UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

The impact of peer effects, substance use, and parental background on educational outcomes

PhD Thesis submitted to the Faculty of Economics and Business

Institute of Economic Research

University of Neuchâtel

For the degree of PhD in Economics

by

Lionel PERINI

Accepted by the dissertation committee:

Prof. Milad Zarin-Nejadan, University of Neuchâtel, thesis director

Prof. Mehdi Farsi, University of Neuchâtel

Prof. Christian Suter, University of Neuchâtel

Dr. Jean-Marc Falter, Swiss National Bank

Defended on 23th May 2014

Faculté des Sciences Economiques
Avenue du 1er-Mars 26
CH-2000 Neuchâtel
www.unine.ch/seco

UNIVERSITÉ DE
NEUCHÂTEL

IMPRIMATUR POUR LA THÈSE

The impact of peer effects, substance use, and parental background on educational outcomes

Lionel PERINI

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

La Faculté des sciences économiques,
sur le rapport des membres du jury

Prof. Milad Zarin (directeur de thèse, Université de Neuchâtel)
Prof. Mehdi Farsi (président du jury, Université de Neuchâtel)
Prof. Christian Suter (FLSH, Université de Neuchâtel)
Prof. Jean-Marc Falter (Banque Nationale Suisse et Université de Genève)

Autorise l'impression de la présente thèse.

Neuchâtel, le 25 septembre 2014

Le doyen
J-M Greher
Jean-Marie Grether

■ Téléphone : +41 32 718 15 00 ■ Fax : +41 32 718 15 01 ■ E-mail : secretariat.seco@unine.ch ■ www.unine.ch

To my family

Summary

This doctoral thesis investigates three issues in economics of education applied to Switzerland which is an interesting case for analyzing topics which refer to education on the grounds that having a highly-qualified workforce is a central element for Swiss competitiveness. As Switzerland has no - or few - natural resources or raw materials, priority has been given to education. If the country is still competitive with respect to larger industrialized countries in fields like sciences, technology or finance, this is mainly due to the quality of its education system which ensures an excellent capacity for innovation and a sophisticated business culture. As a consequence, Switzerland tries to preserve this advantage, which implies a constant reassessment of educational policies. In this setting, this dissertation focuses on three important factors, namely peers effects, substance use and parental background, whose considerations may have an impact on how the Swiss education system works and can be improved.

The first essay of this thesis discusses peer effects at the lower secondary level. At this stage, pupils tend to be strongly influenced by their classmates in the learning process and the nature of these social interactions can give precious insights in the debates on school tracking policies. The objective of the study is to determine if the introduction of a completely non-selective school system in the Swiss education landscape could lead to efficiency and equity gains compared to the current tracking system where students are separated according to their school performances. The identification strategy relies on ability track fixed effects to control for within-school sorting and quantile regression methods to account for heterogeneity issues.

The end of compulsory schooling coincides with a period where adolescents are increasingly exposed to the consumption of addictive products through peers' influence, risky health behaviour or time preferences. For that purpose, the second essay - in collaboration with Joachim Marti of Leeds University - analyzes the relationship between cannabis use and dif-

ferent short-term educational outcomes at the upper secondary level because lifetime cannabis consumption among Swiss teenagers is unusually high in international comparison. We consider a lagged measure of substance use to reduce reverse causality between health and education and individual fixed effects to rule out selectivity effects.

After high school graduation, most students pursue their schooling path at the tertiary level. As individuals differ in their background characteristics (e.g., innate ability, socioeconomic status, or family support), the return to higher education is not expected to be the same for each student. But what types of students benefit most from university education? Using propensity score matching methods, the third essay focuses on the relationship between the predicted probability to complete university education and returns to schooling to analyze if completing a university degree complements or substitutes family background characteristics in generating earnings capability.

Keywords: peer effects, quantile regression, substance use, propensity score matching, returns to education

Contents

Foreword **13**

1 Introduction **14**
1.1 Overview 14
1.2 The Swiss educational system 17
1.3 Presentation of the three essays 20
1.3.1 Peer effects in Swiss lower secondary schools 20
1.3.2 Cannabis use and short-term academic performances . 28
1.3.3 Who benefits most from university education in Switzerland? 34
1.4 References 41

2 Peer effects in Swiss lower secondary schools **51**
2.1 Introduction 51
2.2 Background and literature review 54
2.2.1 Choice of the reference group 54
2.2.2 Identification strategies 54
2.2.3 Efficiency and equity considerations 57
2.3 Data 59
2.3.1 PISA national sample 59
2.3.2 Educational outcomes 61
2.3.3 School track design and ability tracks 61
2.3.4 Peer characteristics variables 63
2.3.5 Control variables 63
2.4 Empirical analysis 65
2.4.1 Identification of mean peer effects 66
2.4.2 Efficiency analysis 66
2.4.3 Equity analysis 67
2.5 Results 68
2.5.1 Mean peer effects 68
2.5.2 Efficiency 69

2.5.3 Equity . . . 73
2.6 Conclusion . . . 79
2.7 References . . . 80
2.8 Appendices . . . 84
2.8.1 School track designs . . . 84
2.8.2 Additional tables . . . 85

3 Cannabis use and short-term academic performances† 91
3.1 Introduction . . . 91
3.2 Data . . . 94
3.3 Empirical approach . . . 99
3.3.1 Pooled OLS and fixed effects . . . 99
3.3.2 Propensity score matching . . . 100
3.4 Results . . . 102
3.4.1 Pooled OLS and fixed effects . . . 102
3.4.2 Propensity score matching . . . 104
3.5 Discussion . . . 107
3.6 References . . . 108
3.7 Appendices . . . 112
3.7.1 Distribution of the propensity score . . . 112
3.7.2 Additional tables . . . 113

4 Who benefits most from university education in Switzerland? 116
4.1 Introduction . . . 116
4.2 Theoretical background and literature review . . . 120
4.2.1 Theoretical framework . . . 120
4.2.2 Returns to schooling in Switzerland: a short literature review . . . 122
4.3 Data . . . 124
4.4 Empirical framework . . . 128
4.4.1 Homogeneous returns to education . . . 128
4.4.2 Heterogeneous returns to education . . . 129
4.5 Results . . . 131
4.5.1 Homogeneity assumption . . . 131
4.5.2 Heterogeneity assumption . . . 132
4.6 Auxiliary analysis . . . 139
4.7 Conclusion . . . 141
4.8 References . . . 144
4.9 Appendices . . . 149
4.9.1 NATE parameter . . . 149

4.9.2 Additional tables 150

5 Conclusions 153

List of Tables

2.1 A selective review of closely related papers 60
2.2 Variables description and summary statistics 64
2.3 Mean peer effets, OLS estimation 70
2.4 Mean peer effects, ability track FE estimation 71
2.5 Nonlinear peer effects and peer heterogeneity, ability track FE estimation 72
2.6 Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE 74
2.7 Peer heterogeneity and educational opportunities, OLS estimation 76
2.8 Peer heterogeneity and educational opportunities, Quantile regressions 77
2.9 School track designs 84
2.10 Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE, Reading 85
2.11 Nonlinear peer effects and peer heterogeneity, Quantile regressions with ability track FE, Mathematics 86
2.12 Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE, Sciences 87
2.13 Peer heterogeneity and educational opportunities, Quantile regression, Reading 88
2.14 Peer heterogeneity and educational opportunities, Quantile regression, Mathematics 89
2.15 Peer heterogeneity and educational opportunities, Quantile regression, Sciences 90

3.1 A selective review of closely related papers 95
3.2 Descriptive statistics 98
3.3 OLS and FE models for both any and frequent cannabis use . 103
3.4 Logistic regressions, any and frequent cannabis use 105
3.5 Matching estimates and Rosenbaum bounds 106

3.6 Any cannabis use, OLS estimation 113
3.7 Frequent cannabis use, OLS estimation 114
3.8 Any cannabis use, FE estimation 115
3.9 Frequent cannabis use, FE estimation 115
3.10 Description of variables related to psychological traits 115

4.1 Literature on returns to schooling in Switzerland: a selective review 125
4.2 Description of variables and summary statistics 127
4.3 Homogeneous returns to university education, OLS estimation 132
4.4 Predicted probability to complete university education, Logistic regressions 133
4.5 Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE) 136
4.6 Heterogeneous returns to university education, Matching-smoothing estimates 139
4.7 Quantile regressions, Men and Women 142
4.8 Detailed information on the propensity score strata 150
4.9 Stratification-multilevel model, OLS estimation, Men 150
4.10 Stratification-multilevel model, OLS estimation, Women . . . 151
4.11 Stratification-multilevel model, Sample selection specification, Women 152

List of Figures

1.1 The Swiss educational system 19
1.2 Analytical framework between health and education 29
1.3 Substance use in Switzerland: an international comparison . . 32
1.4 Internal rates of return: an international comparison 36

2.1 Unconditional versus conditional test scores - Efficiency analysis 75
2.2 Unconditional versus conditional test scores - Equity analysis . 78

3.1 Distribution of the propensity score among cannabis users and non-users 112

4.1 Distribution of the propensity score among matura and university graduates 134
4.2 Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE), Men 137
4.3 Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE), Women 138
4.4 Heterogeneous returns to university education, Matching-smoothing model (MS-HTE) 140

Foreword

After graduating in economics, the idea of writing a doctoral thesis emerged when I was a student in the Master of Sciences in Statistics at University of Neuchâtel where I developed my interest for quantitative methods. Always interested by educational matters, the "III Summer School in Public Economics: Economics of education" which took place at the Faculty of Economics at University of Barcelona in 2010 was an excellent opportunity to improve my knowledge in this area. More precisely, the lectures given by Prof. Hessel Oosterbeek and Prof. Richard Romano were of primary importance to learn more about peer effects and returns to schooling. The availability of the Swiss longitudinal survey "Transition from Education to Employment" (TREE), coupled with the political debates on cannabis liberalization in Switzerland, led to my third contribution which is also related to health economics.

Most of this thesis was written while I was working at the Institute of Economic Research at University of Neuchâtel. I am grateful to my colleagues for their support and the fruitful discussions. I also thank Elke Luedemann, Sylvain Weber, Mehdi Farsi and anonymous referees for their helpful comments. More specifically, I am grateful to Prof. Milad Zarin-Nejadan in his role of thesis advisor for the careful reading of this document and for his relevant advices. Last but not least, I address specific thanks to my co-author and friend Joachim Marti whose collaboration was very instructive and appreciated.

I also would like to note that all remaining errors are my own.

Berne, September 2014

Lionel Perini

Chapter 1

Introduction

1.1 Overview

Education provides substantial value to individuals and the society in general. At the individual level, increasing educational attainment does not only improve earnings prospects but also some other aspects like well-being, prestige, joy of learning ("psychic earnings") or health behaviour. At the societal level, a high-skilled workforce generates productivity gains which translate into higher income for the economy. Well-educated people generate also positive externalities for the society through their social skills, ecological attitude or civic participation. The role of the education system is to increase the level of knowledge of the population, to develop the personality of individuals, to facilitate the socialization process and to encourage the transmission (and preservation) of culture. To summarize, education influences the society from different angles and requires adapted schooling policies.

This doctoral thesis investigates three issues in economics of education applied to Switzerland which is an interesting case for analyzing topics which refer to education on the grounds that having a highly-qualified workforce is a central element for Swiss competitiveness. As Switzerland has no - or few - natural resources or raw materials, priority has been given to education. If the country is still competitive with respect to larger industrialized countries in fields like sciences, technology or finance, this is mainly due to the quality of its education system which ensures an excellent capacity for innovation and a sophisticated business culture. As a consequence, Switzerland tries to preserve this advantage, which implies a constant reassessment of educational policies. In this setting, this dissertation focuses on three important factors, namely peers effects, substance use and parental background, whose considerations may have an impact on how the Swiss education system works and can be improved.

Since the 1960s, economics has significantly contributed to the theoretical framework in the field of education, e.g., human capital theory (Becker, 1964; Mincer, 1974) or signalling theory (Spence, 1973). After a short decline of interest in the 1980s (which also coincides with a reduction of state intervention), the resurgence of research on economics of education since two decades is mainly explained by an increasing demand from policy-makers to obtain quantifiable information to provide answers to policy questions and to justify resources allocation. In this context, the economics of education can help to understand how education might best be produced, how to improve social mobility and what are the monetary and non-monetary outcomes from education. At the same time, this period has been characterized by the emergence of new analytical tools to quantify the benefits of educational programs. More precisely, the recent econometric methods allow to place particular emphasis on establishing causality in economic analyses.

Different identification strategies are proposed in the literature on economics of education. A randomized experiment allows the greatest reliability and validity of parameter estimates given that individuals are randomly selected in the sample. In such a case, the control and treatment groups[1] are equivalent in terms of probability of selection. Any difference between the two groups is due to the treatment and not to differences in the assignment process. However, for different reasons - economic, ethical and logistical -, randomized experiments are not common in social sciences. When the sample is not random, e.g., due to observational data, estimating causal effects is challenging, especially due to the influence of unobserved factors. As a consequence, classical ordinary least squares (OLS) suffer from endogeneity biases. Explicitly, we cannot distinguish between accidental association and causation.

A traditional solution to handle unobserved characteristics is to rely on an instrumental variable (IV) strategy, i.e., finding a variable correlated with the endogenous variable (*inclusion* restriction) but not with the error term (*exclusion* restriction). However, finding an instrument which satisfies the second condition is very hard (Bound et al., 1995; Checchi, 2006)[2]. This

[1]The *control* group refers to a group of subjects who do not receive the factor under study and thereby serve as comparison group with the *treated* group for which treatment results are evaluated.

[2]Control functions (CF) are also concerned by this problem given that their modelling

explains why economists consider alternative identification strategies such as difference-in-differences (measuring differences in outcomes for two groups at two points in time, one of which has been subject to treatment and the other not), regression discontinuity (analyzing discontinuity in outcomes by estimating local treatment effects in a small neighbourhood defined by a forcing variable where observed units are supposed identical) or fixed effects methods (the unobserved school- or individual-specific component is considered as fixed, which purges the estimates from omitted time-invariant school or individual characteristics).

When data do not offer the possibility to use methods accounting directly for unobservables or requiring a longitudinal design, researchers generally impose a conditional independence assumption[3]. In this context, propensity score matching (PSM) - which consists in matching individuals with similar propensity scores and carrying out the analysis on the adjusted data - has clear advantages over parametric approaches such as OLS regression. First, it avoids serious mismatches between treated and control units by matching only similar cases[4]. Second, PSM does not require specifying the functional form of the outcome equation. However, the conditional independence assumption is untestable. Therefore, sensitivity or auxiliary analyses are helpful to test the robustness of the matching estimates.

The first essay of this thesis discusses peer effects at the lower secondary level. At this stage, pupils tend to be strongly influenced by their classmates in the learning process and the nature of these social interactions can give precious insights in the debates on school tracking policies. The objective of the study is to determine if the introduction of a completely non-selective school system in the Swiss education landscape could lead to efficiency and equity gains compared to the current tracking system where students are separated according to their school performances. The identification strategy relies on ability track fixed effects to control for within-school sorting and quantile regression methods to account for heterogeneity issues.

The end of compulsory schooling coincides with a period where adoles-

of the selection process relies on instrumental variables.

[3]Conditional independence assumption (CIA), also called unconfoundedness or ignorability, states that, conditional on an observed set of covariates, there are no unobserved elements that are associated with both treatment and dependent variables (Rosenbaum and Rubin, 1983).

[4]The common support assumption ensures that the range of propensities to be treated is the same for treated and control units.

cents are increasingly exposed to the consumption of addictive products through peers' influence, risky health behaviour or time preferences. For that purpose, the second essay - in collaboration with Joachim Marti of Leeds University - analyzes the relationship between cannabis use and different short-term educational outcomes at the upper secondary level because lifetime cannabis consumption among Swiss teenagers is unusually high in international comparison. We consider a lagged measure of substance use to reduce reverse causality between health and education and individual fixed effects to rule out selectivity effects.

After high school graduation, most students pursue their schooling path at the tertiary level. As individuals differ in their background characteristics (e.g., innate ability, socioeconomic status, or family support), the return to higher education is not expected to be the same for each student. But what types of students benefit most from university education? Using propensity score matching methods, the third essay focuses on the relationship between the predicted probability to complete university education and returns to schooling to analyze if completing a university degree complements or substitutes family background characteristics in generating earnings capability.

The rest of this introduction is structured as follows. Section 1.2 presents the structure of the Swiss education system. Section 1.3 describes the three essays of the dissertation. This section presents the theoretical background, discusses the Swiss context, positions the research questions, explains the identification strategies and summarizes the main findings of the corresponding contributions. References are given in section 1.4.

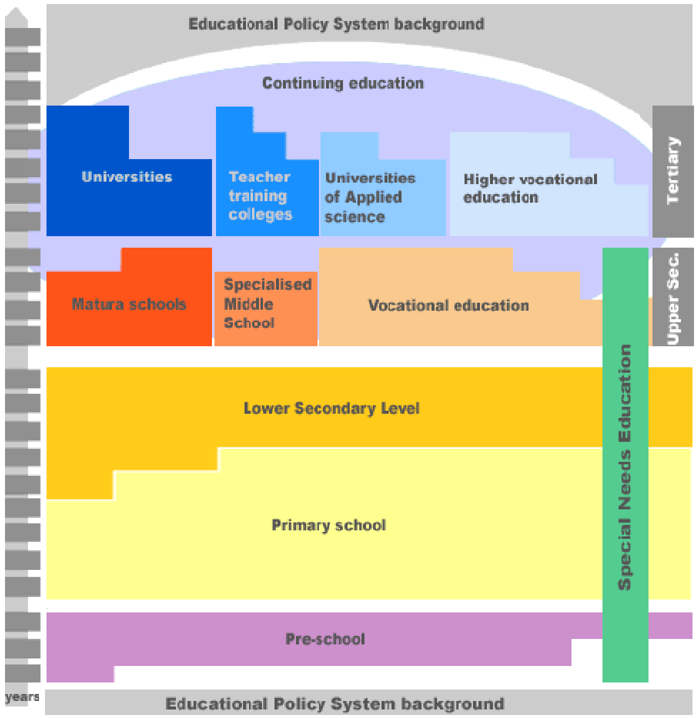
1.2 The Swiss educational system

The chapters of this thesis deal chronologically with three consecutive educational levels. It is therefore convenient to describe first the structure of the Swiss educational system and its peculiarities.

In Switzerland, the school system is extremely diversified because the Swiss Constitution delegates the main responsibility for education to the cantons (i.e., sub-national governments) and communes (i.e., municipalities) on the basis of the subsidiarity principle. Figure 1.1 synthesizes graphically the Swiss educational system. Cantons and their municipalities are responsible (in terms of law-making, financing and realization) for pre-primary, primary and lower secondary levels. At the upper secondary level, cantons are responsible for schools with high requirements (i.e., matura schools) whereas the Confederation (i.e., central government) regulates the field of vocational education. At the tertiary level, cantonal universities and universities of teacher education are under cantonal sovereignty whereas the two Swiss Federal Institutes of Technology, higher vocational schools and universities of applied sciences are legislated and financed by the Confederation.

Even if each entity has its own responsibilities, they are required to collaborate according to the Constitution. The cantonal sovereignty, however, renders the harmonization process difficult. For that purpose, a fourth entity called the Swiss Conference of Cantonal Ministers of Education (EDK) - which is composed of the twenty-six cantonal ministers of education - has been created in 1970 to ensure an intercantonal cooperation in terms of education. This political body - which lies somewhere between the cantonal and federal structures - employs different instruments to find solutions in key areas of education. The instruments take the form of agreements (legally binding), recommendations (not legally binding), or statements. The areas of interest include for instance harmonisation of compulsory schooling, coordination of language instruction and recognition of diplomas across the country.

Finding a balance between harmonization goals (according to the Constitution) and local needs (based on cultural and historical factors) is the main concern at the lower secondary level. For example, duration of studies and school selection procedures are subject to heterogeneous practices. For that purpose, the concordat *Harmos* came into force in 2009 and its main objective is to harmonize the cantonal school structures and meet objectives at the national level over a period of six years. To date, fifteen cantons have agreed on the concordat, seven have refused it and the last four have not yet taken a decision (CSRE, 2014). The concept of ability tracking, however, was not rediscussed in this concordat and still takes different forms across cantons and/or municipalities. By regrouping the different tracking procedures in a broader set of categories, we can identify three school designs. First, students can be sorted in different school types according their cognitive abilities (*separated* system). Each school type possesses its own curricula, teachers and sometimes range of subjects. To limit subdivisions at the lower secondary level which can be source of inefficiency and inequity, two different school designs emerged as alternatives to the traditional separated system. The *cooperative* system separates students in different ability tracks but within the same building while the *integrated* system mixes students in a comprehensive way except for core subjects where differentiated-level courses are proposed.



Source: http://www.fhnw.ch/business/about-us/fhnw-1/swiss-education-system

Figure 1.1: The Swiss educational system

It is worth mentioning that some cantons allow for a mix of the different systems mentioned above. Finally, decision basis regarding the placement into ability track takes different forms such as teacher's recommendations, parental endorsement, school performances at the end of primary school or testing.

The level of the ability track at the lower secondary level determines the schooling path at the upper secondary level. Students who attend higher-ability tracks are prepared for matura schools (i.e., high school) whereas middle- and lower-ability tracks prepare students for vocational education (e.g., professional matura or apprenticeship). In Switzerland, around 86% of the population between 25 and 64 years of age have at least an upper secondary degree which is well above the OECD average of 74% (OECD, 2012). The apprenticeship (also called dual vocational system) - which combines school-based and workplace-based education - offers one of the most successful transitions between education and labour market among OECD countries. After the lower secondary level, around 60% of students pursue this path whereas around 30% choose a matura school (OECD, 2009).

As a general rule, only students with an academic matura can enter into conventional universities. However, since 2005, students with a professional matura also have the possibility to attend these institutions through an instrument called “Passerelle Dubs”[5]. Nevertheless, few of them - less than 4% - take this opportunity (CSRE, 2014). Students with a professional matura are more prone to attend universities of applied sciences given that they have direct access to these institutions. Tertiary graduation rates in Switzerland are below the OECD average, i.e., 31% against 39%, respectively, essentially due to an efficient and attractive dual vocational system where students can rapidly gain financial autonomy (OECD, 2012). The international comparison, however, is somewhat biased given that some educational programs tertiarised in other countries are proposed at the upper-secondary level in Switzerland (CSRE, 2014). The high rate of scientific publications (relative to the Swiss population) and the high positions of Swiss universities in international rankings confirm the excellent reputation of the Swiss higher education system (OECD, 2009). The importance attached by Swiss authorities to good-quality higher education is reflected in the substantial amount of expenditure per tertiary student enrolled in public educational institutions (around 21,577 USD using PPPs in 2009) which is among the highest in OECD countries (OECD average: 13,728 USD using PPPs).

1.3 Presentation of the three essays

1.3.1 Peer effects in Swiss lower secondary schools

Theoretical background

Peer effects - also called social interactions - theory has for objective to capture the impact a social environment or group (e.g., neighbourhood, family, friends or classmates) can exert on individual behaviour. The starting point for academic research on this topic was the publication of the Coleman report (1966) which concludes that the composition of the student body is among the most important inputs explaining educational outcomes. The paper of Duncan et al. (1968) was also a pioneering contribution discussing social interactions in sociology. Economists have long been skeptical about

[5]Students have to pass additional exams after the professional matura. The length of the preparatory courses is around one year.

including social interactions in their discipline (Manski, 2000). While some considered economics essentially as the study of markets and prices, other defined it as a field where the concerns are more oriented toward allocation of resources and incentives. By the 1970s, the development of microeconometrics, labour economics and endogenous growth theory supported the fact that economics cannot be considered from a narrow viewpoint and gave the opportunity to social interactions to be introduced within economic analyses.

Initially identified in education, the concept of peer effects has then been extensively investigated in different fields such as substance use (Krauth, 2005; Fletcher, 2010; McVicar and Polanski, 2012; Moriarty et al., 2012), criminal activity (Bayer et al., 2009), juvenile behaviour (Gaviria and Raphael, 2001), obesity (Trogdon et al., 2008), teen pregnancy (Evans et al., 1992), sexual behaviour (Selvan et al., 2001) or team sport (Ashworth and Heyndels, 2007).

Educational peer effects are of primary interest, basically for two reasons. First, school- or classmates constitute a relevant peer reference group. Indeed, children and adolescents spend a lot of time in the school environment, giving the opportunity to social interactions to arise sizeably. Second, the nature of these interactions may affect the individual learning process. Empirical evidence shows that students learn not only from their teachers but also from their classmates. It is worth mentioning, however, that measuring the channels by which peer effects are transmitted is a daunting task because we generally cannot ask so much from the data[6]. Consequently, most existing literature focuses on the source of peer effects but not on the mechanism of transmission. Overall, empirical evidence on the magnitude of peer effects is rather mixed - essentially due to the heterogeneity in the samples and econometric tools used - but generally reports positive, small yet significant peer estimates.

From a policy perspective, peer effects give precious information on the way pupils should be grouped to maximize cognitive skills accumulation or to achieve equity goals. If low-ability students benefit most from high-ability peers and the latter are not adversely affected by a more heterogeneous environment, mixing students is an efficient strategy to enhance cognitive skills. On the contrary, if students with the same cognitive skills perform better

[6]We can distinguish between two main channels: *effective* learning, namely that students benefit directly from explanations and support from their classmates, or *observational* learning, namely that high-ability students are source of motivation and identification (Bandura, 1986).

together, tracking students may be the appropriate school design. Finally, analyzing if peer group diversity weakens or strengthens the family background effect within the classroom allows to deal with the issue of equality of opportunity.

Swiss context

Ability tracking is the rule at the lower secondary level in Switzerland. It takes the form of ability grouping (when tracking occurs between schools or between classes), level grouping (when tracking occurs within the class) or a combination of both. Whatever the tracking design, the main idea is to create homogeneous groups of students in order to adapt teaching to their specific needs and to improve efficiency. However, this system has some drawbacks and limitations, especially in the Swiss case:

- Ability tracking in Switzerland occurs relatively early, namely at 11-13 years of age, when it is impossible to measure rigorously the effective skills of a child. As ability track enrollment is often based on subjective decisions (e.g. parental or teacher endorsement), this system reduces educational opportunities for disadvantaged pupils given that family's influence is important at early stages of the education process (Vellacot and Wolter, 2004; Kronig, 2007). If parents do not speak the national language, are not interested by education, or have deficiencies in cultural capital, less gifted pupils are confronted to discrimination effects, whose consequences are far-reaching given that these influence strongly the pupil's educational pathway. An experiment in canton Fribourg reported that teacher's recommendations tend to be biased according the parents' socioeconomic background (Baeriswyl et al., 2006). Moreover, Bauer and Riphahn (2006) report that early tracking based on school performances leads to a reinforcement of the effects of students' socioeconomic background on educational outcomes.

- PISA studies revealed significant overlaps in terms of school performances across the different ability tracks (Moser and Angelone, 2008). As a result, reorientations among ability tracks are frequent. If tracking rules are not reliable, i.e., if they do not represent perfectly the true cognitive ability of the student, school efficiency and effectiveness is not ensured. Moreover, field-invariant tracking means that students are strong or weak in all disciplines, which is not necessarily the case.

- The lowest ability tracks no longer meet the labour market needs. Nowadays, even those in manual professions need some intellectual aptitudes to carry out increasingly demanding tasks. Such individuals should have access to a general education of a higher level.

- The number of pupils at the lower secondary level is expected to decrease on average by 3% by the end of 2017, and small schools will probably not survive this demographic trend (CSRE, 2014). Merging the schools in difficulty is generally not satisfactory from the point of view of social cohesion and is complicated by the fact that the Swiss lower secondary level is essentially composed of different school types which are managed separately. This explains also why the cooperative and integrated models have been adopted in several cantons as alternatives to the separated system (CSRE, 2010).

These arguments show that the Swiss tracking system has not fully convinced regarding the efficiency and equity gains that were expected by creating homogeneous classes. Several countries like Sweden, Finland or Poland which have postponed or abolished tracking have not only reduced the correlation between school outcomes and socioeconomic background but have also improved students' average performance. More specifically, Finnish pupils are top performers in the last PISA surveys. As mentioned in OECD (2009), this situation is "perhaps consistent with micro-evidence on peer effects, which suggest that the presence of academically strong pupils reinforces education outcomes of weak pupils, while adverse effects of weak students on strong students may perhaps be smaller or absent." Similarly, Hanushek and Woessman (2006) find that selective systems do not outperform comprehensive ones in terms of schooling performances. The question as to which of the two systems is the most appropriate to increase school efficiency (but also equity) is still open.

