic Literature***, *39*(1), 34–92.

Schlotter, M., Schwerdt, G., & Woessmann, L. 2011. Econometric methods for causal evaluation of education policies and practices: a non-technical guide. ***Education Economics***, *19*(2), 109–137. doi:10.1080/09645292.2010.511821

Shavit, Y., & Muller, W. 2000. Vocational Secondary Education. ***European Societies***, *2*(1), 29–50. doi:10.1080/146166900360710

Silva, M. M. (1998). ***A educação escolar em mudança. Volume II*.** Lisboa: Universidade Técnica de Lisboa, ISCSP.

Wolter, S. C., & Ryan, P. 2011. Aprenticeship. In A. Eric, S. J. M. Hanushek, & L. Woessmann (Eds.), ***Handbook of the Economics of Education, Volume 3*** (pp. 521–576). Amsterdam: North Holland.

7. Legislation

Decree-Law Nº. 47/430 of 2 of January 1967

Legislative Order N.º 194-A/1983 of 21 of October

Decree-Law N.º 26/89 of 21 of January

Decree-Law N.º 286/89 of 26 of August

Decree-Law N.º 74/2004 of 26 of March

Amendment N.º 44/2004 of 25 of May

Decree-Law N.º 24/2006 of 6 of February

Amendment N.º 23/2006 of 7 of April

Decree-Law N.º 35/2007 of 15 of February

Decree-Law N.º 41/2007 of 21 of February

Decree-Law N.º 272/2007 of 26 of July

Legislative Order N.º 74-A of 2008

Decree-law nº.139/2012 of 25 of July

Decree-Law N.º 92/2014 of 20 of June

Ordinance nº.165-B/2015 of 3 of July

8. Annexes

Table I - Professional areas of Vocational courses

Entertainment and performing arts	Forestry and hunting
Agricultural and animal production	Framework of business
Protection of persons and property	Fisheries
Audiovisual and multimedia production	Food industry
Construction and repair of motor vehicles	Metallurgy and metal
Electronics and automation	Accounting and taxation
Tourism and leisure	Computer sciences
Secretariat and administrative work	Philosophy, history and related sciences
Hotels and restaurants	Crafts
Journalism	Library, archive and documentation
Chemical engineering	Trading
Textile, clothing, footwear and leather	Electricity and energy
Floriculture and gardening	Materials (wood, paper, plastic, glass and others)
Extractive industries	Architecture and urbanism
Civil construction	Marketing and advertising
Health services	Support services for children and young people
Dental sciences	Finance, banking and insurance
Social work and guidance	Management and administration
Environmental protection	Design
Safety and health at work	

Source: ANQEP website, http://www.anqep.gov.pt/

Table II - Formation components and workload of VE

Formation Components	Subjects	Total hours[(a)]
Sociocultural	Portuguese	320h
	Foreign language I, II or III[(b)]	220h
	Integration Area	220h
	Information and communication technologies	100h
	Physical education	140h
Scientific	2 to 3 subjects[(c)]	500h
Technical	3 to 4 subjects[(d)]	180h
	Workplace training component [(e)]	420h
Total workload of Course		3100h

(a) Total workload not compartmentalized by the three years of formation cycle, to be managed by the school, considering its pedagogical autonomy while being aware of the total annual workload and in order to optimize the global management of modules and workplace training;
(b) Student chooses one foreign language. If student was taught one foreign language in basic education it is obligated to learn a second language in secondary education;
(c) Basic scientific subjects fixed by own regulation in function of the intended professional qualification;
(d) Subjects of technological, technical and practical nature;
(e) The workplace training component aims to develop and acquire technical, relational and organizational skills relevant to the intended professional qualification and is subject to own regulation;

Source: Decree-Law Nº. 74/2007 of 26 of March.

Table III – Students enrolled in Secondary Education: total (public and private) and by type of education

Years	Type of Education						
	Total	Tech. Course[(1)]	GE	App.[(2)]	ETC[(3)]	VE	Recurrent and others
2000	417.705	69.029	265.601	//	//	29.100	53.975
2001	413.748	65.971	242.452	//	//	30.668	74.657
2002	397.532	59.286	224.641	//	//	33.799	79.806
2003	385.589	54.975	214.242	//	2.353	33.587	80.432
2004	382.212	53.831	212.927	//	2.877	34.399	78.178
2005	376.896	60.697	206.133	//	2.832	36.765	70.469
2006	347.400	53.384	188.764	//	3.422	36.943	64.887
2007	356.711	44.532	196.149	//	5.224	47.709	63.097
2008	349.477	27.361	196.337	//	8.425	70.177	47.177
2009	498.327	22.039	195.688	13.584	4.388	93.438	169.190
2010	483.982	16.543	197.711	17.619	2.320	107.266	142.523
2011	440.895	15.288	198.085	18.669	2.117	110.462	96.274
2012	411.238	12.296	199.321	21.056	2.012	113.749	62.804
2013	398.447	8.220	201.336	33.366	3.025	115.885	36.615