Research questions and data

The first essay seeks to determine if adopting a comprehensive school design at the lower secondary level in Switzerland would be Pareto-improving in terms of efficiency and equity. Using a peer effects framework, the empirical analysis has for objective to determine who benefits most from peers' influence within the classroom and if increasing peer diversity leads to adverse effects on student's performances. The second part of the essay focuses on the equity issue by estimating if peer heterogeneity strengthens or weakens the impact of family background on educational performances. More specifically, the study addresses the three following research questions:

1. *Are less-endowed students positively/negatively affected by high-endowed peers?* If positive, it means that mixing students can generate positive effects for less-endowed students through observational or effective learning. Conversely, we can expect a negative relationship if students with learning difficulties are discouraged by the presence of high-ability peers. In such a case, tracking students in different ability classes is more adapted to satisfy students' specific needs.

2. *Does class heterogeneity affect positively/negatively student's performances?* A positive relationship would mean that mixing students creates positive learning environment, e.g., by giving the opportunity to high-ability students to benefit also themselves from the explanations that they give to their comrades. If negative, this could support the thesis that some disruptive students may reduce class motivation and performance (Lazear, 2001).

3. *Does class heterogeneity weakens or reinforces the family background effect on educational performances?* This question determines if mixing students can lead to equity gains or not.

My research relies on the PISA 2006 Swiss national survey whose questionnaires have been administered to a nationally representative sample of pupils in the 9^{th} grade (i.e., last year of compulsory schooling). These cross-sectional data contain relevant information on pupil's socioeconomic background, ability tracks (high-, middle- or low-ability tracks), tracking systems (full- or partial-tracking) and the field of study (reading, mathematics, sciences). As we can identify the school, ability track, and class the student attends, this dataset is very relevant to estimate peer effects in the Swiss context where ability tracking procedures essentially occur at the class level (but also within the class).

Models and identification

Both theoretical and empirical research on educational peer effects start from the following education production function:

$$Y_i = \gamma_0 + \gamma_1 X_i + \gamma_2 \bar{X}_{(-i)j} + \gamma_3 \bar{Y}_{(-i)j} + \gamma_4 S_s + \varphi_i \tag{1.1}$$

where Y_i denotes the educational outcome of student i, X_i is a vector of observable student characteristics, $\bar{X}_{(-i)j}$ are the average characteristics of peer group j excluding the contribution of student i, $\bar{Y}_{(-i)j}$ is the average outcome of peer group j excluding the contribution of student i, S_s is a set of school or class characteristics and φ_i is an error term. With respect to the terminology adopted by Manski (1993), the coefficient γ_2 represents *contextual* effects (i.e., the impact of peer background characteristics which are exogenous to the peer group formation) whereas γ_3 reflects *endogenous* effects (i.e., the influence of contemporaneous peers' behaviours such as effort, inspiration, or rivalry). The *linear-in-means* model presented in equation (1.1) is subject to three main econometric problems - simultaneity, collinearity and selectivity - which are discussed below. It is worth mentioning that previous literature has managed to overcome some but not necessary all of these issues simultaneously.

First, the reciprocal influence between the individual outcome Y_i and peers' outcomes $\bar{Y}_{(-i)j}$ implies a *reflection* problem (Manski, 1993). The main strategy to solve this problem of simultaneity is to conduct an econometric analysis using a lagged value of peer achievement as instrumental variable (Hanushek et al., 2003; Sund, 2007; Burke and Sass, 2008), but the relevance of this kind of instrument is questionable on the grounds that random shocks or serial correlation may still be present.

Second, it is difficult to separate the impact of contextual effects from endogenous ones on school outcomes. Given that $\bar{X}_{(-i)j}$ influences $\bar{Y}_{(-i)j}$ through Y_j, *collinearity* problems may arise. In this context, the main bulk of the literature relies on a reduced form model which relates the individual "endogenous" outcome on peer group "exogenous" characteristics, ignoring the precise nature of the peer effects parameter. To develop this idea, let us consider the equation (1.1) but where the dependent variable corresponds to the mean peer achievement:

$$\bar{Y}_{(-i)j} = \gamma_0 + \gamma_1 \bar{X}_{(-i)j} + \gamma_2 X_i + \gamma_3 Y_i + \gamma_4 S_s + \varphi_j \tag{1.2}$$

By substituting for mean peer achievement, $\bar{Y}_{(-i)j}$, in equation (1.1), I obtain:

$$\begin{aligned} Y_i &= \frac{\gamma_0(1+\gamma_3)}{(1-\gamma_3^2)} + \frac{(\gamma_1+\gamma_2\gamma_3)}{(1-\gamma_3^2)} X_i + \frac{(\gamma_2+\gamma_1\gamma_3)}{(1-\gamma_3^2)} \bar{X}_{(-i)j} \\ &+ \frac{(1+\gamma_3\gamma_4)}{(1-\gamma_3^2)} S_s + \frac{(\gamma_3\varphi_i+\varphi_j)}{(1-\gamma_3^2)} \end{aligned}$$

or, more simply:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 \bar{X}_{(-i)j} + \beta_3 S_s + \epsilon_i \tag{1.3}$$

where β_2 determines the magnitude of total peer effects as we can no longer distinguish between contextual and endogenous effects. An alternative strategy to solve this issue is to relax the assumption of group interaction by considering peer groups as individual-specific. If each student has its own peer group, we can use the performances or background characteristics of the excluded classmates (i.e., students with whom the individual does not have direct contact) as an instrument for peers' endogenous outcome (Bramoullé et al., 2009; De Giorgi et al., 2009). While this approach has the advantage to give a clear interpretation for each structural parameter, i.e., γ_2 and γ_3, it requires too much from the data to be easily generalized.

Last but not least, *selection bias* occurs when some unobserved components falsify the relationship between the peer variable and the individual outcome. For example, parents can influence the school choice of their offspring when the selection procedure relies on parental endorsement or geographical location. As a result, if students with similar backgrounds are regrouped within the same school, the assignment process is not random and the estimates will be upward biased. In such a case, similar behaviour between pupils is maybe not explained by social interactions but rather by the social environment they face, which reflects the existence of correlated effects (Manski, 1993).

Different strategies are possible to deal with selectivity issues. The most rigorous research design to rule out selection bias consists in conducting natural experiments in which students are randomly assigned to peer reference groups. Sacerdote (2001) focuses on peer effects among college roommates at Dartmouth college whose random assignment implies that there is no correlation between roommates' and individual background characteristics. A second type of natural experiment consists in using a randomly assigned policy treatment to subjects. In this context, Hoxby and Weingarth (2005) observe students before and after they experience policy-induced changes in peers. The authors consider a switch in school reassignment policy in Wake County (US) from balanced schools on the basis of race to balanced schools on the basis of family income. This reassignment allows to measure how student's achievement is influenced by the new peer composition in the classroom, conditional on student's fixed characteristics. The study of Duflo et al. (2008) uses experimental data from Kenya to compare schools in which students were randomly assigned to a first-grade class with other schools in which students were assigned on the basis of prior achievement. Randomized evaluations, however, face several theoretical and practical issues which explain why they are not regularly proposed in social sciences.

When working with observational data, a traditional solution to control for endogenous peer group formation consists in modelling the enrollment process by conducting a two-stage least squares (2SLS) regression where the peer group variable, which generally takes the form of a peer ability measure, is the dependent variable of the first stage regression (Lefgren, 2004; De Paola and Scoppa, 2010). However, finding an instrument which satisfies the exclusion restriction is very difficult. A second possibility is to resort to a fixed effects strategy to control for unobserved time-invariant individual and/or school characteristics. In the literature, most econometric models include a school-specific component (Mc Ewan, 2003; Schneeweis and Winter-Ebmer, 2007; Vigdor and Nechyba, 2007; Fletcher, 2010). For instance:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 \bar{X}_{(-i)j} + \underbrace{\mu_s + \gamma_i}_{\epsilon_i} \tag{1.4}$$

where μ_s represents the school-specific component and ν_i an idiosyncratic error term. Such an analysis is valid as long as there is no tracking within schools. When students are sorted by ability within the school - which is the case at the lower secondary level in Switzerland - , we have to control for the ability track they follow. In comparison with equation (1.4), equation (1.5) instead incorporates an ability track specific component and can be represented as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 \bar{X}_{(-i)} + \underbrace{\mu_k + \nu_i}_{\xi_i} \tag{1.5}$$

The first essay starts from equation (1.5) to estimate linear peer effects and to control for selectivity issues. To answer the first research question, I consider the same model but within a quantile regression framework in order to estimate returns to peers along the ability distribution. By including the standard deviation of the peer variable in the educational production function, I can respond to the second research question. Finally, a model including an interaction term between peer heterogeneity and family background allows to address the third research question by analyzing how class diversity affects the parental background effect on educational outcomes.

Main findings

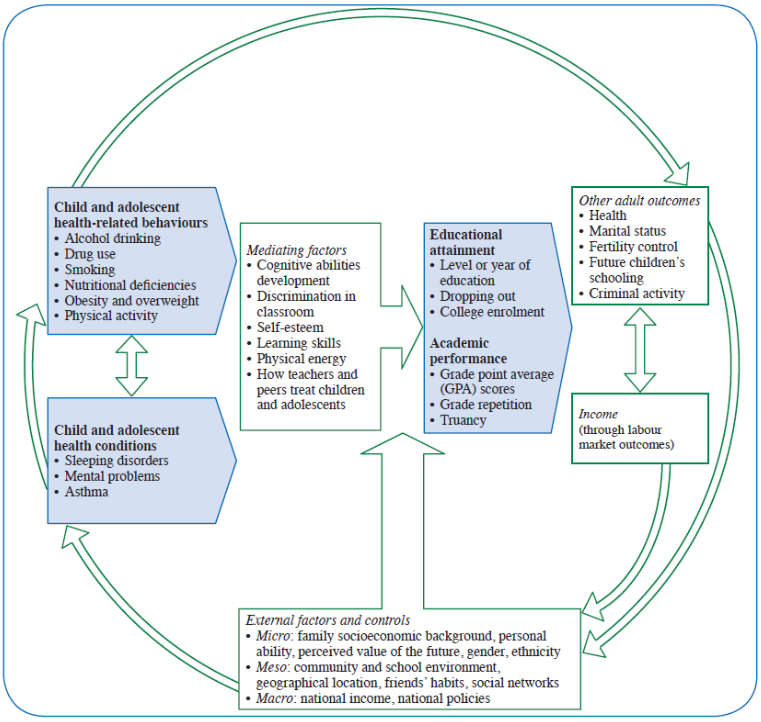
Empirical findings report positive, small but significant average peer effects in reading and sciences after controlling for ability track fixed effects. The average peer coefficient in mathematics, however, is not significant. When accounting for non-linearity in peer effects, different pictures emerge according to the field considered. In reading, the returns to peers decrease continuously along the student's ability distribution. In sciences, only low-achieving students benefit significantly from peer effects while the inverse is true in mathematics. Results also indicate that increasing peer diversity within the classroom would not lead to detrimental effects on the (high-achieving) classmates but would reduce the family background effect on school performances, whatever the field considered. In short, ability-mixed classes could be a potential solution to increase both efficiency and equity in reading and sciences courses whereas mathematical courses seem to be more efficient, although not egalitarian, when students are regrouped with similar peers. In conclusion, equity and efficiency can go hand in hand but this relationship depends on the field considered.

1.3.2 Cannabis use and short-term academic performances

Theoretical background

Economists initially considered health as one form of human capital (Mushkin, 1962; Becker, 1964). The pioneering contribution of Grossmann (1972) demonstrated, however, that health and human capital are two distinct dimensions which interact together in explaining wage differences. More precisely, education is incorporated as input in the health capital model proposed by Grossmann on the grounds that better education improves the efficiency with which gross investments in health operate. Until recently, many contributions have then considered health behaviour as a by-product of education rather than the inverse. For instance, some argue that schooling is a long-term investment which encourages students to stay healthy (Cutler and Lleras-Muney, 2006); others consider that increasing cognitive skills reinforces the knowledge of health issues (Grossman, 1973) while some others assume that higher education is correlated with higher earnings which give the possibility to afford more health expenditures (Kenkel, 1991; Lleras-Muney, 2006). The nature of the relationship between health and education, however, is not unilateral. Health-related behaviours and health conditions can affect schooling attainment through different mechanisms presented in Figure 1.2 taken from Shurcke and de Paz Nieves (2011).

When focusing on health behaviour, we can identify three possible interactions between substance use and education. First, substance use may have an impact on educational outcomes through mediating factors such as motivation, self-esteem, hangover or cognitive functioning (memory, reasoning, concentration). Second, cannabis use may be a consequence of poor schooling outcomes. Finally, health and schooling outcomes may be not directly



Source: Shurcke and de Paz Nieves (2011)

Figure 1.2: Analytical framework between health and education

related but share underlying causes such as inherent ability, family background, peers' influence, time preferences or deviant behaviour. For that purpose, thorough econometric analyses are needed in order to determine if the health-education gradient is based on causal mechanisms or if it reflects only an spurious association between the two variables.

Several studies have demonstrated that cannabis use increases attention disorders and reduces concentration, motor skills or memory (Matsuda et al. 1993; Heyser et al., 1993). Clinical and epidemiological research also confirms that cannabis use may lead to dependence (or addiction) effects and that individuals may have strong difficulties to stop their consumption despite social, psychological or physical impairments (Copeland et al., 2001; Stephens et al., 2002, Budney, 2006; Budney et al., 2006; Budney et al., 2007). However, we cannot determine precisely if the detrimental effects of cannabis use on mediating factors are short-lived or if they are cumulative accross time. While the effects of cannabis on long-term education outcomes (e.g., school dropout, educational attainment or wages) is rather well-documented (Bray et al., 2000; Yamada et al., 1997; Lynksey et al., 2003; MacLeod et al., 2004; Bessey and Backes-Gellner, 2009; McCaffrey et al., 2010), few papers focused on the effects on short-term education outcomes such as grades, absenteeism or concentration. Pacula et al. (2003) take advantage of the panel structure of the NELS data to estimate a difference-in-differences model whose findings show that cannabis use has a negative impact on standardized mathematics tests. Focusing on US adolescents aged 12-18, Roebuck et al. (2004) conclude on the basis of zero-inflated negative binomial regressions that cannabis users are more likely to skip school relative to non-users. Engberg and Morral (2006) find that reductions in substance use may improve school attendance. Using a sample of US adolescents, their methodological approach relies on both random and fixed effects, augmented with a set of time-varying control variables. On the basis of logistic regressions, the study of Caldeira et al. (2008) finds that consuming cannabis leads to concentration problems and higher absenteeism at school.

Swiss context

Switzerland is well above the international average regarding lifetime prevalence of cannabis consumption (Figure 1.3). With one out of three Swiss adolescents who have already smoked cannabis at least once at the age of 15, the situation is worrisome given that early cannabis use is positively associated with future dependence and lower cognitive functioning. In Switzerland, around 12% of males and 5% of females between 15 and 24 years regularly consume cannabis[7] (FSO, 2008).

The situation is especially critical when discussing cannabis consumption within the school environment. According to a report made by the FOPH (Federal Office of Public Health) and Addiction Suisse (2004) intended for Swiss lower secondary schools, about one third ninth-grade teachers have seen students attending their classes affected by cannabis consumption. Risky health behaviours may lead to a deterioration of educational outcomes (e.g., truancy, concentration or performance), explaining why educational and health entities must work together. Early consumption may start at the end of compulsory school and further increase when youths enter into academic or vocational education. Consequently, it is very important to inform pupils relatively early on the potential damages of this substance and

[7] It is worth precising that such alarming rates, however, are not characteristic of tobacco and alcohol whose consumption is close to the international average.

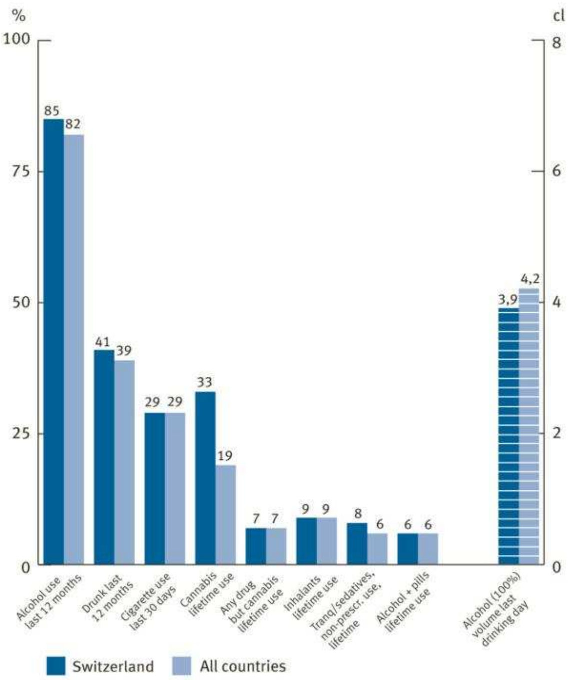
to involve all concerned actors in the discussion.

Cannabis use has been at the heart of political debates in Switzerland during the last years. The Commission for Drug Issues in Switzerland argues in favor of a legalization of cannabis in the country while the federal government is more divided on this question. Consequently, the message addressed to youths is unclear and the negative consequences of cannabis consumption run the risk of being neglected. This situation is amplified by heterogeneous practices regarding law enforcement among Swiss cantons. In October 2013, administrative fines for cannabis consumers entered into force, for an amount of 100 CHF, conditional on the age of the individual (18 years of age) and the quantity (maximum of 10 grammes). However, the application of administrative fines will not raise consciousness regarding cannabis consumption. This new law only reduces the repression's workload for police and justice but does not send clear messages regarding the real risks of cannabis consumption and what constitutes a misuse. In this context, the recent experiences of liberalization made in the United States (Colorado) and Uruguay might bring important elements into the discussion and therefore need to be carefully analyzed.

Research questions and data

My second essay - in collaboration with Joachim Marti - seeks to determine *what is the impact of cannabis consumption on short-term academic performances*. As mentioned before, the theoretical and empirical background sustaining the hypothesis that cannabis use has consequences on short-term educational outcomes is not clearly established in the literature and needs further investigations, especially outside the United States. Six dependent variables are considered (absenteeism, school difficulties, poor grades, lack of motivation, lack of engagement and concentration problems) while frequency of cannabis consumption is measured through any (less than 3 times a month) and frequent (at least once a week) use. The accent is on high-school students because they represent an appropriate treatment group to measure the effects of addictive products on human capital accumulation. From a methodological viewpoint, a particular attention is paid to the role played by unobserved heterogeneity in the obtained results.

The study relies on the longitudinal dataset TREE ("Transition from Education to Employment") which surveys the post-compulsory educational and labour market pathways of some students who participated in the PISA



Source: http://www.espad.org/switzerland (2007)

Figure 1.3: Substance use in Switzerland: an international comparison

2000 international study. The sample has been followed up by means of seven survey panels between 2001 and 2007, with an eighth wave in 2010. A ninth wave took place in 2014. TREE data are relevant to study the impact of cannabis use on education because they include rich information on background characteristics, substance use, school outcomes and psychological traits which give the opportunity to analyze the health-education gradient from different angles.

Models and identification

Let us start by presenting the classical model estimated by OLS:

$$Y_i = \beta_0 + \beta_1 C_i + \beta_2 X_i + \varsigma_i \tag{1.6}$$

where Y_i is the educational outcome, C_i is a measure of substance consumption, X_i is a set of individual characteristics and ς_i is an error term. The first econometric issue - simultaneity - is related to the reciprocal influence between Y_i and C_i whereas a second problem - the omitted variable bias - emerges when C_i is correlated with some unobserved factors included in the error term ς_i. Controlling for unobserved characteristics is of primary importance to interpret the results in causal terms and to propose relevant policy recommendations.

An IV procedure is theoretically the best way to deal with both kinds of biases. Numerous contributions rely on alcohol policies (Dee and Evans, 2003; Renna, 2006; Chatterji, 2006) or religiosity measures (Roebuck et al., 2004; Bessey and Backes-Gellner, 2009) as instruments for substance use. Except for the papers using IV, few studies deal with the problem of inverse causality between health and education. While some authors just mention the issue (Mc Caffrey et al., 2008; Horwood et al., 2010), we decide to replace current substance use by its lagged value. Following this approach, we rewrite equation (1.6) as:

$$Y_{it} = \beta_0 + \beta_1 C_{it-1} + \beta_2 X_i + \varsigma_{it} \tag{1.7}$$

where C_{it-1} is a lagged measure of substance use.

Endogeneity also comes from the fact that other factors may determine both health and education simultaneously (i.e.,"third factor" theory). Among them, we can mention the socioeconomic status, taste for deviance, innate ability, coping mechanisms or social networks. Even with a large set of background variables, selection on observables is generally not sufficient to account for all confounding factors. In the absence of valid instruments, scholars generally resort to random and/or fixed effects methods to account for unobserved time-invariant heterogeneity (Engberg and Morral, 2006; Basla et al., 2011). By introducing individual fixed effects in equation (1.7), I obtain:

$$Y_{it} = \beta_0 + \beta_1 C_{it-1} + \beta_2 X_i + \underbrace{\eta_i + \epsilon_{it}}_{\varsigma_{it}} \tag{1.8}$$

where η_i represents an individual-specific component.

Economists also consider propensity score matching techniques to analyze the impact of health on different schooling outcomes (McCaffrey et al. 2010; Fletcher and Frisvold, 2011). These methods rely on the conditional independence assumption which postulates that potential outcomes are independent of treatment assignment, conditional on the predicted probability to be treated. The advantage of such a strategy is to provide some information on treatment participation, which is not the case when using fixed effects. Nevertheless, this assumption cannot be verified. Therefore, the recent literature proposes some sensitivity analyses in order to estimate the impact of unobserved heterogeneity on parameter estimates, e.g. Rosenbaum bounds (Rosenbaum, 2002) or Altonji's approach (Altonji et al., 2005).

To answer the research question, the identification strategy of the second essay is based on both approaches mentioned above, i.e., fixed effects methods and propensity score matching. The objective is to compare the respective results to obtain consistent findings. The empirical part based on propensity score matching is completed by an auxiliary analysis (Rosenbaum bounds) to determine to what extent a small departure from the conditional independence assumption may invalidate some of the results.

Main findings

Empirical findings suggest that cannabis consumption is positively and significantly associated with skipping school and getting poor grades and that accounting for unobserved heterogeneity is a key issue when we deal with such a research question. Results advocate for a strong coordination between schooling and health policies and for a better awareness of the risks associated with cannabis use. They also bring new elements in the current debate in Switzerland where discussions on liberalization of cannabis are still topical.

1.3.3 Who benefits most from university education in Switzerland?

Theoretical background

In most industrialized countries, a large share of public spending is attributed to education. But what exactly is the return on this investment? While the costs are relatively easy to calculate, the benefits the individual and the society reap from education are harder to quantify. Indeed, what is the real impact of education on earnings? Does education influence directly workers' productivity which in turn enhances earnings on the labour market? Does educational attainment reflect only individual decisions according the principle of comparative advantage? Or is education only a signal reflecting the workers' productivity or non-cognitive abilities? Although there is no clear answer to these questions, empirical evidence shows that students are always rewarded by their schooling investment, i.e., differences in education explain differences in earnings even after controlling for differences in individual attributes.

The traditional way to quantify the private return to education is to estimate the impact of one additional year of schooling on wage which implicitly assumes a linear relationship between schooling and wages (Ferro-Luzzi and Silber, 1998; Trostel et al., 2002; Harmon et al., 2003). However, because many students do not complete their degrees in the standard number of years, estimates based on years of education are often biased (Jaeger and Page, 1996). For that purpose, empirical analyses are often completed by an estimation of returns to schooling based on educational attainment (Brunello and Miniaci, 1999; Psacharopoulos and Patrinos, 2004). Some contributions focus directly on estimating the *sheepskin effects*, which correspond to gain in earnings resulting from the completion of educational programs, controlling for years of education (Jaeger and Page, 1996; Park, 1999; Ferrer and Riddell, 2002, 2008). Therefore, accounting for credential effects imply discontinuities in schooling returns which may differ across educational levels.

Recent contributions in the literature, however, focus now on the fact that the return to education may vary across individuals with the same educational attainment according to the principle of comparative advantage (Willis and Rosen, 1979; Willis, 1987; Heckman et al., 1998; Carneiro et al. 2001; Heckman et al., 2006). Based on the theoretical framework of Roy (1951), the main idea is that students take schooling decisions on the basis of their expected financial gains. Consequently, there is not a single return to schooling in the population but a distribution of returns that depends on individual characteristics and expectations.

Swiss context

Boarini and Strauss (2010) show that the internal rate of return to tertiary education in Switzerland is relatively high (11.3% for men, 10.1% for women), with a top position in international comparison (Figure (1.4)). *A priori*, such a result is very promising given the importance attached by Switzerland to its tertiary education system.

This picture, however, masks the fact that rates of return differ significantly across the different tertiary educational levels in Switzerland. Using a cost-benefit analysis, Wolter and Weber (2005) report an annual rate of return for male workers of 8.7% for higher vocational schools, 10.6% for the

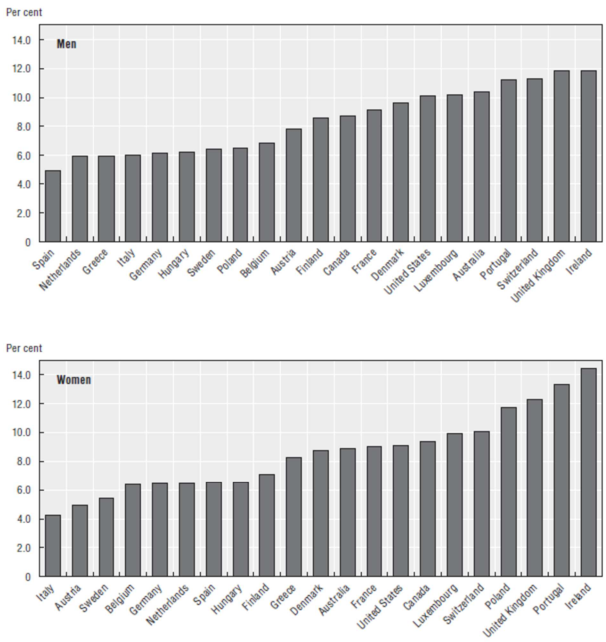


Figure 1.4: Internal rates of return: an international comparison

universities of applied sciences but only 5.4% for conventional universities[8]. In fact, returns are particularly high for universities of applied sciences and tertiary vocational degrees. Different reasons such as the high opportunity costs faced by university students, their limited employability just after graduation or the absence of on-the-job training at university may explain this phenomenon.

Another characteristic of Switzerland is that students with disadvantaged parental backgrounds face strong economic, social, institutional and motivational barriers throughout their educational pathways (Vellacott and Wolter, 2004). Different factors can be mentioned. First, the likelihood of completing university studies depends heavily on the educational background of the parents. Overall, individuals coming from a well-educated family are 1.6 times more likely to attend a university degree (CSRE, 2014). Next, parental income may also be a source of inequality. Indeed, higher education studies require financial resources. Although tuition fees in Switzerland are relatively low, living costs are high and the majority of students have to work to finance their university studies. In general, youngsters whose parents cannot

[8]The rates of return are calculated relative to the next lower level of education.

afford such educational investment have to work part-time (60% on average) to support themselves (CSRE, 2010). Such a situation not only reduces their motivation to attend university but also penalizes them compared to students for whom money is not a problem. Last, migration background is also important. Even if the proportion of foreign students at the tertiary level reaches 24%, around three quarter of these individuals obtained their certificate of entrance for universities abroad (CSRE, 2010). For foreign students who obtained their matura in Switzerland, this rate is just above 6%. Nevertheless, the probability to complete university for second or third generation of immigrants is higher than that of their parents.

Research questions and data

The third essay aims at discovering which kinds of students benefit most from university education in Switzerland by focusing on family background characteristics. This study estimates heterogeneous returns to schooling and analyzes if schooling decisions are driven by the principle of comparative advantage, i.e., if economic agents take schooling decisions on the basis of their expected return to schooling. To answer this question, we need to determine the relationship between returns to university education and the predicted probability of completing this degree. In this context, this contribution addresses the two following questions:

- *"Are university education and parental background substitutes or complements in generating earnings capability?"*
- *Is the distribution of returns identical for both men and women?*

The results are meant to provide precious insights regarding policy recommendations. If completing university education reinforces the parental background effect on earnings capability, it means that individuals have comparative advantage in schooling achievement according to their background characteristics. In such a case, increasing access to university education for less-endowed students is maybe not the optimal policy regarding efficiency gains. However, if completing university education reduces the parental background effect on earnings capability, it means that increasing access to university education for less-endowed students would lead to both efficiency and equity gains. The study also seeks to determine if the distribution of returns to schooling along the propensity score is the same for men and women.

Data used for the empirical analysis come from the Swiss Household Panel which is a yearly panel study started in 1999 and which contains relevant information regarding wages, education and parental background. The main advantage is that I have access to numerous pre-treatment variables (measured at age 15, i.e., before university entrance) to estimate the predicted probability to complete university education.

Models and identification

To obtain the true return to schooling, we should have the possibility to measure the return to education (i.e., the treatment effect) for the same individual, that is,

$$\beta_i = Y_{1i} - Y_{0i}$$

where Y_{1i} is the potential wage of individual i if graduated ($S_i = 1$) and Y_{i0} the potential wage of individual i if non-graduated ($S_i = 0$). More precisely:

$$\begin{aligned} Y_{0i} &= \gamma_0 X_i + u_{0i} \qquad (1.9) \\ Y_{1i} &= \gamma_1 X_i + u_{1i} \qquad (1.10) \end{aligned}$$

where X_i is a set of observed characteristics while u_{1i} and u_{0i} are unobserved components. The fundamental problem of causal inference is that we cannot observe both outcomes for the same individual simultaneously and therefore the estimation of the true β_i is hypothetical (Holland, 1986). The treatment effect literature proposes two solutions to deal with this issue. First, we can invoke a homogeneity assumption where all individuals in the population are supposed to be identical. Numerous contributions rely on this hypothesis to estimate the average return to schooling through OLS or IV methods (Brunello et al., 1999, 2000; Angrist and Krueger, 2001; Trostel et al., 2002). This approach, however, is unrealistic and has little value for policy recommendations. The second solution considers treatment effects at the group level. The objective is to select randomly two groups in the population of which one is subject to treatment and the other not. The parameter obtained - the average treatment effect (ATE) - is representative of the whole population:

$$ATE = E(Y_{1i} - Y_{0i}) = E(Y_{1i}) - E(Y_{0i}) = E(\beta_i) \qquad (1.11)$$

This approach, however, faces two econometric problems related to population diversity when the sample is not random. To discuss them, I write the observed outcome equation resulting from equations (1.9) and (1.10) as:

$$\begin{aligned} Y_i &= (1 - S_i)Y_{0i} + S_i Y_{1i} \\ &= [(\gamma_1 - \gamma_0)X_i + (u_{1i} - u_{0i})]S_i + \gamma_0 X_i + u_{0i} \\ &= \beta_i S_i + \gamma_0 X_i + u_{0i} \end{aligned} \tag{1.12}$$

where

$$\beta_i = (\gamma_1 - \gamma_0)X_i + (u_{1i} - u_{0i})$$

If treated and untreated groups are not randomly formed, both groups may differ even in the absence of treatment. In such a case, econometric analysis suffers from a selection bias, i.e., when $Cov(S_i, u_{0i}) \neq 0$. IV or FE methods are relevant strategies to eliminate this bias on the grounds that a suitable instrument is by definition uncorrelated with the error term whereas FE eliminate any pre-treatment differences between the two state groups. Second, individuals may self-select into treatment on the basis of their unobserved expected economic returns, i.e., when $Cov(S_i, (u_{1i} - u_{0i})) \neq 0$. In such a case, none of IV or FE methods can solve this return bias. Indeed, FE eliminate only pre-treatment differences between both treated and control groups while an IV strategy cannot control for the correlation between S_i and $(u_{1i} - u_{0i})$ when individuals act on the latter to make their schooling choice.

In the presence of heterogeneity in returns to schooling, we can set stronger identifications assumptions on the IV estimator to identify a local average treatment effect (LATE) that measures the average return to schooling for individuals who are concerned by a change in the instrumental variable (Angrist and Imbens, 1994; Angrist and Krueger, 1999; Angrist, 2004; Heckman, Urzua and Vytlacil, 2006; Angrist and Pischke, 2009). In this context, the returns to schooling are only revealed for the (unidentified) subpopulation (called "compliers") affected by the observed changes in the instrumental variable (Imbens and Angrist, 1994; Imbens and Wooldridge, 2008). Consequently, the LATE estimator does not usually measure the average causal effects on all of both treated and untreated units.

The traditional solution to deal with a situation where individuals self-select into higher education on the basis of their unobserved expected returns is to use control functions (CFs). This method consists in a two-stage approach where the estimated residuals of the schooling equation - which take the form of inverse Mills ratios - are integrated in the wage regression and inform us on the direction of both selection and return bias (Garen, 1984; Heckman and Robb, 1985; Heckman and Vytlacil, 1998; Card, 1999; Blundell et al., 2001; Deschênes, 2007). Expressed differently, CFs offer an immediate test of endogeneity by testing if their coefficients are significantly different from zero, which is not possible with the traditional IV estimator. To summarize, CFs allow to recover the average treatment effect when both heterogeneity and self-selection are present whereas the traditional IV estimator is only able to recover a local average treatment effect for a specific instrument-related sub-population. Both methods, however, depend strongly on the relevance of the instrument chosen and most of traditional candidates (e.g., parental education, distance to schooling, educational reforms) have their own drawbacks (e.g., endogeneity, availability, relevance)(Checchi, 2006).

Prior statements led some scholars to rely on selection on observables to estimate heterogeneous returns to schooling. In the absence of a superior alternative, matching individuals with similar propensity scores is the most interesting way to account for observed selection bias and heterogeneity because of the non-parametric property of matching estimators (Rosenbaum and Rubin, 1983). Propensity score matching relies on the conditional independence assumption which can be written as:

$$(Y_{1i}, Y_{0i}) \perp S_i | P(S_i = 1 | X_i)$$

Expressed differently, it assumes that all confounding factors in the relationship between wages and schooling are captured by the propensity score. In the case of returns to schooling, Tsai and Xie (2008), Brand and Xie (2010) and Brand and St-Thomas (2012) propose hierarchical linear models (HLM) based on propensity score matching to explore the pattern of returns to higher education as a function of the propensity score. The first approach relies on a stratification-multilevel method which consists in partitioning the sample into homogeneous propensity score strata and estimating within each subpopulation the return to schooling. The second approach is a matching-smoothing model where the variations in matched wages differences are smoothed by using non-parametric functions. In both cases, the nature of the association between the observed returns to schooling and the propensity score allows us to determine who benefits most from education.

The corresponding essay estimates heterogeneous returns to university education by using both hierarchical linear models described above. To account for sample selection issues, the specification for women is augmented with the Heckman's selection procedure.

Main findings

Empirical findings report that men with the lowest propensity to complete university education obtain the highest returns to schooling on the labour market, meaning that a university degree reinforces more the earnings capability of low propensity men than that of their high propensity peers. This statement, however, only holds when labour market variables are introduced in the regression model, which indicates that working experience is an important factor influencing positively the benefits obtained by low propensity men from completing a university degree. In addition, the different empirical models converge to the conclusion that returns to university education for women are rather homogeneous along the propensity score distribution. In summary, both results lead to the rejection of the comparative advantage assumption which stipulates that individuals with the highest expected returns to schooling benefit most from this degree.

Results indicate that facilitating access to university education for students with disadvantaged parental background, especially men, is relevant, not only in terms of equality of opportunity but also in terms of efficiency given that these students obtain higher returns on their educational investment. Different policy recommendations are possible. In the short run, efforts have to be made to facilitate access to grants or loans for disadvantaged students. In the long run, filling the gap between less- and more-endowed individuals requires adapted schooling policies since early childhood. In this context, the national incentive program that created numerous new childcare places since 2003 (more than 35'000) and the support of special-needs children into the public education system go in the right direction.

1.4 References

Altonji J., T. Elder and C. Taber, Selection on observed and unobserved variables: assessing the effectiveness of catholic schools, *Journal of Political Economy*, 113(1), pp. 151-184, 2005.

Angrist J. and A. Krueger, Instrumental variables and the search for identification: from supply and demand to natural experiments, *Journal of Economic Perspectives*, 15(4), pp. 69-85, 2001.

Angrist J. and G. Imbens, Identification and estimation of local average treament effects, *Econometrica*, 62(2), pp. 467-475, 1994.

Angrist J., Treatment effect heterogeneity in theory and practice, *Economic Journal*, 114(494), pp. 52-83, 2004.

Ashworth J. and B. Heyndels, Selection bias and peer effects in team sport: the effects of age grouping on earnings of German soccer players, *Journal of Sport Economics*, 8(4), pp. 355-377, 2007.

Baeriswyl F., C. Wandeler, U. Trautwein and U. and K. Oswald, Leistungstest, Offenheit von Bildungsgängen und obligatorische Beratung der Eltern. Reduziert das Deutschfreiburger Übergangsmodell die Effekte des sozialen Hintergrunds bei Übergangsentscheidungen?, *Zeitschrift für Erziehungswissenschaften*, 9. Jg., Heft 3, S. 373-392. Verlag für Sozialwissenschaften, Wiesbaden, 2006.

Balsa A., L. Giuliano and M. French, The effects of alcohol use on academic achievement in high school, *Economics of Education Review*, Vol. 30, pp. 1-15, 2011.

Bandura A., *Observational learning*, Social foundations of thought and action: a social cognitive theory, Chapter 2, Prentice-Hall, USA, 1986.

Basu A., J. Heckman, S. Navarro and S. Urzua, Use of instrumental variables in the presence of heterogeneity and self-selection: An application in breast cancer patients, *Health Economics*, 16(11), 2007.

Bauer P. and R. Riphahn, Timing of School Tracking as a Determinant of Intergenerational Transmission of Education, *Economics Letters*, 91(1), pp. 90-97, 2006.

Bauer P. and R. Riphahn, Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second-generation immigrants, *Journal of Population Economics*, 20(1), pp. 121-148, 2007.

Bayer P., R. Hjalmarsson and D. Pozen, Building criminal capital behind bars: peer effects in juvenile corrections, *Quarterly Journal of Economics*, 124(1), pp. 105-147, 2009.

Becker G., *Human capital: a theoretical and empirical analysis, with special reference to education*, New York: Columbia University Press, 1964.

Bessey D. and U. Backes-Gellner, Marijuana consumption, educational outcomes and labour market success: evidence from Switzerland, Leading House Working Paper N°43, University of Zurich, 2009.

Betts J.R. and A. Zau, Peer groups and academic achievement: panel evidence from administrative data, unpublished document, Public Policy Institute of California, 2004.

Blundell R., L. Dearden and B. Sianesi, Estimating the returns to education: models, methods and results, unpublished document, Centre for the Economics of Education, LSE, 2001.

Boarini R. and H. Stauss, What is the private return to tertiary education ? New evidence from 21 OECD countries, OECD Journal: Economic Studies, Vol. 2010, 2010.

Bound J., D. Jaeger and R. Baker, Problems with instrumental variable estimation when the correlation between the instruments and the endogenous explanatory variables is week, *American Statistical Association*, 90(430), pp. 442-450, 1995.

Bramoullé Y., H. Djebbari and B. Fortin, Identification of peer effects through social networks, *Journal of Econometrics*, Vol. 150, pp. 41-55, 2009.

Brand J. and Y. Xie, Who benefits most from college? Evidence for negative selection in heterogeneous returns to higher education, *American Sociological Review*, 75(2), 2010.

Brand J. and J. Simon-Thomas, Causal effect heterogeneity, California Center for Population Research, Working paper 2011-012, 2012.

Bray J., G. Zarkin, C. Ringwalt and J. Qi, The relationship between marijuana initiation and dropping out of high school, *Health Economics*, Vol. 9, pp. 9-18, 2000.

Brunello G. and R. Miniaci, The economic returns to schooling for Italian men. An evaluation based on instrumental variables, *Labour Economics*, 6(4), pp. 509-519, 1999.

Brunello G., S. Comi and C. Lucifora, The returns to education in Italy: a new look at the evidence, IZA Discussion Paper N° 130, 2000.

Budney A., Are specific dependence criteria necessary for different substances: How can research on cannabis inform this issue?, *Addiction*, 101 (Suppl 1), pp. 125-133, 2006.

Budney A., B. Moore, H. Rocha and S. Higgins, Clinical trial of abstinence-based vouchers and cognitive-behavioral therapy for cannabis dependence, *Journal of Consulting and Clinical Psychology*, 74(2), pp. 307-316, 2006.

Budney A., R. Roffman, R. Stephens and D. Walker, Marijuana dependence and its treatment, *Addiction science and clinical practice*, 4(19), pp. 4-16, 2007.

Burke M. and T. Sass, Classroom peer effects and student achievement, Calder Working Paper N° 18, 2008.

Caldeira K., A. Arria, K. O'Grady, K. Vincent and E. Wish, The occurence of cannabis use disorders and other cannabis-related problems among first-year college students, *Addictive behaviours*, Vol. 33, pp. 397-411, 2008.

Card D., The causal effect of education on earnings, in O. Ashenfelter and D. Card (eds.), Vol. 3A, *Handbook of labour Economics*, Amsterdam: Elsevier, pp. 1801-1863, 1999.

Carneiro P., J. Heckman, and E.Vytlacil, Estimating the return to education when it varies among individuals, unpublished manuscript, Department of Economics, University College London, 2001.

Chatterji P., Does alcohol use during high school affect educational achievement?: Evidence from the National Education Longitudinal Study, *Economics of Education Review*, Vol. 25, pp. 482-497, 2006.

Checchi D., *The economics of education: human capital, family background and inequality*, Cambridge University Press, 2006.

Centre suisse de coordination pour la recherche en éducation (CSRE), Swiss education report 2010, Aarau, 2010.

Centre suisse de coordination pour la recherche en éducation (CSRE), Swiss education report 2014, Aarau, 2014.

Coleman J., E. Campbell, C. Hobson, J. McPartlan, A.M. Mood, F.D. Weinfeld and R.L. York, *Equality of educational opportunity*, Government Printing Office, Washington D.C., 1966.

Copeland J., W. Swift, R. Roffman et R. Stephens, A randomized controlled trial of brief cognitive-behavioral interventions for cannabis use disorder, Journal of Substance Abuse Treatment, 21(2), pp. 55-64, 2001.

Cutler D. and A. Lleras-Muney, Understanding differences in health behaviours by education, *Journal of Health Economics*, Vol. 29, pp. 1-28, 2006.

Dee T. and W. Evans, Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates, *Journal of labour Economics*, 21(1), pp. 178-209, 2003.

De Giorgi G., M. Pellizarri and S. Redaelli, Be as careful of the company you keep as of the books you read: peer effects in education and on the labour market, NBER Working Paper N° 14948, 2009.

De Paola M. and V. Scoppa, Peer group effects on the academic performance of Italian Students, *Applied Economics*, 42(17), pp. 2203-2215, 2010.

Deschênes O., Estimating the effects of family background on the return to schooling, *Journal of Business and Economic Statistics*, 25(3), pp. 264-277, 2007.

Duflo E., P. Dupas and M. Kremer, Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya, NBER Working Paper 14475, 2008.

Duncan O., A. Haller and A. Portes, Peer influences on aspirations: a reinterpretation, *American Journal of Sociology*, 74(2), pp. 119-137, 1968.

Engberg J. and A. R. Morral, Reducing substance use improves adolescents' school attendance, Research Report, Drug Policy Research Center, RAND Corporation, 2006.

Entorf H. and M. Lauk, Peer effects, social multipliers and migrants at school: an international comparison, unpublished document, 2006.

Evans W., W. Oates and R. Schwab, Measuring peer group effects: a study of teenage behaviour, *Journal of Political Economy*, 100(5), pp. 966-991, 1992.

Falter J.-M., Educational attainment in Switzerland: the role of observable and non-observable family factors, unpublished manuscript, 2005.

Fergusson D. and L. Horwood, Early onset cannabis use and psychosocial adjustment in young adults, *Addiction*, Vol. 92, pp. 279-295, 1997.

Fergusson D., L. Horwood and N. Swain-Campbell, Cannabis use and psychosocial adjustement in adolescence and young adulthood, *Addiction*, Vol. 97, pp. 1123-1125, 2002.

Fergusson D., L. Horwood and A. Beautrais, Cannabis and educational achievement, *Addiction*, Vol. 12, pp. 1681-1692, 2003.

Ferrer, A. and W. Riddell, The Role of Credentials in the Canadian Labour Market, *Canadian Journal of Economics*, 35, pp. 879-905, 2002.

Ferrer, A. and W. Riddell, Education, Credentials and Immigrant Earnings, *Canadian Journal of Economics*, 41, pp. 186-216, 2008.

Ferro-Luzzi G. and J. Silber, On gender differences in wages in Switzerland, *International Journal of Manpower*, 19(7), pp. 475-485, 1998.

Fertig M., Educational production, endogenous peer group formation and class composition. Evidence from the PISA 2000 study, IZA Discussion Paper N° 714, 2003.

FOPH and Addiction Suisse, Ecole et cannabis, règles, mesures et détection précoce, 2004.

Fletcher J., Social interactions and smoking: evidence using multiple student cohorts, instrumental variables and school fixed effects, *Health Economics*, Vol. 19, pp. 466-484, 2010.

Fletcher J. and D. Frisvold, College selectivity and young adult health behaviours, *Economics of Education Review*, Vol. 30, pp. 826-837, 2011.

Fuentes A., Raising education outcomes in Switzerland, OECD Economics Department Working Papers N° 838, 2011.

Garen J., The returns to schooling: a selectivity bias approach with a continuous choice variable, *Econometrica*, 52(5), pp. 1199-1218, 1984.

Gaviria A. and S. Raphael, School-based peer effects and juvenile behaviour, *Review of Economics and Statistics*, Vol. 83, pp. 257-268, 2001.

Gibbons S. and S. Telhaj, Peer and achievement in England's secondary schools, SERC Discussion Paper 1, 2008.

Grossman M., On the concept of health capital and the demand for health, *Journal of Political Economy*, 80(2), pp. 223-255, 1972.

Grossman M., The correlation between health and schooling, NBER Working Paper N° w0022, 1973.

Hanushek E., J. Kain, J. Markman and S. Rivkin, Does peer ability affect student achievement?, *Journal of Applied Econometrics*, 18(5), pp. 527-544, 2003.

Hanushek E. and L. Woessmann, Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries, Economic Journal, Vol. 116, pp. 63-76, 2006.

Harmon C., H. Oosterbeek and I. Walker, The returns to education: microeconomics, *Journal of Economic Survey*, 17(2), pp. 115-155, 2003.

Heckman J. and R. Robb, Alternative methods for evaluating the impact of interventions, in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of labour Market Data*, New York: Cambridge University Press, pp. 156-245, 1985.

Heckman, J. and E. Vytlacil, Instrumental variables methods for the correlated random coefficient model: Estimating the rate of return to schooling when the return is correlated with schooling, *Journal of Human Resources*, Vol. 23, pp. 974987, 1998.

Heckman J. and X. Li, Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000, *Pacific Economic Review*, 9(3), pp. 155-171, 2004.

Heckman J., S. Urzua and E. Vytlacil, Understanding instrumental variables in models with essential heterogeneity, *Review of Economics and Statistics*, 88(3), 2006.

Heyser C., R. Hampson and S. Deadwyler, Effects of $Delta - 9 - Tetrah.$ on delayed match to sample performance in rats: alterations in short-term memory associated with changes in task specific firing of hippocampal cells, *Journal of Pharmacology and Experimental Therapeutics*, Vol. 264, pp. 294-307, 1993.

Holland P., Statistics and causal inference (with discussion), *Journal of American Statistical Association*, Vol. 81, pp. 954-970, 1986.

Horwood L., D. Fergusson, M. Hayatbakhsh, J. Najman, C. Coffey, G. Patton, E. Silins and D. Hutchinson, Cannabis use and educational achievement: Findings from three Australasian cohort studies, *Drug and Alcohol Dependence*, Vol. 110, pp. 247-253, 2010.

Jaeger D. and M. Page, Degrees matter: new evidence on sheepskin effects in the returns to education, *Review of Economics and Statistics*, 78, pp. 733-740, 1996.

Kenkel D., Health behaviour, health knowledge and schooling, *Journal of Political Economy*, Vol. 99, pp. 287-305, 1991.

Krauth B., Peer effects and selection effects on smoking among Canadian youth, *Canadian Journal of Economics*, 38(3), pp. 735-757, 2005.

Kronig W., Die systematische Zufälligkeit des Bildungserfolgs. Theoretische Erklärungen und and empirische Untersuchungen zur Lernenwicklung und zur Leistungbewertung in unterschiedlichen Schulklassen, Berne: Haupt, 2007.

Lazear E., Education production, *Quarterly Journal of Economics*, 116(3), pp. 777-803, 2001.

Lefgren L., Educational peer effects and the Chicago public schools, *Journal of Urban Economics*, 56(2), pp. 169-191, 2004.

Manski C., Identification of endogenous social effects: The Reflection Problem, *Review of Economic Studies*, 60(3), pp. 531-542, 1993.

Manski C., Economic analysis of social interactions, *Journal of Economic Perspectives*, Vol. 14, pp. 115-136, 2000.

Matsuda L., T. Bonner and S. Lolait, Localization of cannabinoid receptor messenger RNA in rat brain, *Journal of Comparative Neurology*, Vol. 327, pp. 535-550, 1993.

McCaffrey D., R. Pacula, B. Han and P. Ellickson, Marijuana use and high school dropout: the influence of unobservables, *Health Economics*, Vol. 19, pp. 1281-12999, 2010.

McEwan P., Peer effects on student achievement: evidence from Chile, *Economics of Education Review*, Vol. 22, pp. 131-141, 2003.

McVicar D. and A. Polanski, Peer effects in UK adolescent substance use: never mind the classmates?, Melbourne Institute Working Paper Series, N° 8/12, 2012.

Menghrajani P., K. Klaue, F. Dubois-Arber and P.-A. Michaud, Swiss adolescents' and adults' perceptions of cannabis use: a qualitative study, *Health Education Research*, 20(4), pp. 476-484, 2004.

Mincer, J., Investment in human capital and personal income distribution, *Journal of Political Economy*, 66(4), pp. 281-302, 1958.

Mincer J., *Schooling, experience, and earnings*, New York: Columbia University Press, 1974.

Moriarty J., D. Mc Vicar and K. Higgins, Peer effects in adolescent cannabis Use: it's the friends, stupid, Melbourne Institute Working Paper Series, N° 27/12, 2012.

Moser U. and D. Angelone, PISA 2006: Porträt des Kantons Basel-Landschaft, Oberentfelden: Sauerländer, 2008.

Mushkin, S., Health as an investment, *Journal of Political Economy*, 70(2), pp. 129-157, 1962.

OECD, OECD Economic Surveys: Switzerland 2009, OEDC Publishing, 2009.

OECD, *Education at a Glance 2012: OECD indicators*, OECD Publishing, 2012.

OECD, OECD Economic Surveys: Switzerland 2013, OECD Publishing, 2013.

Pacula R. , J. Ringel and K. Ross, Does marijuana use impair human capital formation?, NBER Working Paper N° 9963, 2003.

Park J., Estimation of sheepskin effects using the old and the new measures of educational attainment in the current population survey, *Economic Letters*, 62, pp. 237-240, 1999.

Psacharopoulos G. and H. Patrinos, Returns to investment in education: a further update, *Education Economics*, 12(2), pp. 111-134, 2004.

Rangvid B., Educational peer effects. Quantile regression: evidence from Denmark with PISA 2000 data, Do school matter?, PhD thesis, Aarhus School of Business, 2004.

Renna F., The economic cost of teen drinking: late graduation and lowered earnings, *Health Economics*, Vol. 16, pp. 407-419, 2006.

Roebuck M., M. French and M. Dennis, Adolescent marijuana use and school attendance, *Economics of Education Review*, Vol. 23, pp. 133-141, 2004.

Rosenbaum P. and D. Rubin, The central role of the propensity score in observational studies for causal effects, *Biometrika*, Vol. 70, pp. 41-55, 1983.

Rosenbaum P. *Observational studies*, Second Edition, New York: Springer, 2002.

Roy A., Some thoughts on the distribution of earnings, *Oxford Economic Paper*, Vol. 3, pp. 135-146, 1951.

Sacerdote B.I., Peer effects with random assignment, *Quarterly Journal of Economics*, Vol. CXVI, pp. 681-704, 2001.

Sacerdote B.I., Peer effects in education: How might they work, how big are they and how much do we know thus far?, in E. Hanushek, S. Machin and L. Woessmann (eds.), Vol. 3, *Handbook of the Economics of Education*, Chapter 4, 2011.

Schneeweis N. and R. Winter-Ebmer, Peer effects in Austrian schools, *Empirical Economics*, Vol. 32, pp. 387-409, 2007.

Selvan M., M. Ross, A. Kapadia, R. Mathai and S. Hira, Study of perceived norms, beliefs and intended sexual behaviour among higher secondary school students in India, *AIDS Care*, 13(6), pp. 779-788, 2001.

Stephens R., T. Babor, R. Kadden and M. Miller. The Marijuana Treatment Project: Rationale, design, and participant characteristics, Addiction, 97(Suppl 1), pp. 109-124, 2002.

Suhrcke M. and C. de Paz Nieves, The impact of health and health behaviours on educational outcomes in high-income countries: a review of the evidence, WHO Regional Office for Europe, 2011.

Spence M., Job market signaling, *Quarterly Journal of Economics*, 87(3), pp. 355-374, 1973.

Sund K., Estimating peer effects in Swedish high schools using school, teacher, and student fixed effects, Swedish Institute for Social Research, Working Paper N°8, 2007.

Suter S., Bildung oder Begabung? Eine ökonometrische Untersuchung zu den Ursachen individueller Lohnunterschiede, PNR FNS 43, 2005.

Swiss Federal Statistical Office (SFSO), *PISA 2006: Les competences en sciencess et leur role dans la vie: rapport national*, FSO, 2007.

Swiss Federal Statistical Office (SFSO), Enquête suisse sur la santé 2007. Premiers résultats. 2008.

Trogdon J., J. Nonnemaker and J. Pais, Peer effects in adolescent overweight, *Journal of Health Economics*, 27(5), pp. 1388-1399, 2008.

Trostel P., I. Walker and P. Woolley, Estimates of the economic return to schooling for 28 countries, *Labour Economics*, 9(1), pp. 1-16, 2002.

Tsai S. and Y. Xie, Changes in earnings returns to higher education in Taiwan since the 1990s, *Population Review*, Vol. 47, pp. 1-20, 2008.

Vellacott M. and S. Wolter, Equity in the Swiss education system: dimensions, causes and policy responses, National report from Switzerland contributing to the OECD's review of "Equity in Education", CSRE, 2004.

Vigdor J. and T. Nechyba, Peer effects in North Carolina public schools, in L. Woessman and Paul P. Peterson (eds.), *Schools and the equal opportunity problem*, CESifo Seminar Series, MIT Press Books, 2007.

Willis R. and S. Rosen, Education and self-selection, *Journal of Political Economy*, Vol. 87, pp. 7-36, 1979.

Willis R., Wage determinants: A survey and reinterpretation of human capital earnings functions, in O. Ashenfelter and R. Layard (eds.), Vol. 1,*Handbook of labour Economics*, Chapter 10, 1987.

Wolter S. and B. Weber, La rentabilité de la formation: un indicateur économique essentiel à l'enseignement, *La Vie Economique*, Octobre 2005.

Yamada T., M. Kendix and T. Yamada, The impact of alcohol consumption and marijuana use on high school graduation. *Health Economics*, Vol. 5, pp. 77-92, 1996.

Chapter 2

Peer effects in Swiss lower secondary schools

This paper estimates educational peer effects in Swiss lower secondary schools where different ability tracking designs coexist. Using a cross-sectional survey based on standardized questionnaires, the structure and magnitude of peer effects among classmates are analyzed. The identification strategy relies on ability track fixed effects to control for selectivity issues and quantile regressions to determine returns to peer effects along the conditional ability distribution. Results indicate positive, small but significant average peer effects in reading and sciences whereas the average peer coefficient in mathematics is not significant. In reading and sciences, non-linear peer effects suggest that low-achieving students benefit most from peer effects whereas high-achieving students in mathematics obtain better school performances when they are placed together with similar peers. Class diversity does not affect the overall performance of the classmates but reduces the family background effect on school performances, whatever the field considered. These empirical findings show that mixing students in reading and sciences classes could enhance efficiency and equity while a similar practice in mathematics courses could only improve equity without any gain in efficiency.