(1) Technological course
(2) Apprenticeship
(3) Education and Training Course
Source: Data from School Census - DGEEC/MES in Pordata

Table IV – Inconsistency criteria for control and output variables

Control Variable	Inconsistency criteria
Sex	Cannot change between t-1 and t
Age	Cannot be lower than 14 years at the 10th grade
Nationality	Cannot change between t-1 and t
Guardians' nationality	Cannot change between t-1 and t
Guardians' academic background	The level of qualifications cannot decrease between t-1 and t
Who is the guardian	If student is the Guardian he cannot have qualifications superior to the 10th grade
Basic education	If student was enrolled in basic education in t-1, then the curricular year must have been the 9th grade
Basic education	If student was enrolled in basic education in t-1, then the number of enrolments in t must be one
Curricular year	If student was enrolled in the 10th grade at t, then in t+1 cannot be enrolled in a lower curricular year
Curricular year	If student was enrolled in the 10th grade at t+1, then in t+2 cannot be enrolled in a lower curricular year
Conclusion at t+2	If student completed secondary education in t+2, then in t+1 he must have been in the 11th grade
Conclusion at t+2	If student completed secondary education in t+2, then in that year he must have been enrolled in the 12th grade
Conclusion at t+2	If student completed secondary education in t+2, then he cannot be enrolled in t+3
Conclusion at t+3	If student completed secondary education in t+3, then in t+2 he must have been enrolled in the 12th grade
Conclusion at t+3	If student completed secondary education in t+3, then he must have been enrolled in the same course in t+2 and t+3

Output Variable	Inconsistency criteria
Higher education	If student was enrolled in higher education in t+5, then he must have been enrolled in the 12th grade in t+2 or t+3
Higher education	If student was enrolled in higher education in t+3, then he must have been enrolled in the 12th grade in t+2
Higher education	If student was enrolled in higher education in t+3, then he must have been enrolled in the 11th grade in t+1
Higher education	If student was enrolled in the 10th grade in t, then he cannot be enrolled in higher education in t+4 or t+5
Higher education	If student was enrolled in higher education in t+3, then he cannot be enrolled in secondary education in that year
Higher education	If student was enrolled in the 10th grade in t, then he cannot be enrolled in higher education in t+2
Higher education	If student was enrolled in higher education in t+4, then he must have been enrolled in the 12th grade in t+2 or t+3
Higher education	If student was enrolled in higher education in t+4, then he must have been enrolled in the 11th grade in t+1 or t+2

Source: Own elaboration

Table V – Distribution of control variables by valid and invalid data

Sex	Valid		Invalid	
	N.	%	N.	%
VE (1)	25.792	100	6.148	100
Male	14.048	54,5	3.522	57,3
Female	11.744	45,5	2,626	42,7
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
Male	51.429	43,3	9.467	46
Female	67,265	56,7	11.107	54
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Age	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
15	12.637	49	2.075	33,8
16	8.348	32,4	2.077	33,8
17	3.819	14,8	1.517	24,7
18	988	3,8	479	7,8
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
15	98.871	83,3	16.076	78,1
16	14.874	12,5	3.261	15,9
17	4.171	3,5	1.007	4,9
18	778	0,7	230	1,1
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Nationality	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
Portugal	25.792	100	6.148	100
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
Portugal	118.694	100	20.574	100
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Student has Special Education Needs in t?

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
No	25.792	100	6.148	100
Yes	0	0	0	0
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
No	118.694	100	20.574	100
Yes	0	0	0	0
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Student has Special Education Needs in t-1?

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
No	25.792	100	6.148	100
Yes	0	0	0	0
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
No	118.694	100	20.574	100
Yes	0	0	0	0
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Type of education in t-1

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
Basic education - Regular	25.792	100	6.148	100
n.d.	0	0	0	0
GE (2)	118.694	100	20.574	100
Basic education - Regular	118.694	100	20.574	100
n.d.	0	0	0	0
Total (1) + (2)	144.486	-	26.722	-

Level of Scholar Social Support in t-1

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
None	13.496	52,5	3.308	53,8
Level A	7.182	28	1.709	27,8
Level B/C	5.114	19,5	974	15,8
n.d.	0	0	157	2,6
GE (2)	118.694	100	20.574	100
None	85.806	72,3	12.501	60,8
Level A	16.876	14,2	2.966	14,4
Level B/C	16.012	13,5	2.104	10,2
n.d.	0	0	3.003	14,6
Total (1) + (2)	144.486	-	26.722	-