2.1 Introduction

The magnitude and nature of peer effects are a prominent argument when defining educational policy such as ability tracking, anti-poverty programs, or classroom organization. Since the seminal Coleman report (1966), a growing literature has documented the importance of social interactions on educational outcomes (Hoxby, 2000; Sacerdote, 2001; McEwan, 2003; Vigdor and Nechyba, 2007; Burke and Sass, 2008; Ammermueller and Pischke, 2009; De

Paola and Scoppa, 2010). The relationship between class composition variables and scholastic achievement can provide valuable insights regarding the optimal school design which generally boils down to choosing between two opposite systems: the *selective* (or *tracking*) system where students are separated into different ability groups and the *comprehensive* (or *mixing*) system where students follow ability-mixed classes.

Researches on peer effects and ability tracking are closely related because the existence of social interactions is a crucial element when discussing students' reallocation and the productivity of educational processes. Proponents of a selective system consider that tracking students maximizes student outcomes (measured through the accumulation of cognitive aptitudes) by forming more homogeneous classes where the teacher can adapt his or her program to different kinds of students by focusing on their specific needs. Maximizing efficiency, however, is not the only concern of schooling policy. Other objectives like increasing life chances and social cohesion have to be satisfied too. Advocates of a comprehensive system insist on the fact that mixing students increases educational opportunities by giving the possibility to less-endowed students to benefit from high-achieving peers through direct learning, identification mechanisms or free-riding behaviours. Moreover, the limited means to contest a tracking decision can exacerbate educational inequality (OECD, 2013). At the same time, some parents worry that disruptive students may affect adversely student's behaviour and test performances, a point underlined by Lazear (2001).

There is no clear evidence on which system is definitively the best in terms of efficiency or equity. On the one hand, a comprehensive system might enhance efficiency if students with learning difficulties benefit more from being placed together with high-ability peers while not creating adverse effects on the overall performance of the classroom. On the other hand, even if we expect a mixing system to increase the intergenerational transmission of human capital, ability tracking might improve equity if mobility across ability tracks is encouraged.

The main objective of this research is to find out if grouping students in a completely non-selective way at the Swiss lower secondary level could improve efficiency and equality of opportunity. During the past few years, tracking policies in Switzerland have been subject to several criticisms regarding equity, efficiency, or labour market needs (CSRE, 2010). In this context, we witness the development of within-class sorting which consists in separating students from the same class in different ability level groups for specific disciplines. However, full-tracking policies are still dominant and remains the most attractive school design for the majority of Swiss cantons. This statement is confirmed by the fact that the harmonization of school designs has not been discussed in the concordat *Harmos* whose objective, however, is to harmonize some practices in compulsory education, especially at the lower secondary level.

This study exploits the data from the PISA Swiss national sample 2006 which contains a relevant set of variables on individuals (background variables, test score), schools (location, selection procedure) and tracking systems (tracking design, ability tracks). Linear peer effects are first estimated by OLS. However, this model does not control for within-school sorting, which is a common practice in Swiss lower secondary schools. Therefore, the estimation procedure introduces ability track fixed effects in the identification strategy to account for selectivity issues. This is an improvement compared to previous literature where data limitations often do not allow the researchers to control directly for selective procedures within the school (Mc Ewan, 2003; Vigdor and Nechyba, 2007). In a second step, non-linear peer effects are estimated by quantile regressions to determine the magnitude of peers effects along the conditional ability distribution, which is of primary importance to draw policy recommendations regarding efficiency and equity criteria.

Empirical findings report that accounting for endogeneity issues is essential to obtain unbiased peer estimates. In comparison with the traditional OLS model where I obtain positive, strong and significant peer effects in all fields considered, the introduction of ability track fixed effects reduces the magnitude of the peer coefficients in reading and sciences while the peer coefficient in mathematics loses its significance. In reading and sciences, non-linear peer effects suggest that low-achieving students benefit more from peer effects whereas high-achieving students in mathematics obtain better school performances when they are placed together with similar peers. Class diversity does not affect the overall performance of the classmates but reduces the family background effect on school performances, whatever the field considered. These empirical findings show that mixing students in reading and sciences classes could enhance efficiency and equity while a similar practice in mathematics courses could only improve equity without any gain in efficiency.

The rest of the paper is structured as follows. Analytical background, literature review and empirical evidence are presented in the second part. Data are described in part three. I discuss the empirical framework in part four. The fifth part reports the results and the last part is devoted to the conclusion.

2.2 Background and literature review

The estimation of peer effects entails a number of econometric difficulties, including the endogeneity of the school or class choice (*selection* bias), the reciprocal influence between classmates' behaviour (*simultaneity* bias), and the fact that common unobserved factors (e.g., teacher quality or spatial segregation) jointly determine individual and classmates' performances (*omitted variable* bias). These methodological constraints explain why empirical evidence on peer effects is rather mixed, and their potential to inform policy limited. Various definitions of the school outcomes, choices of peer reference groups, and data limitations further complicate the task of finding a consensus.

2.2.1 Choice of the reference group

The level at which a peer group is defined depends essentially on the survey design. Studies working with PISA international data, which do not include class identifiers, assess the influence of schoolmates on student outcomes. While Fertig (2003) identifies peers as schoolmates, Rangvid (2004) and Schneeweis and Winter-Ebmer (2007) determine peers as pupils who are in the same school and grade. When data provide information at a more disaggregated level, some researchers estimate peer effects at the class level (Hoxby, 2000; Hanushek et al., 2003; Burke and Sass, 2008; Sund, 2007; Ammermueller and Pischke, 2009) while some others are interested in the influence of subgroups within the classroom, e.g., the share of pupils from dissolved families (Bonesronning, 2008) or the share of repeaters (Lavy et al., 2009). However, literature is inconclusive regarding the group level at which peer effects are the strongest (Betts and Zau, 2004; Vigdor and Nechyba, 2007; Burke and Sass, 2008).

2.2.2 Identification strategies

When we discuss peer effects, it is crucial to determine which kinds of social interactions we are talking about and separate them from non-social influences. According to the conceptual framework of Manski (1993, 1995, 2000), there are three arguments that may explain why students belonging to the same peer reference group tend to behave similarly:

- *Endogenous effects* exist when the behaviour of one's peers (e.g., effort, motivation, inspiration, or commitment) influences personal behaviour. Such contemporaneous interactions generate a social multiplier effect because the consequences of introducing a schooling policy not only affect the behaviour of the students of interest but also affect the behaviour of all school- or classmates through their reciprocal influences.

- *Contextual effects* occur when the exogenous characteristics of the peer group (e.g., ability[1], socioeconomic status, or gender) influence the individual's behaviour. Here, however, policy interventions do not create a multiplier effect because these social interactions rely on attributes unaffected by the current behaviour of the individuals.

- *Correlated effects* arise if individuals in the same reference group behave similarly because they face similar environments or share similar characteristics (e.g., teacher quality, living in the same socioeconomic area). Whereas endogenous and contextual effects result from social interactions, correlated effects are not a social phenomenon.

The first endogeneity issue is related to reflexivity because individual and peers' outcomes influence each other (endogenous effects). The usual way to reduce this *reflection* bias (Manski, 1993) consists in using a lagged peer outcome as instrument (Hanushek et al., 2003; Betts and Zau, 2004; Vigdor and Nechyba, 2007; Burke and Sass, 2008; De Paola and Scoppa, 2010). However, this strategy entails two main problems, i.e., the lagged achievement of peers ignores the impact of current peer effort and the presence of serial correlation may still affect the parameter estimates.

A second concern lies in the fact that peer background itself affects peer outcome through individual outcome. Consequently, *collinearity* problems may arise and the related coefficients cannot be identified. The existing literature addresses this issue in two different ways. First, numerous contributions rely on a reduced form model that incorporates only one peer effect variable. Consequently, this variable captures a total social effect which does

[1]When ability is measured before the peer group formation, we can define ability as a contextual (or pretreatment) characteristic. However, when ability is measured after the peer group formation, we can use the test score information as a proxy for peers' performances. In other words, it allows for measuring endogenous (or during treatment) effects.

not account for the precise nature of social interactions (Sacerdote, 2001; Ammermueller and Pischke, 2009; De Paola and Scoppa, 2010). An alternative approach consists in assuming that the peer reference group is individual-specific, i.e., some of the peer groups overlap with one other[2](De Giorgi et al., 2009; Bramoullé et al., 2009). As a result, we can use the educational outcome of the excluded school- or classmates as an instrumental variable for peer achievement. However, such an approach, which may require the use of spatial econometrics, is beyond the scope of this paper, especially due to the data at hand.

Correlated effects are not modeled directly in the econometric model. However, they play an important role if peer group composition is also determined by unobserved factors. For example, if students with higher unobserved abilities or resources are more prone to be oriented towards higher-ability tracks or better schools, peer group composition is not random and peer effect estimates cannot be interpreted in causal terms. Because natural experiments are still in short supply, different strategies has been considered in the literature. Analyzing the case of Denmark where students are mixed during compulsory schooling, Rangvid (2004) estimates a regression model including numerous background attributes and school characteristics to reduce as much as possible the endogeneity problem. The author takes advantage of a large set of data that combines both PISA 2000 and additional register data. An alternative approach to address selectivity bias is to adopt an instrumental variable strategy to explicitly model the enrollment process (Fertig, 2003; Lefgren, 2004; Hoxby and Weingarth, 2005; Gibbons and Telhaj, 2008; Atkinson et al., 2008; De Paola and Scoppa, 2010). Another common strategy is to use fixed effects methods. Many researchers have employed school fixed effects to control for school differences, especially when tracking occurs at the school level (Lefgren, 2004; Schneeweis and Winter-Ebmer, 2007; Gibbons and Telhaj, 2008). This is an appropriate strategy as long as students are not sorted by ability within the school. If this is the case, school fixed effect estimates could be still biased by uncontrolled within-school ability sorting (Mc Ewan, 2003; Vigdor and Nechyba, 2007; Zabel, 2008). A better solution to rule out selection bias consists in combining school fixed effects with teacher and student fixed effects (Sund, 2007; Carman and Zhang, 2008, Burke and Sass, 2008).

[2]Expressed differently, individuals interact no longer in defined groups (where individuals are affected by all other members but by none outside it) but within a social network (where interactions are interdependent)

2.2.3 Efficiency and equity considerations

Linear and average peer effects, however, are not informative regarding policy recommendations. Indeed, nonlinear peer effects and measures of peer diversity are more prone to detect which kinds of students benefit most from social interactions and what is the impact of peer group heterogeneity on individual's school performance. Both information are necessary to draw policy advices regarding efficiency and equity criteria. Table 2.1, largely inspired by the tables presented in Gibbons and Telhaj (2008) and Sacerdote (2011), summarizes the main findings obtained by some important papers closely related to this study. Some of them are discussed below.

Efficiency

Schools and teaching staff have to use efficiently their resources given that means available are limited. They must find solutions which ensure the best success for the learners. We focus here on the question of school placement policies (e.g., tracking versus comprehensive systems). In order to be efficiency-enhancing, a comprehensive system has to meet two conditions. First, peer effects should be stronger for less-endowed students as compared to more-endowed ones. Second, peer diversity should not negatively affect the overall performance in the classroom. In short, the objective is to determine if a reallocation of students in a comprehensive way leads to a *Pareto improving* situation[3].

Numerous studies account for both non-linearity in peer effects and peer heterogeneity to analyze if school desegregation is preferable to school segregation or not. Using country fixed effects and classroom random effects, Vandenbergue (2002) estimates peer effects impact on science and math test scores of secondary school students among 17 OECD countries. Non-linearity is measured in two different ways: first, by introducing a quadratic term for the peer variable in the educational production, and second, by an interaction term between the individual socioeconomic profile and the peer variable. Peer heterogeneity is measured through the standard deviation of the peer variable. Schneeweis and Winter-Ebmer (2007) estimate peer effects in Austrian secondary schools by using both PISA 2000 and 2003 data. Non-linearity is captured first by an interaction between individual and peers' parental background and second, through a quantile regression model. Peer heterogeneity is measured through the standard deviation of the peer variable. Both stud-

[3] In economics, a *Pareto improving* action corresponds to an action that harms no one and helps at least one person.

ies find that peer effects are stronger for less gifted students, but that an increase in peer heterogeneity leads to some adverse effects on student's performances. Consequently, these results cannot give clear recommendations concerning the optimal allocation of students in order to enhance efficiency.

Analyzing the non-selective Danish school system, Rangvid (2004) also considers a quantile regression framework and shows that low-achieving pupils benefit most from schoolmates' interactions, while high-ability students lose nothing from the diversity in the student body. Relying on a rich dataset from Swedish high schools, Sund (2007) finds that low-achieving students benefit most from an increase in both peer average and peer heterogeneity. Non-linear peer effects are captured through the interaction term between mean and standard deviation in peer achievement within the classroom. Both studies satisfy the Pareto improving conditions and indicate that a comprehensive school design is an appropriate system to enhance efficiency.

On the contrary, some findings reveal that high-achieving pupils benefit most from the presence of other high-ability students (Hoxby, 2000; Hanushek et al., 2003; Hoxby and Weingarth, 2005; Gibbons and Telhaj, 2008; Burke and Sass, 2008; Lavy et al., 2009). In such a case, ability tracking appears as the optimal policy to increase efficiency given that individuals perform better when they are sorted with similar peers.

Equity

In parallel to efficiency considerations whose objective is to maximize the accumulation of cognitive skills, the schooling system should also offer equal access to schooling opportunities (equality of opportunity), equal treatment of all persons in training (equality of treatment) and ensure acquisition of basic knowledge for everyone (equality of achievement). We discuss here the concept of equality of opportunity which can be designed as the strategy adopted by the schooling system to face social inequalities.

Equality of opportunity in educational outcomes is a crucial element in the discussion on social cohesion and intergenerational mobility achieved by societies. Many studies refer to this concept when discussing ability tracking (Rees et al., 2000; Figlio and Page, 2002; Bauer and Riphahn, 2006; Hanushek and Woessman, 2006, Brunello and Checci, 2007; Schuetz et al., 2008). The main idea is to analyze how students' educational performance is related to their family background and if ability tracking reinforces or weakens this relation (Schuetz et al., 2008). It is worth mentioning, however, that none of the studies mentioned above rely explicitly on a peer effect framework.

To my knowledge, the study of Raitano and Vona (2011) is the first empirical contribution in the literature that proposes a specific methodology based on peer effects to determine the impact of tracking on equality of opportunity[4]. The main objective of their paper is to assess how school selection procedures and peer variables reinforce or weaken the parental background effect. Using PISA 2006 survey for OECD countries, they show that increasing peer diversity would reduce the parental background effect and therefore would improve equality of opportunity. In their preferred specification which includes country fixed effects, the authors find that an increase by one standard deviation of peer heterogeneity is associated with a reduction of 8.4% in the average family background effect.

2.3 Data

2.3.1 PISA national sample

Initiated by the OECD in 2000, the Program for International Student Assessment (PISA) is an internationally standardized assessment of knowledge and skills acquired by students at the end of compulsory education. Until now, five assessments have been carried out, i.e., every three years. At each wave, a major field (reading, mathematics, or sciences) is examined in depth. Moreover, OECD allows each participating country to generate complementary samples. Consequently, since PISA 2000, Switzerland has taken advantage of this opportunity to generate a PISA *national* sample. In contrast to the international sample that focuses only on 15-year-old students, the PISA Swiss sample is exclusively composed of students attending the ninth grade (i.e., the last year of compulsory education) and additional variables are available to lead a regional analysis.

This study uses the supplementary PISA 2006 data provided by the Swiss Federal Statistical Office (SFSO)[5]. This dataset allows a peer effect analysis at the class level given that data indicate the school, the ability track and the class the student attends. Moreover, I have some information on the

[4]Hidalgo-Hidalgo (2009) proposes a very interesting theoretical paper which addresses the questions of tracking, peer effects and equality of opportunity.

[5]"Base de données suisse PISA 2006 pour la 9ème année" (OFS/CDIP).

Table 2.1: A selective review of closely related papers

Studies	**Context**	**Outcome**	**Peer measure**	**Peer group**	**Methodology**	**Magnitude of average peer effects**	**Non-linearity and Peer Heterogeneity (PH)**
Vandenbergue (2002)	Secondary schools, OECD countries	Test scores (maths and sciences)	Family SES	Class	Country FE and classroom RE	Significant evidence	Diminishing returns in peer effects and negative impact of PH
Mc Ewan (2003)	Secondary schools, Chile	Test scores (Spanish)	Mother's education	Class	School FE	1sd → 0.27sd	Diminishing returns in peer effects
Fertig (2003)	Secondary schools, US	Test scores (reading)	Test scores heterogeneity	School	IV strategy	-	Negative impact of PH
Lefgren (2004)	Chicago public schools, US	Test scores (reading and maths)	Prior test scores	Class	School-year FE and IV strategy	1sd → 0.024sd	-
Rangvid (2004)	Secondary schools, Denmark	Test scores (reading)	Mother's education	School-grade	Additional controls	1sd → 0.08sd	Diminishing returns in peer effects and no detrimental impact of PH
Schneeweis and Winter-Ebmer (2007)	Lower secondary schools, Austria	Test scores (reading and maths)	Parental occupational status	School-grade	School type FE and school FE	No evidence with school FE	Diminishing returns in peers effects for reading and mixed impact of PH
Vigdor and Nechyba (2007)	North Carolina 5th graders, US	Test scores (reading and maths)	Prior test scores	Class	School FE/apparent random assignment	1sd → 0.03sd	-
Sund (2007)	High schools, Sweden	Test scores (GPA)	Prior test scores	Class	School, teacher and student FE	1sd → 0.08sd	Diminishing returns in peer effects and positive effects of PH
Carman and Zhang (2008)	Middle schools, China	Course grade	Prior subject grade	Class	Teacher and student FE	0.1sd → 0.04 sd (maths)	Mixed evidence for non-linearity but no detrimental impact of PH
Gibbons and Thelaj (2008)	State secondary schools, England	Average test scores	Prior test scores	School-grade	Year-to-year changes in school composition	No evidence	Increasing returns in peer effects
Burke and Sass (2008)	Florida public schools, US	Test score gains	Fixed characteristics	Class	Student and teacher FE	1pt → 0.044pt	Increasing returns in peer effects and negative impact of PH
Ammermueller and Pischke (2009)	Europe primary schools	Test scores (reading)	Average peer characteristics	Class	School FE	1sd → 0.07sd	-
De Paola M. and V. Scoppa (2010)	Calabria University, Italy	Secon Level grade	First level grade	Class	2SLS	1sd → 0.19sd	-
Raitano and Vona (2011)	OECD countries	Test score (sciences)	Books at home	School	Country FE and pseudo school FE	-	Decreasing returns in peer effects and PH reduces family background effect

differentiated-ability level courses the student follows when ability tracking occurs within the class. Initially composed of 20,456 pupils, the sample size was reduced for different reasons. First, as some cantons have not opted for a class-based sampling, only 15 cantons are considered in the empirical analysis[6]. Second, classes with less than 6 students are excluded from the sample. Third, two classes per ability track are necessary to ensure within-ability track variation and to avoid fixed effects methods absorbing any variation at the fixed effect group level (i.e., ability track level). At the end, the analytical sample consists of 14,081 students. Except for the core variables measuring test score, family background and peer quality for which missing values are dropped out of the sample, missing values for the other (categorical) variables are treated as another category in order to ensure representativeness in the sample. The following sections define the variables used for the empirical analysis. Summary statistics are presented in Table 2.2.

2.3.2 Educational outcomes

Educational performances in reading, mathematics, and sciences are measured through PISA test scores. As it is common in the literature using PISA or TIMSS (Trends in International Mathematics and Sciences Study) data, I use the first plausible value for the students' actual score[7]. The scale of these variables has been standardized at the OECD level with an average of 500 points and a standard deviation of 100 points.

2.3.3 School track design and ability tracks

The Swiss educational system is organized in a federalist way and therefore involves different actors: Confederation (i.e., central government), cantons (i.e., sub-national governments) and communes (i.e., municipalities). Based on the subsidiarity principle, cantonal and communal authorities enjoy a large degree of autonomy regarding the structure of their schooling system, especially at the lower secondary level. As a result, ability tracking practices differ between and within cantons. If we regroup these heterogeneous practices on the basis of unified criteria, we can distinguish between three school

[6]Participating cantons are Aargau, Bern, Basel Land, St.-Gallen, Schaffhausen, Thurgau, Zürich, Valais, Vaud, Genève, Neuchâtel, Jura, Fribourg, Tessin and Graubünden.

[7]PISA test scores are based on too few items to give a realistic estimation of students' ability. For that purpose, a probability distribution for identifying students' ability is estimated. Plausible values represent random draws from this empirically derived distribution of proficiency values that are conditional on the observed values of the assessment items and the background variables.

track designs[8]:

- The *separated* system tracks students in different school types according to the school performances, i.e., ability tracking occurs at the school level. Consequently, each school has its own curricula and teaching staff. Some schools prepare pupils for university entrance, other for vocational formation (e.g., apprenticeship or professional matura). Expressed differently, the pupils enrolled in the same school follow the same ability track.

- The *cooperative* system sorts students into different ability tracks within a given school, i.e., ability tracking occurs between classes within the same building. As in the separated system, each ability track prepares pupils for different schooling pathways. The advantage of such system is to facilitate the mobility between ability groups which are so-defined located in the same school.

- The *integrative* system mixes pupils in a comprehensive way, except for core subjects like reading and mathematics where pupils from the same class are sent to different level groups on the basis of their aptitudes, i.e., ability tracking occurs within the classroom. Students following high-ability classes are prepared for an academic matura while those following middle- and low-ability classes are more prone to attend a vocational formation.

The separated and cooperative systems are defined as homogeneous given that all students from the same class belong to the same ability track. In such a case, we refer to the concept of *ability grouping*. In opposition to the two former designs, the integrative system is defined as heterogeneous on the grounds that it combines both mixed-ability classes (e.g., sciences) and *level grouping* (e.g., mathematics and reading). Level grouping occurs when students from the same class can belong to different ability tracks regarding the subject of differentiation, e.g., higher-ability track in mathematics but lower-ability track in language instruction. A combination of ability grouping with level grouping also exists in some cantons.

For the empirical analysis, the school track design is accounted for with a dummy variable taking the value 0 if the student belongs to the homogeneous system and the value 1 if she belongs to the heterogeneous system. Unfortunately, due to data limitations, I cannot make a distinction between the separated and the cooperative system. Ability track level is divided into

[8] A short overview is provided in Table 2.9 in the Appendix.

three categories, i.e., high-, middle-, and low-ability track. For sciences, this variable contains a fourth category called "mixed-ability track" given that this field is not subject to level grouping in the integrated system.

2.3.4 Peer characteristics variables

In the homogeneous system, the peer reference group consists of pupils who are in the same class and - by definition - on the same ability track whatever the field considered. For the integrative system, the peer reference group refers to pupils who are in the same level group (reading and/or mathematics) or in the same class (sciences). For that purpose, I have to create field-specific peer groups on the grounds that a student in the integrated system does not necessarily have the same classmates in mathematics or reading. The advantage of such a strategy is to have a relevant set of peers for each situation.

The peer quality variable is measured by the mean parental economic, social, and cultural status in the reference group. This index which serves as proxy for parental background is derived from variables related to parental education, parental occupational status, and an index of home possessions (desk for study, educational software, books, computer, calculator, etc.). Similarly, peer heterogeneity is measured by the standard deviation of the peer variable in the reference group.

2.3.5 Control variables

At the individual level, I control for gender, age, migration background, and own parental background. I add a variable reporting if the language spoken at home is a Swiss national language or not. Parental expectations and the importance attached by parents to the field considered are included in the regression model to reduce the unobserved heterogeneity related to parents' educational preferences.

At the school level, I include a set of school characteristics and a measure of school selection procedure. The former are represented by school size, school location and the proportion of teachers with a university degree in pedagogy while the latter is a school admittance variable based on student's prior records. Finally, I control for the size of the class[9] where classes with

[9]In the integrated system, class size refers to number of students following the same differentiated-level course.

less than six students are excluded from my analysis.

Table 2.2: Variables description and summary statistics

Variables	**Description**	**Mean**	**s.d.**
Test scores			
Reading	Standardized test scores (mean of 500pts and sd of 100pts)	506.608	81.428
Mathematics	Standardized test scores (mean of 500pts and sd of 100pts)	540.273	85.203
Sciences	Standardized test scores (mean of 500pts and sd of 100pts)	516.920	86.212
Peer characteristics			
Peer quality (reading)	Mean parental economic, social and cultural status in the peer reference group	0.170	0.435
Peer quality (mathematics)	Mean parental economic, social and cultural status in the peer reference group	0.171	0.435
Peer quality (sciences)	Mean parental economic, social and cultural status in the peer reference group	0.169	0.432
Peer heterogeneity (reading)	Standard deviation of parental economic, social and cultural status in the peer reference group	0.813	0.190
Peer heterogeneity (mathematics)	Standard deviation of parental economic, social and cultural status in the peer reference group	0.813	0.191
Peer heterogeneity (sciences)	Standard deviation of parental economic, social and cultural status in the peer reference group	0.812	0.186
Parental background			
Parental background	Parental economic, social and cultural status (index with mean of 0 and sd of 1)	0.171	0.872
Background characteristics			
Parental expectation	=1 if expectation is higher education graduated, =0 otherwise	0.172	0.378
	=2 if missing	0.107	0.308
Parental value	=1 if important	0.904	0.295
(reading)	=2 if missing	0.020	0.141
Parental value	=1 if important	0.900	0.300
(mathematics)	=2 if missing	0.020	0.140
Parental value	=1 if important	0.532	0.499
(sciences)	=2 if missing	0.027	0.162
Other language at home	=1 if none of Swiss official language is spoken at home	0.134	0.340
	=2 if missing	0.036	0.187
Migration background	Ref. cat.= natives		
	=1 if immigrant	0.122	0.327
	=2 if immigrant parents	0.095	0.294
	=3 if missing	0.015	0.122
Age	Student's age in years	15.695	0.625
Female	=1 if female	0.504	0.500
School track design			
heterogeneous system	=1 if heterogeneous system , =0 if homogeneous system	0.126	0.332
Ability tracks characteristics			
Ability track level	Ref. cat.= low-ability track		
(reading)	=1 if middle-ability track	0.306	0.460
	=2 if high-ability track	0.397	0.489
Ability track level	Ref. cat.= low-ability track		

(mathematics)	=1 if middle-ability track	0.307	0.461
	=2 if high-ability track	0.403	0.491
Ability track level	Ref. cat.= mixed-ability track		
(sciences)	=1 if low-ability track	0.218	0.413
	=2 if middle-ability track	0.297	0.457
	=3 if high-ability track	0.359	0.480
Class size			
Classe size (reading)	Number of students in the class	15.560	8.983
Classe size (mathematics)	Number of students in the class	15.607	8.941
Classe size (sciences)	Number of students in the class	15.730	8.851
Schools characteristics			
Teacher quality	=1 if more than 50% of teachers held a university degree in pedagogy, =0 otherwise	0.556	0.497
	=2 if missing	0.200	0.399
School size	Ref. cat.= less than 500 students		
	=1 if between 500 and 1000 students	0.444	0.497
	=2 if more than 1000 students	0.085	0.279
	=3 if missing	0.035	0.183
School location	Ref. cat.= village		
	=1 if small town	0.480	0.500
	=2 if town	0.299	0.458
	=3 if city	0.078	0.268
	=4 if missing	0.010	0.100
School admittance			
Admission procedure	=1 if based on prior student's records, =0 otherwise	0.353	0.478
	=2 if missing	0.022	0.145
	Nb of schools	297	
	Nb of classes	893	
	Nb of students in the homogeneous system	12,309	
	Nb of students in the heterogeneous system	1,772	
	Nb of students (total)	14,081	

2.4 Empirical analysis

This section is organized as follows. First, I propose a reduced form model that estimates the mean impact of classmates' quality on educational achievement by using OLS and ability track fixed effects, respectively. Second, I account for non-linearity in peer effects and peer heterogeneity to determine if mixing students can be an efficiency-enhancing policy. Finally, I move to the equity effect by investigating if class heterogeneity has an equalizing impact on student's performance with regards to her parental background.

2.4.1 Identification of mean peer effects

The OLS specification serves as baseline model. The basic linear-in-means model can be represented as follows:

$$\begin{aligned} Y_{icks} &= \beta_0 + \beta_1 \bar{PB}_{(-i)cks} + \beta_2 PB_{icks} + \beta_3 X_{icks} + \beta_4 C_{cks} \\ &+ \beta_5 A_{icks} + \beta_6 SD_s + \beta_7 S_s + \beta_8 SP_s + \epsilon_{icks} \end{aligned} \tag{2.1}$$

where Y_{icks} is the test performance of student i in class c, ability track k and school s, $\bar{PB}_{(-i)cks}$ is the parental background of classmates, excluding the contribution of student i, PB_{icks} is the parental background of student i, X_{icks} is a vector of individual and other background characteristics (i.e., gender, age, immigration status, language at home, parental taste for schooling and parental expectations), C_{cks} is the size of the class, A_{icks} represents the ability track level the student follows, SD_s is the type of school design (i.e., homogeneous or heterogeneous), S_s are school characteristics, SP_s is a measure of school selection procedure and ϵ_{icks} is an error term. Equation (2.1), however, might suffer from selectivity problems, i.e.,

$$Cov(\bar{PB}_{(-i)cks}, \epsilon_{icks}) \neq 0$$

Consequently, estimates of β_1 can be biased. Indeed, even with a rich set of background variables, unobserved factors may still influence the peer group composition. In Switzerland, ability track assignment is based on different criteria such as prior test performances, teacher recommendations, or parental endorsement which are generally not observed by the researcher. In order to reduce selectivity issues, I introduced ability track fixed effects in equation (2.1). My preferred specification is then:

$$\begin{aligned} Y_{icks} &= \beta_0 + \beta_1 \bar{PB}_{(-i)cks} + \beta_2 PB_{icks} + \beta_3 X_{icks} + \beta_4 C_{cks} \\ &+ \underbrace{\mu_k + \nu_{icks}}_{\epsilon_{icks}} \end{aligned} \tag{2.2}$$

where μ_k is an ability track specific component and ν_{icks} is an idiosyncratic error term.

2.4.2 Efficiency analysis

A comprehensive system needs to meet two conditions to enhance efficiency, i.e., decreasing returns in peer effects and no negative impact of peer diversity on student's achievement. I consider two strategies to account for non-linearity in peer effects.

The first approach interacts the peer variable with the parental background to detect if peer effects are stronger for pupils with disadvantaged parental background. I also introduce the standard deviation of the peer variable to explicitly control for class diversity because average peer effects can reflect either homogeneous or heterogeneous groups of pupils. I have then:

$$\begin{aligned} Y_{icks} &= \alpha_0 + \alpha_1 \bar{PB}_{(-i)cks} + \alpha_2 PB_{icks} + \alpha_3 X_{icks} + \alpha_4 C_{cks} \\ &+ \alpha_5 (\bar{PB}_{(-i)cks} \cdot PB_{icks}) + \alpha_6 \tilde{PB}_{(-i)cks} + \mu_k + \nu_{icks} \end{aligned} \tag{2.3}$$

where $\tilde{PB}_{(-i)cks}$ represents the standard deviation of the peer variable, i.e., heterogeneity in the peer reference group. This specification, however, only reports the effect of the class compositional variables on the average student whereas the most important question is to find for which kind of students the peer group matters.

The second approach considers the same set of covariates and fixed components but within a quantile regression framework which analyzes peer effects for different subgroups of pupils, hierarchically structured by school performances. The quantile regression method has several advantages such as the reduced weight attached to outliers, the robustness to potential heteroscedasticity and the semi-parametric form of the model. I obtain the following specification:

$$\begin{aligned} Q_\theta(Y_{icks}) &= \alpha_{\theta 0} + \alpha_{\theta 1} \bar{PB}_{(-i)cks} + \alpha_{\theta 2} PB_{icks} + \alpha_{\theta 3} X_{icks} + \alpha_{\theta 4} C_{cks} \\ &+ \alpha_{\theta 5} (\bar{PB}_{(-i)cks} \cdot PB_{icks}) + \alpha_{\theta 6} \tilde{PB}_{(-i)cks} + \mu_k + \nu_{icks} \end{aligned} \tag{2.4}$$

where θ represents the θ^{th} quantile of the considered variables.

2.4.3 Equity analysis

To account for equality of opportunity, I consider a model inspired by the contributions of Schuetz et al. (2008) and Raitano and Vona (2011) which analyzes the relationship between student's test scores, parental background and sorting policies. As a starting point, assuming that observed school selection variables (e.g., student's prior records or teacher recommendations) can perfectly predict *ex-ante* the peer group formation is unrealistic. Other considerations (residential segregation or idiosyncratic preferences) can also constrain students' choices. For that purpose, scholars consider that accounting for peer variables, which are an *ex-post* measure of peer group formation, can provide reliable information on how class composition and parental background interact with each other.

The empirical model regresses the individual test's score on parental background, peer heterogeneity, school admittance procedure, school track design, individual background characteristics and a set of interaction terms between parental background and the variables of interest. I deliberately do not control for ability track fixed effects on the grounds that ability track enrollment is assumed to be strongly correlated with parental background and therefore may falsify the magnitude of the parental background gradient. I have then:

$$\begin{aligned} Y_{icks} &= \gamma_0 + \gamma_1 PB_{icks} + \gamma_2 (PB_{icks} \cdot \tilde{PB}_{(-i)cks}) + \gamma_3 (PB_{icks} \cdot SD_s) \\ &+ \gamma_4 (PB_{icks} \cdot SP_s) + \gamma_5 \tilde{PB}_{(-i)cks} + \gamma_6 X_{icks} + \gamma_7 SD_s + \gamma_8 SP_s \\ &+ \vartheta_{icks} \end{aligned} \quad (2.5)$$

The objective is to determine on the basis of the interaction variables mentioned in equation (2.5) if peer heterogeneity, the school admittance procedure and the school track design reinforce or weaken the impact of parental background on student's performances. The main focus is on the interaction term between peer heterogeneity and parental background to know if the former reinforce or weaken the effect of the latter on school performances. As before, I also consider a quantile regression approach to detect the potential differences along the test score distribution.

2.5 Results

2.5.1 Mean peer effects

Results from the baseline OLS regressions (ref. equation (2.1)) are presented in Table 2.3. The estimations report positive, strong, and significant average peer effects in all fields. Coefficients related to parental background, parental expectation, and language at home follow the expected signs. The value attached by parents to the field of interest influences positively and significantly the score obtained by their offspring in reading and sciences whereas the inverse holds for mathematics. A potential explanation may reside in the fact that parental interest for literacy or environmental issues can be more easily transmitted to children than their interest for mathematics. My results show that natives obtain higher test scores than pupils with migration backgrounds and that males perform better in mathematics and sciences, whereas females obtain better results in reading. As all students are in the ninth grade, the negative impact of age on school performances may be explained by the fact that older students generally reflect repeaters. Concerning the school track design, an integrative system seems to reduce reading performances while having no significant effect on mathematics performances[10]. In sciences, students who are grouped in a comprehensive way perform better than pupils who are enrolled in low ability tracks but worse than pupils from middle and high ability tracks. Overall, pupils in higher-ability tracks obtain better results in all fields considered. The coefficients related to class size are negative, small, and significant in each field except in mathematics where the coefficient is not significant. Finally, I notice that PISA test scores are higher in schools whose enrollment process is based on prior student ability. However, OLS estimation may be problematic regarding endogeneity biases and therefore coefficients need to be interpreted cautiously.

Results from fixed effects regressions (ref. equation (2.2)) are presented in Table 2.4. Compared to the OLS regressions, the introduction of ability track fixed effects reduces significantly the magnitude of peer effects in the three fields considered. Moreover, the peer effect coefficient in mathematics is no longer significant. These results reflect the existence of strong selection effects in the peer group composition. By interpreting my peer estimates in terms of standard deviation, I obtain that a one-standard-deviation increase in peer quality produces an significant increase of 0.042 and 0.035 of a standard deviation in reading and sciences test scores, respectively. Concerning the other control variables, minor differences exist between OLS and fixed effects regressions. Only two coefficients (related to parental value in mathematics and class size) lose their significance.

2.5.2 Efficiency

The policy relevance of mean peer effects is limited because it is crucial for policy makers to know which subgroup of pupils is most affected by peer effects and what is the potential impact of heterogeneous classes on educational outcomes. The first strategy consists in including two additional peer variables in the regression model. First, I introduce the interaction between the peers' parental background and the own parental background. A negative (positive) coefficient for the interaction term would indicate that pupils with low (high) parental background are more sensitive to the peer group's

[10]The variable *heterogeneous system* is not included in the regression model for sciences because it is perfectly collinear with the variable *Ability track* given that mixed-ability classes correspond by definition to the integrated system.

Table 2.3: Mean peer effets, OLS estimation

Dependent variable	**Reading test score**	**Mathematics test score**	**Sciences test score**
Explanatory variables	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**
Peer characteristics			
Peer quality	21.338***	14.623***	17.533***
	(3.549)	(3.592)	(3.569)
Parental background			
Parental background	6.654***	6.629***	8.124***
	(0.766)	(0.727)	(0.760)
Background characteristics			
Natives (reference category)			
Immigrant	-16.090***	-21.199***	-26.635***
	(2.186)	(2.096)	(2.137)
Immigrant parents	-22.207***	-25.454***	-31.578***
	(2.328)	(2.496)	(2.370)
Age	-5.119***	-9.585***	-2.903***
	(1.147)	(1.164)	(1.139)
Female	15.083***	-27.039***	-20.329***
	(1.185)	(1.238)	(1.207)
Other language at home	-16.144***	-12.471***	-18.667***
	(2.204)	(2.148)	(2.148)
Parental expectation	7.947***	8.588***	12.524***
	(1.686)	(1.707)	(1.744)
Parental value	4.898**	-5.718**	17.869***
	(2.200)	(2.207)	(1.323)
Ability tracks characteristics			
Ability track: mixed			(ref. cat)
Ability track: low	(ref. cat)	(ref. cat)	-39.421***
			(3.974)
Ability track: middle	46.642***	52.148***	16.035***
	(3.278)	(3.119)	(3.586)
Ability track: high	73.887***	88.324***	50.007***
	(4.134)	(4.162)	(4.153)
Class size			
Class size	-0.203*	-0.145	-0.381***
	(0.121)	(0.115)	(0.137)
School track design			
heterogeneous system	-7.428**	-5.414	
	(3.622)	(3.301)	
School characteristics			
Teacher quality	8.206***	16.026***	9.365***
	(2.576)	(2.685)	(2.432)
School size: low (reference category)			
School size: middle	-10.017***	-12.785***	-14.728***
	(2.713)	(2.618)	(2.588)
School size: high	-5.231	-5.947	-9.734**
	(4.136)	(3.930)	(3.863)
School location: village (reference category)			
School location: small town	7.997**	10.236***	2.205
	(3.237)	(3.420)	(3.223)
School location: town	7.958**	7.788**	2.404
	(3.740)	(3.861)	(3.680)
School location: city	-0.140	-22.124***	-12.436**
	(5.150)	(5.540)	(5.102)

School sorting policies			
Admission procedure	16.457***	14.488***	19.432***
	(2.302)	(2.399)	(2.181)
Constant	524.419***	649.140***	554.680***
	(18.911)	(18.962)	(17.893)
R-squared	0.356	0.398	0.414
N	14,081	14,081	14,081

Standard errors clustered at the class level.
*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.4: Mean peer effects, ability track FE estimation

Dependent variable	**Reading test score**	**Mathematics test score**	**Sciences test score**
Explanatory variables	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**
Peer characteristics			
Peer quality	8.020***	3.476	6.994**
	(3.030)	(2.888)	(3.084)
Parental background			
Parental background	5.486***	5.782***	7.461***
	(0.776)	(0.719)	(0.754)
Background characteristics			
Natives (reference category)			
Immigrant	-12.045***	-15.408***	-20.110***
	(1.941)	(1.854)	(1.923)
Immigrant parents	-18.553***	-20.880***	-24.948***
	(2.097)	(2.178)	(2.171)
Age	-10.939***	-18.074***	-10.987***
	(1.011)	(0.999)	(1.006)
Female	13.315***	-28.906***	-21.858***
	(1.086)	(1.112)	(1.111)
Other language at home	-15.739***	-10.890***	-17.777***
	(2.141)	(2.006)	(2.108)
Parental expectation	11.966***	12.649***	15.608***
	(1.589)	(1.531)	(1.569)
Parental value	8.884***	-2.709	13.556***
	(2.074)	(2.133)	(1.168)
Class size			
Class size	-0.032	-0.038	-0.094
	(0.111)	(0.120)	(0.127)
Constant	729.414***	905.462***	763.731***
	(18.598)	(23.999)	(21.635)
Ability track FE	Yes	Yes	Yes
R-squared	0.462	0.507	0.509
N	14,081	14,081	14,081

Standard errors clustered at the class level.
*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

influence. Second, I consider the standard deviation of the peer variable. Both variables can give valuable information regarding the potential gains in efficiency that one could obtain by adopting mixed-ability classes.

All interaction terms between peer effects and parental background reported in Table 2.5 are non-significant. More interestingly, an increase in class heterogeneity does not decrease significantly the school performances of the average student, whatever the field considered. These findings, however, do not allow to consider a comprehensive system as efficiency-enhancing on the grounds that Pareto conditions are not fully satisfied, i.e., there is no diminishing returns in peer effects according to family background.

Table 2.5: Nonlinear peer effects and peer heterogeneity, ability track FE estimation

Dependent variable	**Reading test score**	**Mathematics test score**	**Sciences test score**
Explanatory variables	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**
Peer characteristics			
Peer quality	7.803***	3.288	6.692**
	(3.023)	(2.887)	(3.083)
Peer quality*PB	-0.264	0.620	1.474
	(1.597)	(1.523)	(1.622)
Peer heterogeneity	-2.841	-1.157	-0.664
	(3.700)	(3.645)	(3.790)
Parental background			
Parental background	5.315***	5.581***	7.148***
	(0.860)	(0.808)	(0.848)
Background characteristics			
Natives (reference category)			
Immigrant	-12.024***	-15.412***	-20.121***
	(1.943)	(1.855)	(1.923)
Immigrant parents	-18.500***	-20.941***	-25.097***
	(2.108)	(2.192)	(2.178)
Age	-10.937***	-18.064***	-10.967***
	(1.011)	(1.000)	(1.007)
Female	13.310***	-28.912***	-21.864***
	(1.086)	(1.112)	(1.111)
Other language at home	-15.709***	-10.887***	-17.795***
	(2.089)	(2.004)	(2.105)
Parental expectation	12.005***	12.620***	15.519***
	(1.591)	(1.530)	(1.574)
Parental value	8.888***	-2.698	13.542***
	(2.072)	(2.134)	(1.167)
Class size			
Class size	-0.030	-0.039	-0.093
	(0.111)	(0.120)	(0.127)
Constant	732.336***	906.255***	763.616***
	(18.933)	(24.353)	(21.960)
Ability track FE	Yes	Yes	Yes
R-squared	0.462	0.507	0.509
N	14,081	14,081	14,081

Standard errors clustered at the class level.
*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

The second strategy (ref. equation (2.4)) gives the possibility to focus on the evolution of peer effects along the conditional ability distribution by using quantile regressions. Empirical findings from Table 2.6[11] report very interesting information concerning non-linearities in peer effects. For reading, we see a clear decrease in peer coefficients along the distribution. In sciences, we can see that only pupils in the two first quantiles benefit significantly from a higher peer average. On the contrary, high-achieving students in mathematics perform better when they are surrounded by similar peers given that peer effects estimates are only significant in the two last quantiles of the distribution. In all fields, class heterogeneity does not decrease significantly student's own achievement. These results argue in favor of adopting comprehensive classes in reading and sciences because the conditions to be Pareto-improving are now met: low-ability students turn out to be most affected by a higher peer average without affecting the overall performance of the classroom. For mathematics, however, maintaining ability tracking turns out to be the best way to ensure school efficiency because only pupils at the top of the conditional ability distribution obtain positive and significant peer effects estimates.

Quantile regressions estimate peers effects along the conditional ability distribution. This means for instance that the identified "weak" students are weak conditional on their background characteristics but not necessarily in absolute terms. In this context, a relevant analysis consists in analyzing if all results obtained from the quantile regressions can also be valid for the unconditional distribution. For that purpose, Figure 2.1 plots the residuals from the FE model against absolute test scores. The graphical analysis shows that conditional and unconditional test scores are closely related, which allows us to interpreted our estimates in a more robust way to draw policy recommendations.

2.5.3 Equity

To determine if class heterogeneity reinforces or weakens the impact of parental background on student's performances, I focus on the interaction term between peer heterogeneity and parental background (ref. equation (2.5)). A negative (positive) coefficient would indicate that class heterogeneity reduces (increases) the impact of parental background on schooling performances.

Results from Table 2.7 show that class heterogeneity reduces the family background effect whatever the field considered. Moreover, we can see that a school admission procedure based on prior school performances reinforces the

[11] Detailed results are presented in Table 2.10 to 2.12 in the Appendix.

Table 2.6: Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Reading					
Peer characteristics					
Peer quality	10.777*	9.163***	8.412**	6.826**	1.560
	(6.006)	(3.382)	(3.472)	(2.995)	(3.606)
Peer quality*PB	0.199	-1.505	-0.134	0.909	-1.180
	(2.018)	(2.001)	(2.258)	(1.509)	(1.925)
Peer heterogeneity	-4.852	-5.194	-5.050	1.686	7.106
	(4.535)	(3.252)	(5.809)	(4.502)	(5.304)
Pseudo r-squared	0.303	0.295	0.274	0.257	0.252
Mathematics					
Peer characteristics					
Peer quality	0.791	1.458	3.997	7.401***	5.981*
	(4.302)	(3.390)	(3.424)	(2.859)	(3.911)
Peer quality*PB	-1.399	-0.871	1.008	0.158	1.553
	(3.020)	(2.158)	(1.980)	(1.588)	(1.521)
Peer heterogeneity	-2.956	0.546	-0.676	-5.615	-1.973
	(6.223)	(3.928)	(3.682)	(3.936)	(4.524)
Pseudo r-squared	0.321	0.317	0.310	0.301	0.296
Sciences					
Peer characteristics					
Peer quality	8.341***	9.638***	6.327	3.674	-1.019
	(2.394)	(2.688)	(4.080)	(3.540)	(3.646)
Peer quality*PB	0.143	0.335	0.757	0.892	1.367
	(1.860)	(1.747)	(1.960)	(1.611)	(2.339)
Peer heterogeneity	-2.113	-6.186	-3.027	0.510	-0.702
	(6.712)	(6.555)	(5.201)	(6.355)	(7.161)
Pseudo r-squared	0.327	0.325	0.311	0.295	0.289
Ability track FE	Yes	Yes	Yes	Yes	Yes
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively. The model also controls for migration status, age, gender, language at home, parental expectation, parental value for the field considered and class size.

parental background effect and that the integrated system decreases it. To summarize, these findings speak in favor of adopting comprehensive classes to improve equality of opportunity at the lower secondary level, irrespective of the field analyzed. It is worth to point out, however, that such conclusions cannot be definitively interpreted as causal because I do not control for unobserved characteristics in this specification.

As for efficiency, I estimate quantile regressions whose results are presented in Table 2.8[12]. First, interaction terms between peer heterogeneity

[12]Detailed results are presented in Tables 2.13 to 2.15 in the Appendix.

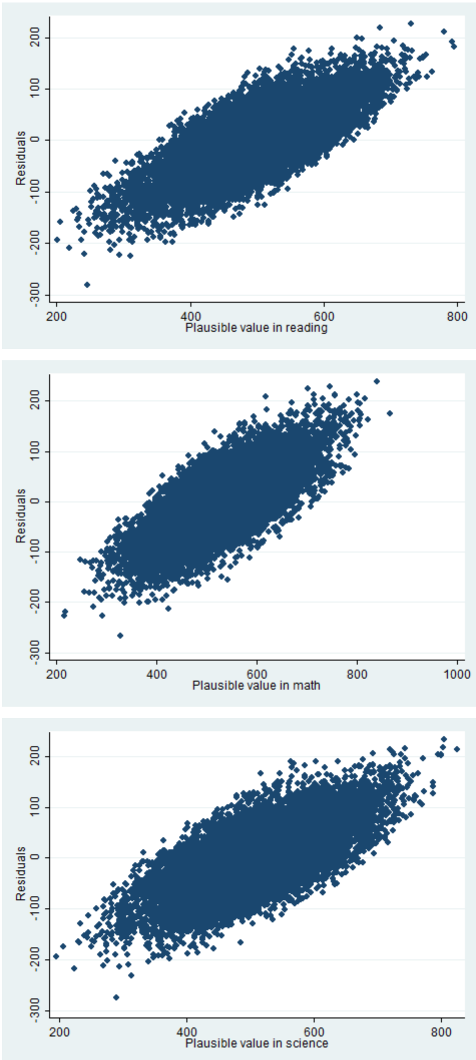


Figure 2.1: Unconditional versus conditional test scores - Efficiency analysis

Table 2.7: Peer heterogeneity and educational opportunities, OLS estimation

Dependent variable	**Reading test score**	**Mathematics test score**	**Sciences test score**
Explanatory variables	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**	**Coefficients (Robust s.e.)**
Parental background			
Parental background	23.860***	28.652***	27.257***
	(3.614)	(3.787)	(3.966)
Interactions variables			
Peer heterogeneity*PB	-8.292**	-13.852***	-10.752**
	(3.875)	(4.095)	(4.371)
heterogeneous system*PB	-10.001***	-6.917***	-10.691***
	(2.408)	(2.478)	(2.344)
Admission procedure*PB	5.977***	5.823***	6.025***
	(1.916)	(2.074)	(1.925)
Main effects			
Peer heterogeneity	-6.529	-7.429	-6.049
	(6.917)	(7.602)	(7.648)
heterogeneous system	-12.571**	-10.412***	-10.291***
	(3.184)	(3.660)	(3.304)
Admission procedure	17.517***	20.233***	23.261***
	(2.778)	(2.984)	(2.758)
Background characteristics			
Natives (reference category)			
Immigrant	-20.763***	-28.675***	-32.475***
	(2.312)	(2.335)	(2.351)
Immigrant parents	-31.707***	-37.581***	-42.036***
	(2.619)	(2.788)	(2.722)
Age	-12.450***	-16.771***	-10.753***
	(1.302)	(1.317)	(1.277)
Female	19.120***	-23.219***	-16.181***
	(1.348)	(1.427)	(1.364)
Other language	-18.396***	-14.649***	-20.959***
	(2.452)	(2.457)	(2.448)
Parental expectation	25.261***	27.378***	27.437***
	(1.820)	(1.875)	(1.834)
Parental value	2.781	-6.693**	24.536***
	(2.472)	(2.554)	(1.531)
Constant	693.936***	825.570***	684.100***
	(20.785)	(21.478)	(20.818)
R-squared	0.207	0.229	0.269
N	14,081	14,081	14,081

Standard errors clustered at the class level.
*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

and parental background reveal that mixing pupils with different parental economic, social, and cultural status reduces the parental background effect on schooling performances along the entire ability distribution in reading and mathematics whereas this reduction is only significant at the top of the distribution in sciences. The integrative system reduces the parental background effect in reading and sciences whereas results for mathematics are inconclusive. Finally, the variable related to the school admission procedure based on prior student's records is positively correlated with family background in

Table 2.8: Peer heterogeneity and educational opportunities, Quantile regressions

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Reading					
Parental background	25.350***	25.475***	24.588***	23.381***	20.055***
	(6.491)	(3.190)	(2.649)	(2.634)	(2.862)
Peer heterogeneity *PB	-7.717	-8.603***	-9.599**	-8.898**	-7.637*
	(6.209)	(2.896)	(4.054)	(4.169)	(4.639)
heterogeneous system*PB	-14.083***	-12.601***	-9.547***	-8.256***	-4.580***
	(3.357)	(2.290)	(1.142)	(0.968)	(1.570)
Admission procedure*PB	5.161**	6.270***	6.962***	5.184**	8.288**
	(2.461)	(2.058)	(1.988)	(2.502)	(4.212)
Pseudo r-squared	0.129	0.126	0.114	0.099	0.094
Mathematics					
Parental background	29.830***	29.850***	27.089***	28.609***	28.540***
	(6.998)	(5.455)	(2.291)	(2.278)	(7.426)
Peer heterogeneity*PB	-14.908*	-15.171**	-11.279***	-15.675***	-16.033**
	(8.661)	(6.673)	(2.863)	(3.464)	(7.895)
heterogeneous system*PB	-5.192	-6.088	-7.037***	-2.864	-2.967
	(4.878)	(4.173)	(1.858)	(3.993)	(3.079)
Admission procedure*PB	6.586***	6.879***	4.770***	7.224*	6.547**
	(0.510)	(1.148)	(0.407)	(3.697)	(3.053)
Pseudo r-squared	0.127	0.129	0.130	0.120	0.112
Sciences					
Parental background	24.138***	30.947***	28.281***	27.549***	30.079***
	(9.006)	(11.533)	(7.072)	(5.506)	(8.157)
Peer heterogeneity*PB	-6.502	-14.242	-11.484	-12.283***	-15.130***
	(11.078)	(13.978)	(9.201)	(5.832)	(7.479)
heterogeneous system*PB	-15.356***	-12.577***	-12.226***	-6.967	-7.579
	(5.735)	(3.664)	(3.114)	(4.709)	(4.630)
Admission procedure*PB	6.660***	5.837***	6.447***	6.796***	6.277***
	(1.517)	(2.157)	(1.728)	(1.646)	(1.611)
Pseudo r-squared	0.150	0.157	0.150	0.137	0.132
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.
The model also controls for migration status, age, gender, language at home, parental expectation and parental value for the field considered.

the different fields considered. On the basis of these findings, we see that most of variables analyzed here have a homogeneous impact on student's performances, whatever the level of cognitive skills considered.

Figure 2.2 plots the residuals from the OLS model against the raw test scores and reports, as for efficiency, that both conditional and unconditional test scores are very similar. As a consequence, estimates from quantile regressions are more interpretable for drawing policy advices.

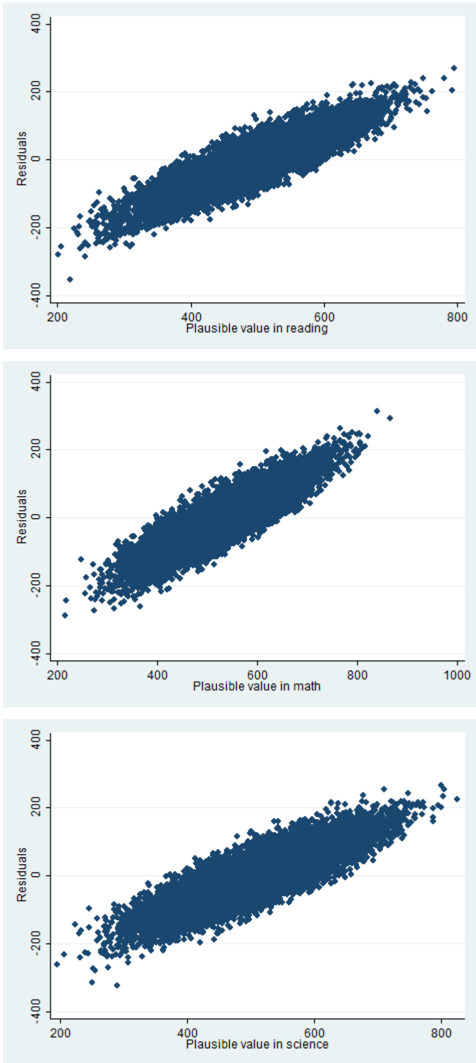


Figure 2.2: Unconditional versus conditional test scores - Equity analysis

2.6 Conclusion

Revisiting the organizational design at the Swiss lower secondary level is relevant on the grounds that ability tracking seems to be positively associated with social inequalities, does not manage to form homogeneous classes in terms of students' skills, and is not well-adapted to the current labour market needs. Moreover, school segregation prevents low-achieving students to glean positive peer effects from a regular contact with more advanced students. However, most parents have concerns that creating ability-mixed classes can affect the quality of instruction in the classroom and reduce the motivation of brighter students.

On the basis of peer effect theory that analyzes the magnitude and nature of social interactions between classmates, this study investigates which kind of effects in terms of efficiency and equity we could expect from introducing a completely non-selective system in the Swiss educational landscape. This research question is addressed by exploiting the relevant resources available in the PISA 2006 Swiss national sample which allows to estimate peer effects at the class level and control for both between- and within-school sorting. The peer quality variable is represented by the average parental socioeconomic background within the classroom and ability track fixed effects are introduced in the empirical model to reduce correlated effects. Non-linear peer effects are estimated with quantile regressions in order to analyze which kinds of pupils benefit most from peer effects along the conditional ability distribution.