Guardians' academic qualifications

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
Up to 1st cycle	8.302	32,3	1.685	27,4
2nd cycle	7.859	30,2	952	15,5
3rd cycle	5.083	19,6	678	11
Secondary	2.962	11,4	333	5,4
Bachelor	809	3,2	127	2,1
Postgraduate	73	0,3	9	0,1
Others	704	3	152	2,5
n.d.	0	0	2.212	36
GE (2)	118.694	100	20.574	100
Up to 1st cycle	17.549	14,8	2.679	13
2nd cycle	23.961	20,1	1.840	8,9
3rd cycle	24.546	20,7	2.237	10,9
Secondary	27.066	22,8	2.031	9,9
Bachelor	21.775	18,4	1.950	9,5
Postgraduate	2.100	1,8	143	0,7
Others	1.697	1,5	140	0,7
n.d.	0	0	9.554	46,4
Total (1) + (2)	144.486	-	26.722	-

Student has computer at home in t?

	Valid		Invalid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
No	10.742	41,6	3.006	48,9
Yes	15.050	58,4	3.142	51,1
GE (2)	118.694	100	20.574	100
No	39.974	33,7	9.322	45,3
Yes	78.720	66,3	11.252	54,7
Total (1) + (2)	144.486	-	26.722	-

Grade of National exam of Portuguese in t-1

	Valid		Non-valid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
<=50%	13.698	53,1	2.092	34
[51-74%]	11.443	44,4	1.822	29,6
[75-100]%	651	2,5	95	1,5
n.d.	0	0	2.139	34,8
GE (2)	118.694	100	20.574	100
<=50%	25.091	21,1	4.189	20,4
[51-74%]	67.612	57	10.571	51,4
[75-100]%	25.991	21,9	3.653	17,8
n.d.	0	0	2.161	10,5
Total (1) + (2)	144.486	-	26.722	-

Grade of National exam of Mathematics in t-1

	Valid		Non-valid	
	Nº	%	Nº	%
VE (1)	25.792	100	6.148	100
<=50%	17.973	69,7	2.879	46,8
[51-74%]	6.810	26,4	1.072	17,4
[75-100]%	1.009	3,9	156	2,5
n.d.	0	0	2.041	33,2
GE (2)	118.694	100	20.574	100
<=50%	39.826	33,6	6.666	32,4
[51-74%]	44.496	37,5	7.068	34,4
[75-100]%	34.372	29	4.857	23,6
n.d.	0	0	1.983	9,6
Total (1) + (2)	144.486	-	26.722	-

Table VI – Distribution of control variables before and after matching

Sex

	Valid data (%)	Matched data (%)
VE	100	100
Male	54,50	52,90
Female	45,50	47,10
GE	100	100
Male	43,30	46,70
Female	56,70	53,30

Computer at home

	Valid data (%)	Matched data (%)
VE	100	100
Yes	41,6	39,1
No	58,4	60,9
GE	100	100
Yes	33,7	31,9
No	66,3	68,1

Grade of final exam of Mathematics (t-1)

	Valid data (%)	Matched data (%)
VE	100	100
<=50%	69,7	71,3
[51-74]%	26,4	25,4
[75-100]%	3,9	3,3
GE	100	100
<=50%	33,6	50,9
[51-74]%	37,5	40,4
[75-100]%	29,0	8,7

Grade of final exam of Portuguese (t-1)

	Valid data (%)	Matched data (%)
VE	100	100
<=50%	53,1	51
[51-74]%	44,4	47
[75-100]%	2,5	2
GE	100	100
<=50%	21,1	29,2
[51-74]%	57	66,6
[75-100]%	21,9	4,2

Age

	Valid data (%)	Matched data (%)
VE	100	100
15	49,0	58,1
16	32,4	31,6
17	14,8	9,3
18	3,8	1,0
GE	100	100
15	83,3	81,1
16	12,5	15,1
17	3,5	3,5
18	0,7	0,3

Scholar Social Support

	Valid data (%)	Matched data (%)
VE	100	100
None	52,5	56,6
Level A	28,0	25,8
Level B/C	19,5	17,6
GE	100	100
None	72,3	71,9
Level A	14,2	15,4
Level B/C	13,5	12,8

Guardians' academic qualifications

	Valid data (%)	Matched data (%)
VE	100	100
Up to 1st cycle	32,3	32,4
2nd cycle	30,2	31,7
3rd cycle	19,6	20,3
Secondary	11,4	11,7
Bachelor	3,2	2,7
Postgraduate	0,3	0,1
Others	3,0	1,1
GE	100	100
Up to 1st cycle	14,8	19,8
2nd cycle	20,1	26,8
3rd cycle	20,7	23,5
Secondary	22,8	21,3
Bachelor	18,4	8,0
Postgraduate	1,8	0,1
Others	1,5	0,5