OLS results show that peer effects are sizeable and that classmates' quality represents a strong predictor of student's performances. The magnitude of peer effects estimates, however, decreases when ability track fixed effects are included in the model but they remain positive and significant in reading and sciences. I account for non-linearities in peer effects by using a quantile regression framework which analyzes peer effects for different types of pupils ranked according to their schooling performances. In reading, results report positive, significant and decreasing peer coefficients over the conditional test score distribution. In sciences, only students at the bottom of the distribution benefit significantly from peer effects whereas it is the opposite for mathematics. Class diversity in terms of parental background has no adverse effects on school performances, whatever the field and quantiles considered. Finally, the specification measuring equality of opportunity indicates that peer heterogeneity reduces the impact of the family background on school performances in all fields considered.

Main findings suggest that adopting mixed-ability classes in reading and sciences could lead to Pareto-improving redistribution of students across classes and/or schools. In mathematics, however, maintaining ability tracking seems to be the best practice to ensure school efficiency but does not lead to any gains in terms of equity. Such a difference between mathematics and the other fields may be explained by the importance of prior knowledge acquired during primary school in mathematics. Bridging the gap in mathematics is then more challenging than in reading or sciences where peers effects play a more compensatory role.

Adopting a mixed-ability system except for mathematics, however, is hardly applicable. Indeed, the introduction of the integrated and cooperative systems in some Swiss cantons has been the result of a long trial period. Taking a step further seems complicated, especially after the recent postponement of the age of first tracking to 13 years old in most of Swiss cantons. However, this topic should be addressed in the future political agenda (e.g., through the development of some pilot experiments) given that equality of opportunity is a central discussion at the international level.

The main caveat of this study is that we cannot control for teacher fixed effects while teacher's observed and unobserved characteristics play an important role on students' achievement. A second caveat lies in the fact the empirical analysis does not account for the costs resulting from a change of school design. For instance, detracking would imply additional formation costs for teachers given that teaching in mixed ability classes necessitates additional skills. A third caveat is that my dataset does not offer the possibility to distinguish between endogenous and contextual peer effects, which explains why I rely on a reduced form model estimating a total educational peer effect.

2.7 References

Ammermueller A. and J.S. Pischke, Peer effects in European primary schools: evidence from the Progress in International Reading Literacy Study, *Journal of labour Economics*, 27(3), pp. 315-348, 2009.

Atkinson A., S. Burgess, P. Gregg, C. Propper and S. Proud, The impact of classroom peer groups on pupil GCSE results, CMPO Working Paper N° 08/187, 2008.

Bauer P. and R. Riphahn, Timing of School Tracking as a Determinant of Intergenerational Transmission of Education, *Economics Letters*, 91(1), pp. 90-97, 2006.

Betts J.R. and A. Zau, Peer groups and academic achievement: panel evidence from administrative data, unpublished document, Public Policy Instite of California, 2004.

Bonesronning H., Peer group effects in education production: is it about congestion?, *Journal of Socio-Economics*, Vol. 37, pp. 328-342, 2008.

Bramoullé Y., H. Djebbari and B. Fortin, Identification of peer effects through social networks, *Journal of Econometrics*, 150(1), pp. 41-55, 2009.

Brunello G. and D. Checchi, Does school tracking affect equality of opportunity? New international evidence, Economic Policy, CEPR and CES and MSH, Vol. 22, pp. 781-861, 2007.

Burke M. and T. Sass, Classroom peer effects and student achievement, Calder Working Paper N° 18, 2008.

Carman K. and L. Zhang, Classroom peer effects and academic achievement: Evidence from a Chinese middle school, unpublished document, 2008.

Coleman J., E. Campbell, C. Hobson, J. McPartlan, A. Mood, F. Weinfeld and R. York, *Equality of educational opportunity*, Government Printing Office, Washington D.C., 1966.

Centre suisse de coordination pour la recherche en éducation (CSRE), Swiss education report 2010, Aarau, 2010.

De Giorgi G., M. Pellizarri and S. Redaelli, Be as careful of the company you keep as of the books you read: peer effects in education and on the labour market, NBER Working Paper N° 14948, 2009.

De Paola M. and V. Scoppa, Peer group effects on the academic performance of Italian students, *Applied Economics*, 42(17), pp. 2203-2215, 2010.

Entorf H. and M. Lauk, Peer effects, social multipliers and migrants at school: an international comparison, unpublished document, 2006.

Fertig M., Educational production, endogenous peer group formation and class composition. Evidence from the PISA 2000 study, IZA Discussion Paper N° 714, 2003.

Figlio D. and M. Page, School choice and the distributional effects of ability tracking: does separation increase inequality?, *Journal of Urban Economics*, Vol. 51, pp. 497-514, 2002.

Fuentes A., Raising education outcomes in Switzerland, OECD Economics Department Working Papers N° 838, 2011.

Gaviria A. and S. Raphael, School-based peer effects and juvenile behaviour, *The Review of Economics and Statistics*, Vol. 83, pp. 257-268, 2001.

Gibbons S. and S. Telhaj, Peer and achievement in England's secondary schools, SERC Discussion Paper 1, 2008.

Hanushek E., J. Kain, J. Markman and S. Rivkin, Does peer ability affect student achievement?, *Journal of Applied Econometrics*, 18(5), pp. 527- 544, 2003.

Hidalgo-Hidalgo M., Tracking can be more Equitable than Mixing: Peer effects and college attendance, Working Paper 09.04, Universidad Pablo de Olavide Sevilla, 2009.

Hoxby C., Peer effects in the classroom: learning from gender and race variation, NBER Working Paper N° 7867, 2000.

Hoxby C. and G. Weingarth, Taking race out of the equation: school reassignment and the structure of peer effects, unpublished manuscript, 2005.

Lavy V., M.D. Paserman and A. Schlosser, Inside the black box of ability peer effects: evidence from variation in the proportion of low achievers in the classroom, NBER Working Paper, 2009.

Lazear E., Education production, *Quarterly Journal of Economics*, 116(3), pp. 777-803, 2001.

Lefgren L., Educational peer effects and the Chicago public schools, *Journal of Urban Economics*, 56(2), pp. 169-191, 2004.

Manski C., Identification of endogenous social effects: the reflection problem, *Review of Economic Studies*, 60 (3), pp. 531-542, 1993.

Manski C., *Identification problems in the social sciencess*, Harvard University Press, Massachusetts, 1995.

Manski C., Economic analysis of social interactions, *Journal of Economic Perspectives*, Vol. 14, pp. 115-136, 2000.

McEwan P.J., Peer effects on student achievement: evidence from Chile, *Economics of Education Review*, Vol. 22, pp. 131-141, 2003.

Raitano M. and F. Vona, Assessing students' equality of opportunity in OECD countries: the role of national and school-level policies, OFCE Working Paper, 2011.

Rangvid B., Educational peer effects. Quantile regression: evidence from Denmark with PISA 2000 data, Do schools matter?, PhD thesis, Aarhus School of Business (Denmark), 2004.

Rees D., D. Brewer and L. Argys, How should we measure the effect of ability grouping on student performance?, *Economics of Education Review*, Vol. 19, pp. 17-20, 2000.

Sacerdote B.I., Peer effects with random assignment: Results for Dartmouth Roommates, NBER Working Paper N° 7469, January 2001.

Sacerdote B.I., Peer effects in education: How might they work, how big are they and how much do we know thus far?, in E. Hanushek, S. Machin and L. Woessmann (eds.), Vol. 3, *Handbook of the Economics of Education*, Chapter 4, 2011.

Schneeweis N. and R. Winter-Ebmer, Peer effects in Austrian schools, *Empirical Economics*, Vol. 32, pp. 387-409, 2007.

Schuetz G., H. Ursprung and L. Woessman, Education policy and equality of opportunity, *Kyklos*, 61(2), pp. 279-308, 2008.

Sund K., Estimating peer effects in Swedish high schools using school, teacher, and student fixed effects, Swedish Institute for Social Research, Working Paper N°8, 2007.

Vandenberghe V., Evaluating the magnitude and the stakes of peer effects analyzing sciences and math achievement across OECD, *Applied Economics*, 34(10), pp. 1283-1290, 2002.

Vigdor J. and T. Nechyba, Peer effects in North Carolina public schools, in L. Woessman and Paul P. Peterson (eds.), *Schools and the equal opportunity problem*, CESifo Seminar Series, MIT Press Books, 2007.

Zabel J., The impact of peer effects on student outcomes in New York City Public Schools, *Education Finance and Policy*, 3(2), p. 197-249, 2008.

2.8 Appendices

2.8.1 School track designs

Table 2.9: School track designs

School track design	**Grouping procedure**	**Definition**	**Ability tracks**	**Peer reference group**
Separated system	Homogeneous	Ability grouping (ability tracking *at the school level*)	High-, middle- and low-ability tracks	Class level
Cooperative system	Homogeneous	Ability grouping (ability tracking *at the class level*)	High-, middle- and low-ability tracks	Class level
Integrated system	Heterogeneous	Level grouping (ability tracking *within the class*) for reading and mathematics and mixed-ability classes for sciences	Differentiated-level courses (high, middle or low)	Class level (sciences) and within the class (reading and mathematics)

2.8.2 Additional tables

Table 2.10: Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE, Reading

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Peer characteristics					
Peer quality	10.777*	9.163***	8.412**	6.826**	1.560
	(6.006)	(3.382)	(3.472)	(2.995)	(3.606)
Peer quality*PB	0.199	-1.505	-0.134	0.909	-1.180
	(2.018)	(2.001)	(2.258)	(1.509)	(1.925)
Peer heterogeneity	-4.852	-5.194	-5.050	1.686	7.106
	(4.535)	(3.252)	(5.809)	(4.502)	(5.304)
Parental background					
Parental background	4.470***	4.919***	4.630***	4.479***	6.006***
	(1.061)	(0.805)	(0.745)	(0.571)	(1.165)
Background characteristics					
Natives (reference category)					
Immigrant	-10.673**	-10.175***	-12.950***	-14.546***	-15.533***
	(4.321)	(2.944)	(2.210)	(2.925)	(3.572)
Immigrant parents	-17.689***	-20.218***	-21.902***	-19.077***	-17.015***
	(3.969)	(2.768)	(3.339)	(4.069)	(4.639)
Age	-12.537***	-12.195***	-11.091***	-8.017***	-9.129***
	(1.328)	(1.359)	(0.868)	(0.921)	(0.858)
Female	15.220***	14.229***	11.758***	11.093***	11.474***
	(2.029)	(1.611)	(1.101)	(1.620)	(1.263)
Other language at home	-18.055***	-16.624***	-14.270***	-16.875***	-12.658***
	(3.490)	(3.604)	(3.789)	(3.029)	(3.617)
Parental expectation	12.154***	12.183***	11.653***	12.935***	13.474***
	(2.793)	(2.371)	(1.850)	(1.496)	(2.847)
Parental value	13.572***	10.028***	8.806**	7.958***	5.007*
	(3.773)	(3.877)	(3.441)	(2.765)	(2.666)
Class size					
Class size	-0.093	-0.20	0.001	0.108	0.170
	(0.168)	(0.198)	(0.165)	(0.181)	(0.210)
Constant	674.985***	705.449***	732.642***	729.148***	798.108***
	(36.081)	(28.693)	(20.812)	(22.674)	(25.933)
Ability track FE	Yes	Yes	Yes	Yes	Yes
Pseudo r-squared	0.303	0.295	0.274	0.257	0.252
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.11: Nonlinear peer effects and peer heterogeneity, Quantile regressions with ability track FE, Mathematics

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Peer characteristics					
Peer quality	0.791	1.458	3.997	7.401***	5.981*
	(4.302)	(3.390)	(3.424)	(2.859)	(3.911)
Peer quality*PB	-1.399	-0.871	1.008	0.158	1.553
	(3.020)	(2.158)	(1.980)	(1.588)	(1.521)
Peer heterogeneity	-2.956	0.546	-0.676	-5.615	-1.973
	(6.223)	(3.928)	(3.682)	(3.936)	(4.524)
Parental background					
Parental background	4.717***	5.686***	5.024***	5.843***	4.080***
	(1.044)	(0.588)	(0.915)	(0.614)	(1.139)
Background characteristics					
Natives (reference category)					
Immigrant	-12.656***	-14.436***	-13.787***	-15.493***	-18.518***
	(2.035)	(2.501)	(1.321)	(2.263)	(2.810)
Immigrant parents	-21.083***	-21.624***	-17.370***	-18.604***	-20.703***
	(3.816)	(3.886)	(3.693)	(4.047)	(4.818)
Age	-19.929***	-17.838***	-17.671***	-17.760***	-16.514***
	(1.262)	(1.244)	(1.558)	(1.426)	(1.535)
Female	-27.176***	-28.465***	-30.649***	-30.247***	-29.762***
	(2.269)	(2.347)	(1.616)	(1.500)	(1.459)
Other language at home	-11.267***	-10.716***	-13.592***	-12.249***	-9.172***
	(2.680)	(2.073)	(3.255)	(2.786)	(3.131)
Parental expectation	12.488***	12.414***	11.676***	12.660***	13.488***
	(2.419)	(2.437)	(2.148)	(2.034)	(1.850)
Parental value	-2.115	-3.668	-4.221	-2.587	-3.131
	(2.396)	(2.399)	(2.781)	(2.050)	(2.264)
Class size					
Class size	-0.485***	-0.199	0.012	0.377*	0.376
	(0.162)	(0.146)	(0.227)	(0.199)	(0.145)
Constant	866.860***	856.034***	895.550***	943.242***	944.788***
	(19.581)	(38.048)	(26.071)	(18.738)	(26.300)
Ability track FE	Yes	Yes	Yes	Yes	Yes
Pseudo r-squared	0.321	0.317	0.310	0.301	0.296
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.12: Nonlinear peer effects and peer heterogeneity, Quantile regression with ability track FE, Sciences

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Peer characteristics					
Peer quality	8.341***	9.638***	6.327	3.674	-1.019
	(2.394)	(2.688)	(4.080)	(3.540)	(3.646)
Peer quality*PB	0.143	0.335	0.757	0.892	1.367
	(1.860)	(1.747)	(1.960)	(1.611)	(2.339)
Peer heterogeneity	-2.113	-6.186	-3.027	0.510	-0.702
	(6.712)	(6.555)	(5.201)	(6.355)	(7.161)
Parental background					
Parental background	7.438***	6.838***	6.873***	7.518***	7.026***
	(1.178)	(0.969)	(0.717)	(1.080)	(0.780)
Background characteristics					
Natives (reference category)					
Immigrant	-18.858***	-18.943***	-20.645***	-19.577***	-21.257***
	(1.596)	(1.895)	(1.859)	(2.380)	(1.763)
Immigrant parents	-27.958***	-26.910***	-23.765***	-25.031***	-24.595***
	(3.823)	(3.458)	(2.728)	(2.567)	(3.150)
Age	-12.280***	-11.046***	-10.442***	-10.156***	-9.307***
	(1.381)	(1.197)	(1.560)	(1.475)	(1.391)
Female	-18.729***	-20.081***	-22.796***	-23.724***	-24.726***
	(1.108)	(1.278)	(1.279)	(1.590)	(1.445)
Other language at home	-18.210***	-19.083***	-16.858***	-17.957***	-15.973***
	(3.961)	(3.066)	(3.236)	(1.763)	(2.500)
Parental expectation	16.064***	15.732***	14.815***	16.143***	15.119***
	(2.687)	(2.002)	(1.068)	(1.691)	(2.062)
Parental value	12.238***	11.844***	12.952***	14.412***	16.677***
	(1.496)	(0.868)	(0.908)	(1.460)	(1.768)
Class size					
Class size	-0.110	-0.120	-0.133	-0.126	-0.082
	(0.166)	(0.107)	(0.171)	(0.107)	(0.127)
Constant	729.333***	736.066***	750.190***	770.483***	805.358***
	(29.197)	(18.585)	(26.348)	(36.571)	(59.203)
Ability track FE	Yes	Yes	Yes	Yes	Yes
Pseudo r-squared	0.327	0.325	0.311	0.295	0.289
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.13: Peer heterogeneity and educational opportunities, Quantile regression, Reading

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Parental background					
Parental background	25.350***	25.475***	24.588***	23.381***	20.055***
	(6.491)	(3.190)	(2.649)	(2.634)	(2.862)
Interactions variables					
Peer heterogeneity *PB	-7.717	-8.603***	-9.599**	-8.898**	-7.637*
	(6.209)	(2.896)	(4.054)	(4.169)	(4.639)
heterogeneous system*PB	-14.083***	-12.601***	-9.547***	-8.256***	-4.580***
	(3.357)	(2.290)	(1.142)	(0.968)	(1.570)
Admission procedure*PB	5.161**	6.270***	6.962***	5.184**	8.288**
	(2.461)	(2.058)	(1.988)	(2.502)	(4.212)
Main effects					
Peer heterogeneity	-13.747***	-16.356***	-4.138	-2.415***	-0.050
	(1.954)	(3.864)	(2.937)	(0.630)	(2.537)
heterogeneous system	-5.917***	-10.318***	-14.499***	-16.240***	-16.269***
	(1.567)	(2.655)	(1.105)	(1.345)	(1.337)
Admission procedure	15.147***	15.177***	16.836***	19.930***	18.417***
	(2.081)	(1.639)	(0.709)	(0.633)	(2.075)
Background characteristics					
Natives (reference category)					
Immigrant	-15.255***	-19.416***	-23.180***	-25.245***	-26.119***
	(5.220)	(5.828)	(1.509)	(1.407)	(1.815)
Immigrant parents	-30.041***	-33.640***	-34.459***	-32.005***	-29.098***
	(2.982)	(3.051)	(2.755)	(1.994)	(1.322)
Age	-16.749***	-14.565***	-11.771***	-10.238***	-9.395***
	(1.786))	(1.383)	(1.394)	(1.198)	(1.156)
Female	23.342***	22.491***	16.641***	16.146***	15.593***
	(0.376)	(2.022)	(0.431)	(0.605)	(0.792)
Other language	-18.270***	-18.662***	-20.106***	-14.708***	-17.490***
	(0.971)	(1.938)	(3.190)	(2.172)	(2.030)
Parental expectation	28.978***	28.946***	25.176***	22.754***	21.768***
	(1.780)	(1.253)	(0.905)	(1.944)	(1.616)
Parental value	6.297	4.501	1.662	-0.393	-1.661
	(6.850)	(8.257)	(7.636)	(3.392)	(1.076)
Constant	685.365***	684.059***	687.696***	710.103***	722.697***
	(22.064)	(18.128)	(15.641)	(16.841)	(14.447)
Pseudo r-squared	0.129	0.126	0.114	0.099	0.094
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.14: Peer heterogeneity and educational opportunities, Quantile regression, Mathematics

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Parental background					
Parental background	29.830***	29.850***	27.089***	28.609***	28.540***
	(6.998)	(5.455)	(2.291)	(2.278)	(7.426)
Interactions variables					
Peer heterogeneity*PB	-14.908*	-15.171**	-11.279***	-15.675***	-16.033**
	(8.661)	(6.673)	(2.863)	(3.464)	(7.895)
heterogeneous system*PB	-5.192	-6.088	-7.037***	-2.864	-2.967
	(4.878)	(4.173)	(1.858)	(3.993)	(3.079)
Admission procedure*PB	6.586***	6.879***	4.770***	7.224*	6.547**
	(0.510)	(1.148)	(0.407)	(3.697)	(3.053)
Main effects					
Peer heterogeneity	-18.611**	-16.099**	-10.410***	2.451	6.077*
	(7.858)	(6.581)	(3.884)	(4.295)	(3.405)
heterogeneous system	-5.683	-7.510***	-10.379***	-14.116***	-13.634***
	(3.506)	(2.179)	(3.371)	(4.848)	(0.766)
Admission procedure	16.821***	18.442***	21.615***	21.142***	22.665***
	(1.095)	(1.610)	(1.624)	(1.348)	(1.625)
Background characteristics					
Immigrant	-23.778***	-23.910***	-32.242***	-31.182***	-32.472***
	(3.531)	(1.969)	(3.462)	(6.063)	(4.854)
Immigrant parents	-33.911***	-35.191***	-38.239***	-39.925***	-40.183***
	(5.273)	(3.287)	(4.152)	(3.987)	(5.845)
Age	-18.008***	-17.306***	-16.598***	-15.620***	-12.722***
	(0.339)	(1.717)	(0.501)	(0.576)	(0.911)
Female	-23.414***	-23.117***	-25.061***	-24.492***	-23.373***
	(1.885)	(3.356)	(1.591)	(1.177)	(1.539)
Other language	-12.740*	-13.221***	-16.195***	-14.997***	-15.475***
	(7.269)	(2.146)	(0.754)	(0.748)	(2.684)
Parental expectation	30.746***	30.790***	29.634***	27.928***	25.443***
	(4.229)	(0.565)	(3.282)	(0.377)	(1.056)
Parental value	-3.611	-5.004**	-5.999**	-11.335***	-5.714**
	(3.968)	(2.522)	(2.859)	(1.108)	(2.562)
Constant	771.022***	787.382***	826.349***	856.277***	828.686***
	(11.418)	(31.620)	(4.621)	(7.252)	(17.701)
Pseudo r-squared	0.127	0.129	0.130	0.120	0.112
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 2.15: Peer heterogeneity and educational opportunities, Quantile regression, Sciences

	Quantile				
	0.15	**0.25**	**0.50**	**0.75**	**0.85**
Parental background					
Parental background	24.138***	30.947***	28.281***	27.549***	30.079***
	(9.006)	(11.533)	(7.072)	(5.506)	(8.157)
Interactions variables					
Peer heterogeneity*PB	-6.502	-14.242	-11.484	-12.283***	-15.130***
	(11.078)	(13.978)	(9.201)	(5.832)	(7.479)
heterogeneous system*PB	-15.356***	-12.577***	-12.226***	-6.967	-7.579
	(5.735)	(3.664)	(3.114)	(4.709)	(4.630)
Admission procedure*PB	6.660***	5.837***	6.447***	6.796***	6.277***
	(1.517)	(2.157)	(1.728)	(1.646)	(1.611)
Main effects					
Peer heterogeneity	-13.965***	-10.870***	-9.749**	2.110	6.130
	(4.710)	(3.685)	(3.931)	(5.538)	(4.620)
heterogeneous system	-7.381***	-10.257***	-11.698***	-11.704***	-15.057***
	(1.750)	(1.060)	(3.916)	(0.960)	(2.647)
Admission procedure	21.622***	23.803***	22.911***	23.090***	24.159***
	(0.075)	(0.383)	(0.653)	(0.291)	(1.569)
Background characteristics					
Immigrant	-29.767***	-33.258***	-33.556***	-36.079***	-32.773***
	(3.240)	(6.368)	(3.327)	(3.992)	(2.073)
Immigrant parents	-44.527***	-46.246***	-44.047***	-39.832***	-39.851***
	(6.782)	(3.330)	(1.543)	(2.576)	(3.618)
Age	-14.122***	-11.852***	-9.751***	-9.050***	-8.837***
	(1.476)	(1.513)	(0.610)	(0.586)	(1.232)
Female	-11.610***	-14.172***	-17.624***	-18.283***	-19.886***
	(1.364)	(2.401)	(1.884)	(0.582)	(1.666)
Other language	-21.846***	-19.280***	-22.289***	-19.939***	-19.371***
	(3.018)	(2.626)	(4.096)	(4.706)	(4.908)
Parental expectation	31.511***	32.081***	29.712***	26.083***	25.501***
	(3.103)	(0.311)	(1.430)	(2.305)	(0.915)
Parental value	21.914***	23.989***	26.216***	25.524***	25.501***
	(1.090)	(0.984)	(2.634)	(1.018)	(0.915)
Constant	665.618***	653.217***	671.923***	702.664***	722.117***
	(26.579)	(28.252)	(7.043)	(10.388)	(20.410)
Pseudo r-squared	0.150	0.157	0.150	0.137	0.132
N=14,081					

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Chapter 3

Cannabis use and short-term academic performances†

In this paper we use longitudinal data on Swiss adolescents to investigate the impact of cannabis use on short-term educational performance. We focus our analysis on high school students and analyze various outcomes, including absenteeism, grades, and concentration. We exploit the panel nature of the data and control for a rich set of individual and family characteristics measured at the end of compulsory school. Results from both fixed effects regressions and propensity score matching indicate that high school students who smoke cannabis skip one additional half day of school per month and are 10-20% more likely to obtain poor grades. In addition, our empirical approaches highlight the importance of taking unobserved heterogeneity into account when assessing the impact of substance use on education.

† In collaboration with Joachim Marti (Leeds University)

3.1 Introduction

Data from the recent Addiction Monitoring Survey conducted in Switzerland reveal particularly high levels of cannabis consumption in the country, especially among adolescents and young adults. The prevalence of past 30 days use almost reaches 10% in the 15-24 age group, and nearly 50% of young adults aged 20-24 report having smoked cannabis at least once in their lifetime. Additionally, approximately one-fifth of adolescent and young adult consumers report smoking cannabis daily or almost daily (Gmel et al., 2012). The high levels of cannabis consumption observed in the country may be partly explained by a lack of understanding of its potential harmful consequences. A qualitative study conducted in various age groups in Switzerland shows that, overall, people do not have a very clear perception of the risks of cannabis use and of what constitutes misuse (Menghrajani et al., 2005). The authors call for a more developed prevention approach, including better provision of information about the risks of consumption. This is especially important considering the fairly permissive legislation of cannabis use in Switzerland that may send confusing signals about risks and social norms. Formally, the product is not legal but consequences are limited for consumers and the law is enforced with various degrees of severity in the country.

A consequence of cannabis use that has attracted increasing attention is its potential impact on cognitive abilities and, ultimately, on the accumulation of human capital among youths. The focus on adolescents is especially relevant given that prior studies reveal that this age group is particularly vulnerable to substance use (Monti et al., 2005; Solowij and Michie, 2007; Jager and Ramsey, 2008). Moreover, international evidence shows that adolescents are using cannabis at younger ages and that early initiation often corresponds to worse cognitive outcomes (Solowij and Battisti, 2008; Fontes et al., 2011; Gruber et al., 2012; Meier et al., 2012). Solowij and Battisti (2008) find that the harmful effects of cannabis use on neuropsychological functioning persist beyond the period of intoxication. In a study that assesses cognitive performance of marijuana users, Gruber et al. (2012) conclude that exposure to the product during adolescence affects brain development and finds that age of onset, frequency of consumption and level of consumption influence the strength of this relationship. Fontes et al. (2011) find similar results and emphasize the particularly detrimental effects of early onset. In a recent study using longitudinal data from New Zealand, Meier et al. (2012) compare the evolution of cognitive functioning before (age 13) and after (age 38) initiation of cannabis use and show worse deteriorations in outcomes among early and persistent users. In addition to these adverse consequences, cannabis consumption may lead to addiction effects in the case of regular and prolonged use (Stephens et al., 2002, Budney, 2006; Budney et al., 2006; Budney et al., 2007).

Many studies have found evidence of an association between consumption of the product and poor schooling outcomes (Bray et al., 2000; Yamada et al., 1997; Lynksey et al., 2003; MacLeod et al., 2004; Horwood et al., 2010) but only a few employ empirical strategies that address potential bias arising from reverse causality (e.g., psychological distress due to school difficulties may increase the perceived benefits of consumption) or unobserved heterogeneity (Register et al., 2001; Pacula et al., 2003; Roebuck et al., 2004; van Ours and Williams, 2009; Bessey and Backes-Gellner, 2009; McCaffrey et al., 2010). For instance, Register et al. (2001) analyze the impact of drug use on the number of years of education by using data from the National Longitudinal Study of Youth (NLSY). By incorporating the predicted probability of drug use into a regression framework for educational attainment (two-step estimation process), their results report that drug use has a significant negative impact on the number of years of schooling. Van Ours and Williams (2009) focus on the impact of age at initiation on dropout rates among Australian adolescents. Using bivariate duration models, they show that early onset of cannabis use has a detrimental impact on years of education completed and significantly increases the likelihood of school dropout. Bessey and Backes-Gellner (2009) analyze the impact of onset of cannabis consumption on educational outcomes and labor market success in Switzerland. The authors estimate a multivariate probit model coupled with an IV strategy and find that cannabis use exerts a negative impact on educational attainment. Results also reveal that cannabis consumption is positively associated with the probability of working less than 80%. McCaffrey et al. (2010) find evidence of an impact of heavy and persistent cannabis use on high school dropout using propensity score matching. While they argue that their results are probably driven by time-varying unobserved heterogeneity rather than by effects on cognitive abilities, the mechanisms remain unclear.

The main stream of this literature focuses on educational attainment outcomes, such as dropout rates or the number of years of education completed. Intermediate (short-term) outcomes are rarely considered, leaving underlying mechanisms poorly understood. Only scarce econometric evidence exists on the impact of cannabis use on outcomes such as grades, absenteeism or concentration. Notable exceptions are papers by Pacula et al. (2003), Roebuck et al. (2004), Engberg and Morral (2006) and Caldeira et al. (2008). Pacula et al. (2003) provides evidence of a negative impact of cannabis use on standardized test scores but shows that the estimated effects considerably shrink after accounting for unobserved heterogeneity. Roebuck et al. (2004) estimate the impact of cannabis use on truancy by using a zero-inflated negative binomial regression analysis. Their results show that cannabis users skip more school days than non-users. Engberg and Morral (2006) analyze if decreases in cannabis use improve adolescent school attendance. The authors consider a longitudinal study of US youths aged 12-19. The fixed effect regression models report that the elimination of cannabis use is associated with increased likelihood of school attendance. The study of Caldeira et al. (2008) reports the prevalence of cannabis use disorders and other cannabis-related problems in a large cohort of first-year US college students. Their results reveal that concentrations problems are among the most prevalent cannabis-related problems. Table 3.1 proposes a selective review of closely related papers.

In this paper, we build on this body of work and investigate the pathways through which cannabis use may affect short-term academic outcomes of adolescents in Switzerland, including concentration problems, learning difficulties, absenteeism and poor grades. We compare results from individual fixed effects models and propensity score matching and assess the sensitivity of the latter to potential unobserved heterogeneity using Rosenbaum bounds (Rosenbaum, 2002). We control for several usually unobserved personality traits such as persistency and self-esteem as well as for a rich set of family and individual characteristics measured at completion of compulsory education. We find consistent evidence that cannabis use reduces school attendance and increases the likelihood of poor educational performance among high school students. However, our empirical results highlight the importance of accounting for unobserved heterogeneity in the substance use-education relationship.

3.2 Data

Our data come from the Swiss Transition from Education to Employment (TREE) survey[1], which is nationally representative and longitudinal. TREE collects information on a series of education, work, and health-related variables, along with rich information on psychological traits. The baseline TREE sample consists of a subsample of 5,528 adolescents who responded to the OECD Program for International Student Assessment (PISA) questionnaire in 2000, which takes place at the end of compulsory schooling (i.e., at approximately 15 years old). We are able to match the TREE survey information with the PISA responses and therefore have access to a wide variety of background characteristics for each respondent. This baseline data includes information such as family characteristics (including educational support), intermediate school quality indicators and measures of cognitive ability (e.g.,

[1]The Swiss youth panel study TREE (Transitions from Education to Employment; tree.unibas.ch) has been running since 2000 and has since been funded by the Swiss National Science Foundation, the University of Basel, the Swiss Federal Office of Statistics, the Federal Office of Professional Education and Technology, and the cantons of Bern, Geneva and Ticino. Distribution: Dataservice, FORS, Lausanne. The dataset is available to all interested researchers and can be ordered at the Data Archive of the Swiss Foundation for Research in Social Sciences (FORS) in Lausanne, Switzerland.

Table 3.1: A selective review of closely related papers

Studies	Cannabis indicator	Education indicator	Data	Methodology	Findings
Short-term outcomes					
Pacula et al. (2003)	Light vs heavy use	Standardized test score	NELS	Difference-in-differences	15% reduction in performance on standardized mathematics tests
Roebuck et al. (2004)	Chronic vs non-chronic use	Absenteeism	NHSDA	ZINB	Marijuana users skipped 0.5-1 more days than non-users
Engberg and Morral (2006)	Users vs non-users	School attendance	PETS-A	RE and FE	Elimination of marijuana use is positively associated with school attendance
Caldeira et al. (2008)	Users vs non-users	Concentration problems	Interviews in a US public university	Logistic regressions	Concentration problems are among the most prevalent cannabis-related problems.
Long-term outcomes					
Yamada et al. (1996)	Users vs non-users	High school graduation	NLSY	Probit regressions	Significant adverse effects of marijuana use on high school graduation rates
Bray et al. (2000)	Users vs non-users	High school dropout	Longitudinal survey of students in a US public school	Logistic regressions	Positive relationship between marijuana initiation and dropout. Odds of dropping out is about 2.3 times that non-users
Register et al. (2001)	Users vs non-users	Years of education	NLSY	2SLS	Reduction of around 1 year in educational attainment
Lynksey et al. (2003)	Frequent vs infrequent users	High school completion	Cohort study in Victoria (Australia)	Logistic regressions	Increased risk of early school-leaving (2-5 times higher than infrequent users) but diminishing with age
van Ours and Williams (2009)	Onset of initiation	Years of education	ANDSHS	Bivariate duration models	Reduction in years of education depends on the age of onset. Larger impact for females.
Bessey and Backes-Gellner (2009)	Onset of initiation	Educational attainment and labor market success	Swiss Health Survey	Multivariate probit with an IV strategy	Negative impact on educational attainment and employment level
Mc Caffrey et al. (2010)	Users vs non-users	High school dropout	Alert Plus	PSM	Not significant
Horwood et al. (2010)	Onset of initiation	High school completion, university enrolment and degree attainment	CHDS, VAHCS, MUSP	Logistic regressions	Significant association between onset and all outcomes (contribution to 17% to failure)

NHSDA: National Household Survey on Drug Abuse. ZINB: Zero-inflated negative binomial. ANDSHS: Australian National Drug Strategy Household Survey. PETS-A: Persistent Effects of Treatment Study-Adolescent. RE and FE: Random- and fixed effects. CHDS: The Christchurch Health and Development Study. VAHCS: The Victoria Adolescent Health Cohort Study. MUSP: The Mater-University of Queensland Study of Pregnancy and Outcomes.

reading and math test scores).

In Switzerland, a majority of adolescents are involved in professional tracks after compulsory school (i.e., vocational school, apprenticeship). Because these educational programs differ widely in terms of academic content and study hours, we focus on high-school students, a more homogeneous population that follow a full-time education program. We therefore focus on adolescents enrolled in an academic matura school. Additionally, because our identification strategy relies on individual-level changes in cannabis use and because we use lagged cannabis use in our models, we restrict the sample to students that were observed for each year between 2001 and 2003, i.e., between waves 1 and 3 (N=1,416). It is worth noting that we could have included a fourth wave. However, the duration of studies varies across regional (cantonal) systems (from 3 to 4 years). Therefore, students in 2004 may be a significantly different selected group than students in their first three years of high school. We drop respondents with incomplete information on cannabis use and on other control variables and obtain an analysis sample of approximately 1,100 high school students. The exact size of each analysis sample depends on the outcome under investigation and on the specification used.

In our analysis, we focus on six short-term outcomes that measure different aspects of schooling. First, we consider *absenteeism*, which is defined as the number of days the student was absent from school during the previous month. Second, we create an index for *school difficulties* whose values range from 1 to 5. More precisely, the index is based on the aggregation of the five following questions (each question is represented with a dummy variable whose value is equal to one if the response is "often" and zero otherwise): "If I don't study during the weekend, I can hardly satisfy school requirements", "I have too much work at school", "I can hardly manage the amount of homework", "The subjects of the lessons change so fast, that I have trouble to keep up" and "At school I often feel out of my depth". Next, we create two binary indicators reflecting *lack of engagement* and *lack of motivation* that equal one if the respondent answers "no" to the following statements, respectively: "I work very concentrated at school" and "Usually I am fully present at school." Then, we create a binary outcome *Poor grades* which takes the value one if the student mentions having had at least one failing grade in her last grade report. Our last outcome is a binary indicator of recent *concentration problems* ("Over the last month, did you suffer from lack of concentration?"). Although we are estimating reduced form equations for each of these outcomes, it may be conceptually important to distinguish between "performance" outcomes and mechanisms. Our performance outcome is *Poor grades*, while all other outcomes pertain to the education production function itself: exposure to education (school days skipped), concentration, engagement, motivation and learning ability. It is beyond the scope of this paper to estimate a full structural education production function.

Our main variable of interest is the frequency of cannabis use. The questionnaire asks about the frequency of consumption over the month preceding the interview with possible answers ranging from "never" to "daily use" (i.e., never, 1-3 times a month, 1-2 times a week, 3-5 times a week, and daily). We construct two dummy variables. First, we create an indicator for *any use* that takes the value 1 if the individual has smoked cannabis at least once during the month preceding the interview. Then, we create an indicator for *frequent use* that makes the distinction between frequent users (at least once a week) and never- and occasional users (i.e., never, 1-3 times a month).

The PISA survey includes an extensive set of individual and family characteristics such as gender, living in a nuclear family, parental education, parental wealth, parental socioeconomic status and number of siblings. It also collects information on household educational support (i.e., parental educational support, number of books at home and educational resources at home) and on educational outcomes during the last year of secondary school (i.e., reading and math test scores). We were able to match each respondent to its related information collected in the PISA 2000 survey and therefore obtain a rich set of baseline (i.e., pre-high-school) relevant characteristics. The opportunity to control for baseline ability measured with reading and math test scores is particularly appealing to our approach.

In addition, the TREE survey itself includes a large set of variables measuring psychological traits, non-cognitive skills and substance use. We exploit this information and use a series of scales measured at Wave 1 reflecting persistency, self-efficacy, self-esteem and positive attitude. Each of these psychological variables is constructed by aggregating answers to a series of questions[2]. We also create dummy variables reflecting alcohol and tobacco consumption at Wave 1. Summary statistics for all relevant variables are presented in Table 3.2.

[2]Details on the construction of these variables are provided in Table 3.10 in the Appendix.

Table 3.2: Descriptive statistics

Variables	**Mean**	**s.d.**	**Min**	**Max**
Cannabis use: average over the three years (TREE, Waves 1 to 3)				
Any use	0.16	0.37	0	1
Frequent use	0.07	0.25	0	1
Outcomes: average over the three years (TREE, Waves 1 to 3)				
School days skipped (per month)	1.57	2.49	0	20
School difficulties index	1.00	1.08	0	5
Poor grades	0.51	0.49	0	1
Lack of engagement	0.14	0.35	0	1
Lack of motivation	0.22	0.42	0	1
Concentration problems	0.13	0.33	0	1
Control variables measured in 2000 (PISA)				
Female	0.62	0.49	0	1
More than 100 books at home	0.70	0.41	0	1
Nuclear family	0.84	0.36	0	1
Index of family wealth	0.14	0.76	-2.31	3.38
Index of family educational ressources	0.47	0.61	-3.42	0.76
Index of family educational support	0.01	0.90	-1.49	3.35
Tertiary education: mother	0.29	0.45	0	1
Tertiary education: father	0.50	0.50	0	1
Number of siblings	2.1	2	0	20
Index of socioeconomic status	59.5	16	16	90
Reading test score	5.10	0.85	2	9.5
Math test score	5.02	0.99	1.8	10
Control variables measured in 2001 (TREE, Wave 1)				
Persistency	12.81	3.32	4	16
Self-efficacy	12.40	3.42	4	16
Self-esteem	7.70	5.55	-15	16
Positive attitude	24.15	4.44	6	30
Any alcohol use	0.66	0.47	0	1
Any tobacco use	0.25	0.43	0	1

The *index of family wealth* reflects goods and characteristics of the household (dishwasher, students own room, Internet connection, number of mobile phones, televisions, computers, cars, and number of bathrooms). The *index of family educational resources* reflects the availability of a dictionary, a quiet place to study, a desk for study, textbooks, and of calculators at home. The *index of family educational support* reflects the frequency at which family members are involved with the student's schoolwork: mother, father, and siblings. The index of PISA International Socio-Economic Index of Occupational Status that ranges from 16 to 90 is used as a measure of *socioeconomic status* (Ganzeboom et al., 1992). *Reading test score* reflects student's ability in reading. *Math test score* reflects student's ability in mathematics.

3.3 Empirical approach

Our objective is to uncover the impact of cannabis consumption on a series of short-term educational outcomes. The main empirical challenge is that any observed correlation between cannabis use and poor educational outcomes may be due to the influence of common unobserved factors; or it may be that low performance at school increases the propensity to engage in risky behaviours. In this paper, we exploit the longitudinal nature of our data and a rich set of control variables to overcome these potential issues. More precisely, we estimate fixed effects regressions as well as propensity score matching to reduce the selection bias while the issue of simultaneity is accounted for by considering a lagged value of cannabis consumption. The use of two empirical strategies allows us to assess the robustness of our results and provides different ways to evaluate the importance of unobserved heterogeneity.

An alternative option to uncover the causal impact of cannabis use on short-term academic performance would have been to use an instrumental variable approach. However, credible instruments are challenging to find in substance use research (French and Popovici, 2009), especially in the case of illegal drug use. Indeed, a variable that would impact academic performance only through cannabis use is not available in our case and we therefore rely on changes in consumption over time for identification.

3.3.1 Pooled OLS and fixed effects

We start by estimating a series of OLS and linear probability models to investigate the association between cannabis use and educational outcomes. More precisely, we model the association between lagged cannabis use, $C_{i,t-1}$, and contemporaneous educational outcomes, Y_{it}. Our baseline specification is:

$$Y_{it} = \alpha_0 + \alpha_1 C_{i,t-1} + \alpha_2 X_i + \delta_t + \nu_{it} \tag{3.1}$$

where X_i represents a vector of baseline characteristics, δ_t is a wave indicator that accounts for trends, and ν_{it} is an idiosyncratic error term. We use a lagged measure of cannabis use in order to mitigate potential bias arising from reverse causality. Also, it is worth noting that for the outcomes related to lack of engagement, lack of motivation, poor grades and concentration problems, we decided to estimate linear probability models for ease of interpretation and better comparability between outcomes.

For the OLS results to be considered as unbiased, we must make the assumption that lagged consumption is exogenous. However, it is likely that some unobserved individual characteristics affect both consumption and the outcomes of interest (e.g., time preferences, peer influence, rebelliousness or preference for deviant behaviour). We therefore exploit the longitudinal nature of our data and extend (3.1) by controlling for individual fixed effects, η_i:

$$Y_{it} = \alpha_0 + \alpha_1 C_{i,t-1} + \alpha_2 X_i + \delta_t + \underbrace{\eta_i + \epsilon_{it}}_{\nu_{it}} \tag{3.2}$$

Practically, we use the within-estimator that purges the estimates from the influence of time-invariant individual characteristics:

$$(Y_{it} - \bar{Y_i}) = \alpha(C_{i,t-1} - \bar{C_i}) + \delta(t - \bar{t}) + (\epsilon_{it} - \bar{\epsilon_i}) \tag{3.3}$$

Our main identifying assumption is that after controlling for fixed individual characteristics, there are no other unobserved factors that both influence lagged cannabis use and educational outcomes.

3.3.2 Propensity score matching

To assess the robustness of our results, we use propensity score matching (PSM), a non-parametric approach that relaxes the linearity assumption inherent to the use of OLS and FE estimators. Precisely, we compare short-term educational outcomes of cannabis users to those of a matched group of non-users (or occasional users) with similar observed characteristics. Due to dimensionality issues, performing exact matching with a large number of covariates is challenging (Dehejia and Wahba, 2002). PSM overcomes this problem by matching individuals based on their estimated probability to belong to the treatment group (i.e., their propensity score) (Rosenbaum and Rubin, 1983). With this procedure, we can compare individuals who are similar in terms of observed characteristics but who differ in their use of cannabis. Identification relies on the assumption that there are no remaining unobserved characteristics correlated with both cannabis onset and educational outcomes.

In our approach, we exploit the longitudinal nature of our data and define treated adolescents as those who reported smoking cannabis at any frequency at Wave 2 but who did not smoke at Wave 1. In other words, our treatment of interest is the onset of cannabis use between Waves 1 and 2 (in alternative specifications, we modify the treatment of interest and focus on the onset of frequent cannabis use). To avoid reverse causality issues, educational outcomes are measured in Wave 3. As for the pooled OLS and FE specifications, we have then a lagged measure of cannabis use.

We start by estimating the probability of cannabis initiation, i.e., the probability to belong to the treatment group. We use pre-treatment characteristics as defined above, measured both in PISA and at Wave 1, that are likely to influence both cannabis use and education and estimate logit models of the form:

$$ln\left(\frac{P_i}{(1-P_i)}\right) = \beta_0 + \beta_1 X_i + \gamma\mu_i \tag{3.4}$$

where P_i is the predicted probability to start consuming cannabis (regularly) at Wave 2, X_i is a vector of pre-determined characteristics and μ_i represents unobserved heterogeneity. The parameter γ reflects potential correlation remaining between unobserved characteristics and the participation decision. We first assume conditional independence, which implies that, after controlling for X_i, γ is equal to zero.

The next step consists in forming pairs of treated and untreated individuals that have similar predicted probabilities to be treated (i.e., similar propensity scores). We use several matching estimators, including kernel matching and bias-corrected nearest-neighbour matching with single and multiple neighbours (Abadie and Imbens, 2011). We then estimate the average treatment effect on the treated (ATT) by comparing educational outcomes between the two groups at Wave 3. We have, for each outcome k:

$$\tau^k_{ATT} = E\left[Y^k_{1i}\,|C_i = 1, P_i\right] - E\left[Y^k_{0i}\,|C_i = 1, P_i\right] \tag{3.5}$$

where C_i is the treatment variable (i.e., starting consuming (regularly) cannabis at Wave 2) and k the outcome of interest at Wave 3 (with $k \in 1, ..., 6$).

In order to evaluate the sensitivity of our results to potential unobserved heterogeneity, we use Rosenbaum bounds (Rosenbaum, 2002). This method examines how the confidence intervals around the ATT are affected by different assumptions about the value of γ in (3.4). To get some intuition about this procedure, consider two individuals from a matched pair, indexed by i

and j, who have the same values of observed covariates. Rosenbaum (2002) has shown that, in the presence of unobserved characteristics that affect the participation decision, these two individuals may differ in their odds of receiving treatment by a factor Γ (see Rosenbaum, 2002, 2003, 2005):

$$\frac{1}{\Gamma} \leq \frac{P_i(1-P_j)}{P_j(1-P_i)} \leq \Gamma \tag{3.6}$$

where $\Gamma = e^{\gamma}$. If the (untestable) conditional independence assumption holds, Γ is equal to one. The sensitivity analysis proposed by Rosenbaum computes the range of significance levels for several values of Γ and therefore informs us about how sensitive our findings are to potential biased treatment assignment[3]. This approach does not provide a formal test of the CIA but allows researchers to gauge the sensitivity of their findings to potential selection on unobservables.

3.4 Results

3.4.1 Pooled OLS and fixed effects

Table 3.3 provides results for any cannabis use (panel A) and frequent cannabis use (panel B). For each outcome (i.e., absenteeism, school difficulties, lack of engagement, lack of motivation, poor grades and concentration problems), the table shows the coefficient of interest obtained with both the OLS and FE specifications[4] (ref. equations (3.1) and (3.2)).

OLS results for absenteeism suggest that cannabis use increases the number of school days skipped among high school students, irrespective of frequency of use. After controlling for fixed unobserved factors the coefficient remains significant for any use only: fixed effects results show that cannabis users skip on average 0.6 additional school days per month as compared to non-users. We do not find consistent evidence of an impact of cannabis use on the index of self-reported school difficulties, except in the OLS model for frequent use. Frequent cannabis use has a positive impact on self-assessed lack of attention in the classroom. Fixed effects results indicate that frequent users are approximately 13% more likely to report attention deficit in

[3]It is worth noting that we make the assumption of a potential positive selection bias (i.e., unobserved factors that are positively correlated with both cannabis use and poor educational outcomes).

[4]Tables 3.6 to 3.9 with full results are presented in the Appendix.

Table 3.3: OLS and FE models for both any and frequent cannabis use

Panel A: Any cannabis use	**Absenteeism**		**School difficulties**		**Lack of engagement**		**Lack of motivation**		**Poor grades**		**Concentration problems**	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE
Any cannabis (lag)	0.662***	0.626**	0.081	-0.030	0.090***	0.025	0.086**	0.051	0.043	0.090*	0.034	-0.021
	(0.181)	(0.282)	(0.082)	(0.099)	(0.031)	(0.046)	(0.035)	(0.054)	(0.036)	(0.048)	(0.026)	(0.041)
Panel B: Frequent cannabis use												
	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE
Frequent cannabis (lag)	1.094***	0.590	0.502***	0.124	0.144***	0.133*	0.065	0.680	0.143***	0.228***	0.130***	0.032
	(0.306)	(0.370)	(0.143)	(0.156)	(0.050)	(0.080)	(0.048)	(0.069)	(0.045)	(0.059)	(0.041)	(0.053)

Each coefficient represents a separate regression. Each OLS model controls for gender, family type, family wealth, family educational resources, family educational support, mothers education, fathers education, number of siblings, socioeconomic status of the parents, reading test score, math test score, each measured in 2000. Models also control for alcohol consumption, tobacco use, persistency, self-efficacy, self-esteem, and positive attitude, each measured at Wave 1 (i.e., in 2001). A Wave dummy is included in each model. Robust standard errors are in parentheses. Full results are displayed in the Appendix.*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

the classroom. Models for motivation do not suggest any association between cannabis use and this outcome except in one OLS specification. The most sizeable effects of cannabis use are found for the "poor grades" outcome. Fixed effects estimates are positive and significant in all models. Overall, results suggest that cannabis consumption increases the probability of receiving poor grades in the last grade report by 9 to 23 percentage points, with stronger effects found among frequent users. Finally, we find a positive and significant association between frequent cannabis use and recent concentration problems at school in the OLS model. However, these results do not hold when individual fixed effects are controlled for.

3.4.2 Propensity score matching

We now turn to the results obtained with propensity score matching. Logistic estimates for both any and frequent cannabis use are displayed in Table 3.4 (ref. equation (3.4)). We observe that being a female and having grown-up in a nuclear or wealthy family are negatively associated with the onset of any cannabis use. Baseline tobacco and alcohol consumption increases the probability of cannabis initiation. Interestingly, some psychological traits seem to play a protective role, including persistency and positive attitude towards life. Figure 3.1 in the Appendix shows the distribution of the propensity scores for both treatment and control groups and therefore provides an assessment of the overlap condition.

Table 3.5 reports the ATT estimates (ref. equation (3.5)) for each outcome and also includes the critical values of Γ obtained with Rosenbaum bounds. The interpretation of the critical values is discussed below. As matching algorithm, we decided to rely on nearest neighbour and kernel matching. Nearest neighbour with one neighbour (NN), respectively with five neighbours (NN(5)), assigns a weight one to the closest non-treated observation(s) and zero to all others. Kernel matching defines a neighbourhood for each treated unit and constructs the counterfactual using all untreated units with this neighbourhood, not only the closest unit. It assigns a positive weight to all units within the neighbourhood and a zero weight to all others.

The ATTs for absenteeism are positive and significant for both any and frequent cannabis use with kernel and NN(5) matching[5]. Estimates suggest

[5]Bootstrap standard errors are used to assess the statistical significance of the ATT estimates.

Table 3.4: Logistic regressions, any and frequent cannabis use

Models	**Any use**	**Frequent use**
Explanatory variables	**Coefficients**	**Coefficients**
Female	-0.432**	-1.148***
	(0.186)	(0.255)
More than 100 books at home	-0.022	.0.120
	(0.213)	(0.285)
Nuclear family	-0.419**	-0.268
	(0.209)	(0.278)
Index of family wealth	-0.250**	-0.269*
	(0.115)	(0.153)
Index of home educational ressources	0.210	0.130
	(0.135)	(0.177)
Index of family educational support	-0.107	0.018
	(0.102)	(0.139)
Tertiary education: mother	0.232	0.504*
	(0.200)	(0.265)
Tertiary education: father	-0.175	-0.465
	(0.204)	(0.283)
Number of siblings	0.046**	0.016
	(0.023)	(0.032)
Index of socioeconomic status	0.007	0.015
	(0.006)	(0.009)
Language test score	0.004	-0.169
	(0.125)	(0.174)
Math test score	-0.170	-0.040
	(0.107)	(0.144)
Any alcohol use	0.942***	0.727**
	(0.232)	(0.353)
Any tobacco use	1.565***	1.891***
	(0.180)	(0.260)
Persistency	-0.170***	-0.263***
	(0.050)	(0.068)
Self-efficacy	0.169***	0.156**
	(0.055)	(0.073)
Self-esteem	0.012	-0.026
	(0.021)	(0.028)
Positive attitude	-0.103***	-0.074*
	(0.031)	(0.039)
Constant	0.428	0.428
	(0.989)	(1.320)
LR chi(2)	228.83	172.45
Prob > Chi2	0.000	0.000
Pseudo R-squared	0.200	0.242
N	1196	1196

*,**and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

that cannabis users skip on average 0.5 to 1.6 more school days per month than non-users. Results for school difficulties and poor grades are significant across all matching methods but only when we consider frequent cannabis use. Findings for poor grades are of similar magnitude than those obtained with the fixed effects specifications above and suggest that cannabis users have a 12 to 19% higher probability of obtaining poor grades.

Table 3.5: Matching estimates and Rosenbaum bounds

Matching techniques	**NN**	**Kernel**	**NN(5)**
Absenteism			
Any cannabis use	0.440	0.594**	0.672***
Γ	1	1	1
Frequent cannabis use	0.710	0.996**	1.159***
Γ	1	1.15	1.35
School difficulties			
Any cannabis use	-0.086	0.064	0.043
Γ	1	1	1
Frequent cannabis use	0.446*	0.370*	0.386**
Γ	1.6	1.05	1.05
Lack of engagement			
Any cannabis use	0.121*	0.093**	0.086*
Γ	1.71	1	1
Frequent cannabis use	0.067	0.141***	0.109
Γ	1	1.1	1
Lack of motivation			
Any cannabis use	0.096	0.068	0.098
Γ	1	1	1
Frequent cannabis use	0.013	0.035	0.019
Γ	1	1	1
Poor grades			
Any cannabis use	0.048	0.058	0.078
Γ	1	1	1
Frequent cannabis use	0.187**	0.146*	0.120*
Γ	1.35	1.2	1.1
Concentration problems			
Any cannabis use	0.022	0.020	0.001
Γ	1	1	1
Frequent cannabis use	0.127	0.083	0.084
Γ	1	1	1

Cannabis use (any and frequent) are in a lagged form.
Γ values represent the level at which p-values are critical.
NN refers to nearest neighbour matching with 1 control case.
NN(5) refers to nearest neighbour matching with 5 control cases.

To assess the robustness of these results to potential selection bias, we turn to the interpretation of the critical values of Γ. These values reflect the minimum amount of selection on unobservables that would produce estimates that are no longer statistically significant. For example, in the case of absenteeism, the critical value of Γ equals 1.35 (NN(5) matching, frequent use), meaning that the presence of unobserved characteristics that would make individuals 35% more likely to be in the treatment group would bias the results. These values do not indicate whether our estimates are spurious but provide an indication of how confident we can be in interpreting our estimates as being unbiased. Overall, even if results for absenteeism, school difficulties and poor grades are consistently significant, the critical values of Γ for these outcomes never exceed 1.6. However, as mentioned in DiPrete and

Gangl (2004), these results are "worst-case scenarios." In other words, they do not indicate the presence of selection bias but only tell use how strong the selection bias should be to invalidate our conclusions.

3.5 Discussion

In this paper, we investigate the impact of cannabis use on short-term educational outcomes among high school students. We exploit a Swiss longitudinal dataset that follows a cohort of adolescents annually starting at the end of compulsory school and that collects information on educational outcomes, substance use, and on a wide range of individual characteristics. We consider six different outcomes and are able to control for a rich set of baseline characteristics at both the individual and family level. Results obtained with two distinct empirical strategies consistently show that cannabis users skip school more often and are more likely to obtain poor grades than non-users. More precisely, we observe strong effects of cannabis use on an indicator of exposure to schooling (i.e., school days skipped) and on an indicator of performance (i.e., grades).

These results are in line with previous findings (Pacula et al., 2003; Roebuck et al., 2004; Engberg and Morral, 2006) and should be taken into account in the development of future messages on the risks of cannabis use. With the unclear signals sent by a relatively permissive legislation and an increasingly widespread use of this product for medical purposes, adolescents may underestimate the full consequences of cannabis use. Besides the role played by information campaigns and school-based programs, professional workers in social or health services also stress the importance of parental implication to increase awareness among adolescents that even occasional use might impair their ability to effectively engage in school and may reduce their overall performance. Such a statement also advocates for a closer collaboration between health and education policies.

This study has, however, several important limitations. First, self-reported measures are used for both cannabis use and educational outcomes. These two groups of variables may be subject to intentional misreporting and results may therefore suffer from attenuation bias. Second, we are not able to assess whether our findings on poor grades are driven by impaired cognitive ability or by reduced attendance. Additional analyses are needed to investigate these potential mechanisms in more details and to define the proper interventions. Finally, the information on cannabis use only informs use on the frequency of use but does not provide insights in the intensity of use, neither on the context in which the product is more often consumed.

3.6 References

Abadie, A. and G. Imbens, Bias-corrected matching estimators for average treatment effects, *Journal of Business and Economic Statistics*, 29(1), pp. 1-11, 2011.

Bessey D. and U. Backes-Gellner, Marijuana consumption, educational outcomes and labor market success: Evidence from Switzerland, Leading House Working Paper n° 43, 2009.

Bray, J., G. Zarkin, C. Ringwalt and J. Qi, The relationship between marijuana initiation and dropping out of high school, *Health Economics*, 9(1), pp. 9-18, 2000.

Budney A., Are specific dependence criteria necessary for different substances: How can research on cannabis inform this issue?, *Addiction*, 101 (Suppl 1), pp. 125-133, 2006.

Budney A., B. Moore, H. Rocha and S. Higgins, Clinical trial of abstinence-based vouchers and cognitive-behavioral therapy for cannabis dependence, *Journal of Consulting and Clinical Psychology*, 74(2), pp. 307-316, 2006.

Budney A., R. Roffman, R. Stephens and D. Walker, Marijuana dependence and its treatment, *Addiction science and clinical practice*, 4(19), pp. 4-16, 2007.

Caldeira K., A. Arria, K. O'Grady, K. Vincent and E. Wish, The occurence of cannabis use disorders and other cannabis-related problems among first-year college students, *Addictive behaviours*, Vol. 33, pp. 397-411, 2008.

Dehejia, R. and S. Wahba, Propensity score-matching methods for nonexperimental causal studies, *Review of Economics and statistics*, 84(1), pp. 151-161, 2002.

DiPrete T. and M. Gangl, Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimations with imperfect instruments, *Sociological Methodology*, 34(1), pp. 271-310, 2004.

Engberg J. and Morral A., Reducing substance use improves adolescents school attendance, *Addiction*, Vol. 101, pp. 1741-1751, 2006.

Fontes, M., K. Bolla, P. Cunha, P. Almeida, F. Jungerman, R. Laranjeira, R. Bressan and A. Lacerda, Cannabis use before age 15 and subsequent executive functioning, *British Journal of Psychiatry*, 198(6), pp. 442-447, 2011.

French M. and I. Popovici, That instrument is lousy! In search of agreement when using instrumental variables estimation in substance use research, *Health Economics*, Vol. 20, pp. 127-146, 2011.

Ganzeboom H., P. De Graaf and J. Treiman, A standard international socioeconomic index of occupational status, *Social sciences Research*, 21(1), pp. 1-56, 1992.

Gmel G., H. Kuendig, E. Maffli, L. Notari, M. Wicki, A. Georges, E. Grisel-Staub, M. Müller, F. Dubois-Arber, J.-P. Gervasoni, S. Lucia, A. Jeannin, A. Uchtenhagen and M. Schaub, Monitorage suisse des addictions: Rapport annuel - Données 2011, Berne 2012.

Gruber, S., K. Sagar, M. Dahlgren, M. Racine and S. Lukas, Age of onset of marijuana use and executive function, *Psychol Addict behaviour*, 26(3), pp. 496-506, 2012.

Horwood, L., D. Fergusson, M. Hayatbakhsh, J. Najman, C. Coffey, G. Patton, E. Silins and D. Hutchinson, Cannabis use and educational achievement: Findings from three Australasian cohort studies, *Drug and Alcohol Dependence*, 110(3), pp. 247-253, 2010.

Jager G. and N. Ramsey, Long-term consequences of adolescent cannabis exposure on the development of cognition, brain structure and function: on overview of animal and human research, *Curr Drug Abuse Rev*, Vol. 1, pp.114-123, 2008.

Lynskey, M., C. Coffey, L. Degenhardt, J. Carlin and G. Patton, A longitudinal study of the effects of adolescent cannabis use on high school completion, *Addiction*, 98(5), pp. 685-692, 2003.

MacLeod, J., R. Oakes, A. Copello, I. Crome, M. Egger, M. Hickman, T. Oppenkowski, H. Stokes-Lampard and G. Davey Smith, Psychological and social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies, *Lancet*, 363(9421), pp. 1579-1588, 2004.

McCaffrey, D., R. Liccardo Pacula, B. Han and P. Ellickson, Marijuana use and high school dropout: the influence of unobservables, *Health Economics*, 19(11), pp. 1281-1299, 2010.

Meier, M., A. Caspi, A. Ambler, H. Harrington, R. Houts, R. Keefe, K. McDonald, A. Ward, R. Poulton and T. Moffitt, Persistent cannabis users show neuropsychological decline from childhood to midlife, *Proceedings of the National Academy of sciencess*, 109(40), pp. E2657-E2664, 2012.

Menghrajani, P., K. Klaue, F. Dubois-Arber and P.-A. Michaud, Swiss adolescents and adults perceptions of cannabis use: a qualitative study, *Health Education Research*, 20(4), pp. 476-484, 2005.

Monti P., R. Miranda, K. Nixon, K. Sher, H. Swartzwelder, S. Tapert et al., Adolescence: booze, brains, and behaviour, *Alcohol Clin Exp Res*, Vol. 29, pp. 207-220, 2005.

Pacula R., J. Ringel and K. Ross, Does marijuana use impair human capital formation?, Rand Health Working Paper, WE-121, 2003.

Register, C., D. Williams and P. Grimes, Adolescent drug use and educational attainment, *Education Economics*, 9(1), pp. 1-18, 2001.

Roebuck, M., M. French and M. Dennis, Adolescent marijuana use and school attendance, *Economics of Education Review*, 23(2), pp. 133-141, 2004.

Rosenbaum, P., *Observational Studies*, Second Edition, New York: Springer, 2002.

Rosenbaum, P., Does a dose-response relationship reduce sensitivity to hidden bias?, *Biostatistics*, 4(1), pp. 1-10, 2003.

Rosenbaum, P., Sensitivity analysis in observational studies, in B.S. Everitt and D.C. Howell (eds.), *Encyclopedia of statistics in behavioural sciences*, pp. 1809-1814, Chichester: John Wiley Sons, Ltd, 2005.

Rosenbaum, P. and D. Rubin, The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70(1), pp. 41-55, 1983.

Solowij N. and P. Michie, Cannabis and cognitive dysfunction: parallels with endophenotypes of schizophrenia, *J Psychiatry Neurosci*, Vol. 32, pp. 30-52, 2007.

Solowij N. and R. Battisti, The chronic effects of cannabis on memory in humans: a review, *Curr Drug Abuse Review*, Vol. 1, pp. 81-98, 2008.

Stephens R., T. Babor, R. Kadden and M. Miller. The Marijuana Treatment Project: Rationale, design, and participant characteristics, Addiction, 97(Suppl 1), pp. 109-124, 2002.

Van Ours, J. and J. Williams, Why parents worry: Initiation into cannabis use by youth and their educational attainment, *Journal of Health Economics*, 28(1), pp. 132-142, 2009.

Yamada T., M. Kendix and T. Yamada, The impact of alcohol consumption and marijuana use on high school graduation, *Health Economics*, Vol. 5, pp. 77-92, 1996.

3.7 Appendices

3.7.1 Distribution of the propensity score

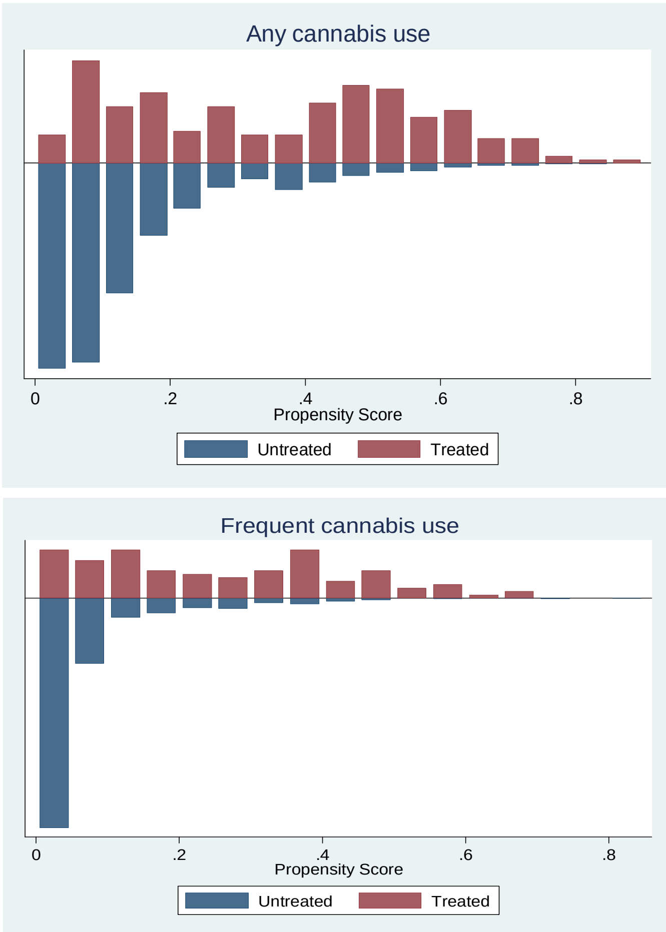


Figure 3.1: Distribution of the propensity score among cannabis users and non-users

3.7.2 Additional tables

Table 3.6: Any cannabis use, OLS estimation

	Absenteeism	**School difficulties**	**Lack of engagement**	**Lack of motivation**	**Poor grades**	**Concentration problems**
Any cannabis use	0.662***	0.0808	0.0896***	0.0857***	0.043	0.034
	(0.181)	(0.082)	(0.031)	(0.035)	(0.036)	(0.026)
Female	0.130	0.230***	-0.040*	-0.045*	-0.062**	0.045***
	(0.131)	(0.060)	(0.021)	(0.024)	(0.028)	(0.017)
More than 100 books at home	-0.389*	-0.006	0.004	-0.048*	-0.079**	0.005
	(0.200)	(0.076)	(0.0235)	(0.0275)	(0.033)	(0.021)
Nuclear family	-0.253	-0.067	-0.060**	0.011	-0.070**	-0.015
	(0.194)	(0.080)	(0.029)	(0.028)	(0.034)	(0.024)
Family wealth	0.198**	-0.024	0.010	0.013	0.053***	0.015
	(0.085)	(0.040)	(0.012)	(0.014)	(0.018)	(0.013)
Home educ. resources	-0.110	-0.031	-0.034**	-0.030*	-0.013	-0.018
	(0.100)	(0.050)	(0.017)	(0.018)	(0.020)	(0.013)
Family educ. support	-0.024	0.084**	-0.015	-0.006	0.001	-0.012
	(0.072)	(0.034)	(0.011)	(0.013)	(0.014)	(0.009)
Tertiary education: mother	0.079	0.088	0.007	0.058**	-0.008	-0.009
	(0.153)	(0.066)	(0.022)	(0.026)	(0.031)	(0.019)
Tertiary education: father	0.057	-0.064	-0.014	-0.021	-0.024	0.014
	(0.145)	(0.066)	(0.021)	(0.024)	(0.031)	(0.019)
Number of siblings	-0.005	-0.001	0.000	0.006	0.010***	-0.002
	(0.019)	(0.009)	(0.003)	(0.004)	(0.003)	(0.002)
Socioeconomic status	0.005	-0.001	0.001	0.001*	-0.002*	-0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Language test score	0.040	0.041	-0.000	-0.001	-0.019	-0.008
	(0.083)	(0.043)	(0.014)	(0.016)	(0.020)	(0.012)
Math test score	-0.155**	-0.073**	-0.003	0.024*	-0.088***	-0.013
	(0.070)	(0.035)	(0.011)	(0.013)	(0.016)	(0.010)
Any alcohol use	0.155	-0.010	0.022	0.026	-0.032	-0.009
	(0.131)	(0.061)	(0.018)	(0.021)	(0.028)	(0.017)
Any tobacco use	0.561***	0.121	0.056**	0.059**	0.048	0.080***
	(0.163)	(0.076)	(0.026)	(0.030)	(0.033)	(0.024)
Persistency	-0.085**	0.065***	-0.032***	-0.029***	-0.026***	-0.015***
	(0.035)	(0.017)	(0.006)	(0.007)	(0.008)	(0.005)
Self-efficacy	0.079**	-0.091***	0.012**	0.015**	-0.004	0.003
	(0.039)	(0.019)	(0.006)	(0.007)	(0.008)	(0.006)
Self-esteem	-0.011	-0.023***	-0.000	0.001	-0.000	-0.007***
	(0.017)	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)
Positive attitude	0.026	0.004	-0.005	-0.012*	-0.003	-0.004
	(0.023)	(0.012)	(0.004)	(0.004)	(0.00482)	(0.00358)
Time trend	0.357***	-0.043	0.001	0.016	-0.015	0.007
	(0.107)	(0.036)	(0.013)	(0.015)	(0.017)	(0.012)
Constant	1.533*	1.495***	0.542***	0.470***	1.764***	0.523***
	(0.805)	(0.376)	(0.123)	(0.135)	(0.152)	(0.112)
N	1867	1977	2004	2006	1997	2101
R^2	0.060	0.085	0.081	0.073	0.094	0.077

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 3.7: Frequent cannabis use, OLS estimation

	Absenteeism	**School difficulties**	**Lack of engagement**	**Lack of motivation**	**Poor grades**	**Concentration problems**
Frequent cannabis use	1.094***	0.502***	0.144***	0.065	0.143***	0.130***
	(0.306)	(0.143)	(0.050)	(0.048)	(0.0454)	(0.0412)
Female	0.175	0.255***	-0.034	-0.043*	-0.055*	0.052*
	(0.132)	(0.060)	(0.021)	(0.024)	(0.028)	(0.017)
More than 100 books at home	-0.392**	-0.010	0.004	-0.047*	-0.080**	0.004
	(0.200)	(0.075)	(0.024)	(0.028)	(0.033)	(0.021)
Nuclear family	-0.250	-0.060	-0.060**	0.010	-0.070**	-0.014
	(0.195)	(0.080)	(0.030)	(0.029)	(0.034)	(0.024)
Family wealth	0.197**	-0.018	0.010	0.012	0.054***	0.016
	(0.085)	(0.040)	(0.012)	(0.014)	(0.018)	(0.013)
Home educ. resources	-0.100	-0.027	-0.033*	-0.029	-0.012	-0.017
	(0.100)	(0.049)	(0.017)	(0.018)	(0.020)	(0.013)
Family educ. support	-0.030	0.084**	-0.016	-0.007	0.000	-0.012
	(0.072)	(0.034)	(0.011)	(0.013)	(0.014)	(0.009)
Tertiary education: mother	0.071	0.083	0.005	0.057**	-0.009	-0.010
	(0.153)	(0.065)	(0.022)	(0.026)	(0.031)	(0.019)
Tertiary education: father	0.069	-0.056	-0.012	-0.021	-0.021	0.016
	(0.143)	(0.065)	(0.021)	(0.024)	(0.031)	(0.019)
Number of siblings	-0.008	-0.004	-0.000	0.006	0.010***	-0.003
	(0.019)	(0.009)	(0.003)	(0.004)	(0.003)	(0.002)
Socioeconomic status	0.005	-0.002	0.001	0.001*	-0.002*	-0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Language test score	0.049	0.044	0.001	-0.000	-0.018	-0.007
	(0.083)	(0.043)	(0.014)	(0.016)	(0.020)	(0.012)
Math test score	-0.164**	-0.073**	-0.005	0.022*	-0.088***	-0.013
	(0.070)	(0.036)	(0.011)	(0.013)	(0.016)	(0.010)
Any alcohol use	0.176	-0.017	0.025	0.031	-0.032	-0.010
	(0.130)	(0.061)	(0.018)	(0.022)	(0.028)	(0.017)
Any tobacco use	0.596***	0.062	0.061**	0.077***	0.038	0.070***
	(0.159)	(0.072)	(0.026)	(0.029)	(0.033)	(0.023)
Persistency	-0.083**	0.070***	-0.031***	-0.030***	-0.024***	-0.014***
	(0.035)	(0.017)	(0.006)	(0.007)	(0.00774)	(0.00501)
Self-efficacy	0.083**	-0.093***	0.013**	0.0160**	-0.005	0.003
	(0.040)	(0.019)	(0.006)	(0.007)	(0.008)	(0.006)
Self-esteem	-0.008	-0.022***	0.000	0.001	-0.000	-0.007***
	(0.017)	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)
Positive attitude	0.022	0.005	-0.005	-0.012***	-0.003	-0.004
	(0.023)	(0.012)	(0.004)	(0.004)	(0.005)	(0.004)
Time trend	0.351***	-0.053	-0.000	0.016	-0.017	0.005
	(0.107)	(0.036)	(0.013)	(0.015)	(0.017)	(0.012)
Constant	1.522*	1.415***	0.539***	0.483***	1.747***	0.504***
	(0.790)	(0.377)	(0.122)	(0.136)	(0.153)	(0.110)
N	1867	1977	2004	2006	1997	2101
R^2	0.062	0.095	0.083	0.070	0.098	0.083

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 3.8: Any cannabis use, FE estimation

	Absenteeism	School difficulties	Lack of engagement	Lack of motivation	Poor grades	Concentration problems
Any cannabis use	0.626**	-0.030	0.025	0.051	0.090*	-0.021
	(0.282)	(0.100)	(0.046)	(0.054)	(0.0476)	(0.0414)
Time trend	0.453***	0.001	0.017	0.026*	-0.013	0.014
	(0.107)	(0.035)	(0.013)	(0.015)	(0.017)	(0.012)
Constant	1.453***	0.998***	0.148***	0.202***	0.518***	0.130***
	(0.064)	(0.023)	(0.010)	(0.011)	(0.011)	(0.009)
N	1867	1977	2004	2006	1997	2101
R^2	0.032	0.000	0.002	0.005	0.004	0.002

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 3.9: Frequent cannabis use, FE estimation

	Absenteeism	School difficulties	Lack of engagement	Lack of motivation	Poor grades	Concentration problems
Frequent cannabis use	0.590	0.124	0.133*	0.0676	0.228***	0.0316
	(0.370)	(0.156)	(0.080)	(0.069)	(0.059)	(0.053)
Time trend	0.461***	-0.004	0.014	0.026*	-0.017	0.013
	(0.107)	(0.035)	(0.013)	(0.016)	(0.017)	(0.012)
Constant	1.512***	0.987***	0.144***	0.206***	0.519***	0.125***
	(0.054)	(0.018)	(0.008)	(0.008)	(0.008)	(0.006)
N	1867	1977	2004	2006	1997	2101
R^2	0.029	0.001	0.008	0.005	0.012	0.002

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 3.10: Description of variables related to psychological traits

Personality trait	Related questions
Persistency	If I decide to accomplish something, I manage to see it through
	I complete whatever I start
	Even if I encounter difficulties, I persisently continue
	I even keep at a painstaking task until I have carried it through
Self-efficacy	I can always manage to solve difficult problems if I try hard enough
	I am confident that I could deal efficiently with unexpected events
	Thanks to my resourcefulness, I know how to handle unforeseen
	I can usually handle whatever comes my way
Self-esteem	On the whole, I am satisfied with myself
	I feel that I have a number of good qualities
	I am able to do things as well as most of other people
	I feel that I am a person of worth, at least on an equal plane with others
	At times, I think I am not good at all
	I certainly feel useless at times
	I wish I could have more respect for myself
	All in all, I am included to fill that I am a failure
Positive attitude	My future looks bright
	I am happy to live
	I am happy with the way my life plan unfolds
	What ever happens, I can see the positive side of it
	My live seems to be meaningfull

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.
Possible values for persistency: “Completely false”, “Mostly false”, “Mostly true”, “Completely true”.
Possible values for self-efficacy: “Completely false”, “Mostly false”, “Mostly true”, “Completely true”.
Possible values for self-esteem: “Not at all true”, “Slightly true”, “Moderately true”, “Very true”, “Completely true”. Possible values for positive attitude: “Completely false”, “Mostly false”, “Somewhat false”, “Somewhat true”, “Mostly true”, “Completely true”.
In the analyses, each of these psychological traits is measured at baseline (i.e. in Wave 1). For persistency, self-efficacy and positive attitude, we construct indices by simply taking the sum of all items (possible values therefore range from 4 to 16 for persistency and self-efficacy and from 6 to 30 for positive attitude). The index of self-esteem is the sum of the first four items (“positive” self-esteem) minus the sum of the last four items (“negative” self-esteem). Self-esteem therefore ranges from -16 to 16.

Chapter 4

Who benefits most from university education in Switzerland?

Recent literature on private returns to education considers diversity in the population, heterogeneity in wage gains and self-selection into schooling. This research addresses these issues by analyzing to what extent returns associated with completing a university degree in Switzerland depend on the propensity to attend and complete this degree. Using data from the Swiss Household Panel and propensity score matching models, I find that low propensity men - after controlling for labour market variables - benefit most from a university degree while returns for women are rather homogeneous along the propensity score distribution. This finding suggests that completing university increases more the earnings capability of men with disadvantaged family backgrounds than that of men with more favorable background, refuting the hypothesis of comparative advantage at school. An auxiliary analysis focusing on the relationship between returns to education and inherent ability within a quantile regression framework leads to similar conclusions.

4.1 Introduction

While Switzerland's competitiveness is mainly based on innovation and human capital formation, graduation rates at university are particularly modest (and even low for individuals with migration backgrounds) in comparison with other high-innovation countries (OECD, 2013). The attractiveness of tertiary vocational education, the development of the universities of applied sciences[1] and the low expected returns associated with completing a university degree may explain such a phenomenon. Indeed, Wolter and Weber (2005) report that the wage premium obtained after university graduation is the lowest among all educational levels in the Swiss education system. This statement has been confirmed by a recent report from the Swiss Coordination Centre for Research in Education (CSRE, 2010). However, no prior study focusing on returns to schooling in Switzerland investigates if all students benefit to the same extent from a university degree in Switzerland while accounting for heterogeneity in returns to schooling is crucial to draw policy recommendations.

While traditional human capital theory (Becker, 1964; Mincer, 1974) assumes that higher education provides students with skills that are equally rewarded in the labour market (i.e., *productivity* explanation), the *selection* explanation considers that the positive relationship between wage and higher education results from a self-selection process based on individual heterogeneous attributes. Contrary to the traditional formulation of human capital where returns to schooling are implicitly assumed to be homogeneous, the selection explanation considers heterogeneity in returns to schooling, i.e., the impact of schooling on wages may differ across individuals with identical educational levels.

The literature on the selection explanation confronts two different viewpoints. First, the *positive selection* explanation assumes that selection in higher education is a rational decision based on expected gains in income, skills or knowledge, net of opportunity costs of pursuing educational investment (Willis and Rosen, 1979; Card, 1995, 2001). In other words, youths self-select into schooling on the basis of the principle of comparative advantage (Willis and Rosen, 1979; Carneiro et al., 2007; Heckman et al., 2006), which implies that high propensity students obtain higher returns to schooling because of their ability, motivation or favorable parental backgrounds. Some recent researches in sociology (Tsai and Xie, 2008; Brand and Xie, 2010), however, consider that the decision to attend higher education is not always rational because norms, expectations or encouragements may differ by family background, leading to different selections mechanisms (Coleman, 1988; Smith and Powell, 1990; Morgan, 2005). Individuals facing low labour market opportunities may have stronger economic incentives to invest in

[1] Compared to PhD-granting institutions, universities of applied sciences are more oriented towards practice (e.g., by giving the possibility to follow on-the-job training programs) and use other acceptance criteria (e.g., students with a professional matura are allowed to enter directly in these institutions).

higher education. One consequence of this *negative selection* hypothesis is that returns to higher education can be higher for low propensity individuals if we assume that individuals with favorable backgrounds may have access to superior labour market positions even in the absence of higher education. Expressed differently, completing a university degree may augment more the labour market opportunities of less-endowed individuals than that of their more-endowed peers.

In this context of self-selection where earnings level results from a combination between individual background attributes and the appropriate level of schooling, this study is a first attempt to determine what type of individuals benefits most from a university degree in Switzerland. For that purpose, I compare university and matura graduates on the grounds that the latter is the unique group having a direct access to university education [2]. The focus on parental background attributes is particularly relevant for the Swiss case given that access to university education for less-gifted individuals is full of economic, institutional, social or motivational barriers (Vellacott and Wolter, 2004). This research also aims at completing the literature on the intergenerational transmission of educational attainment in Switzerland whose results suggest that schooling choice is essentially determined by parental education and family income (Falter, 2005; Bauer and Riphahn, 2007; Cattaneo et. al, 2011).

The estimation procedure relies on hierarchical models based on propensity score matching similar to those used in Tsai and Xie (2008), Brand and Xie (2010) and Xie et al. (2011). This study extends their analyses in two ways. First, I consider Heckman selection models to account for sample selection bias. Second, the empirical analysis is completed by quantile regressions to analyze the relation between inherent ability and returns to schooling. While most of previous literature on schooling returns in Switzerland used data from the Swiss labour Force Survey (SLFS), this paper resorts to the Swiss Household Panel (SHP)[3] which contains numerous parental background variables of primary importance when estimating the probability to complete higher education.

[2]Since 2005, professional matura graduates have also access to university education but have to pass an additional formation of one year, with final examination ("Passerelle Dubs"). However, only 3% of these students uses this possibility (CSRE, 2010).

[3]This study has been realized using the data collected by the Swiss Household Panel (SHP), which is based at the Swiss Centre of Expertise in the Social Sciences FORS. The project is financed by the Swiss National Science Foundation.

The first part of the empirical study relies on a stratification-multilevel method that consists in estimating the returns to schooling across different propensity score strata. The different specifications lead to the conclusion that low propensity students benefit most from completing a university degree. The negative association between the propensity score and the returns to schooling, however, is only significant for men when including labour market variables. For women, the absence of a significant relation between these two variables in all models leads to the rejection of the assumption of heterogeneous returns along the propensity score distribution. The second part relies on a matching-smoothing method that fits a non-parametric regression to smooth the variation in matched wages' differences between matura and university graduates along the propensity score. The graphical analysis shows that returns to university education are rather homogeneous for both genders even if some local patterns can also sustain my previous statement for men. Finally, results from quantile regressions indicate that men with low inherent ability benefit most from university education. Expressed differently, education acts as a substitute for inherent ability in the generation of earnings capability. For women, returns are homogeneous along the conditional wage distribution.

The main findings of this study can be summarized as follows: *(i)* returns to university education for men decrease significantly along the propensity score when labour market experience is accounted for, which leads to the rejection of the comparative advantage hypothesis, *(ii)* women obtain homogeneous returns to university education, whatever the specification considered, and *(iii)* university education and inherent ability can be considered as substitutes in the human capital accumulation.

This chapter is organized as follows. The second section presents the theoretical framework and provides a short review of the literature on returns to schooling in Switzerland. Data are described in section three. Empirical models are presented in the fourth section while results are reported in section five. Section six presents an auxiliary analysis based on a quantile regression framework. The last section is devoted to the discussion.

4.2 Theoretical background and literature review

4.2.1 Theoretical framework

The main objective of causal inference is to determine the return to schooling for an individual i which can be written as:

$$\beta_i = Y_i^1 - Y_i^0 \tag{4.1}$$

where the two right-hand side terms of equation (4.1) correspond to the potential wages resulting from university or matura graduation, respectively. However, the fundamental problem of causal inference is that we cannot observe the same individual for two different treatment status simultaneously (Holland, 1986). As it is not possible to estimate the *individual* return to university education β_i, the literature concentrates on the *average* return to schooling $E[\beta_i]$ which is the expected value of the difference between the two potential outcomes and corresponds to the average treatment effect (ATE):

$$\text{ATE} = E[\beta_i] = E[Y_i|S_i = 1] - E[Y_i|S_i = 0] \tag{4.2}$$

Let us consider the following wage regression:

$$Y_i = \alpha_i + \beta_i S_i + u_i \tag{4.3}$$

where Y_i corresponds to the wage of individual i, S_i is a binary variable reflecting the highest educational attainment (with $S_i = 1$ corresponding to a university degree and $S_i = 0$ a matura degree) and u_i corresponds to unobserved heterogeneity. The conditional wage expectations can then be written as:

$$\begin{aligned} E[Y_i|S_i = 1] &= \alpha_i + E[\beta_i|S_i = 1] + E[u_i|S_i = 1] && (4.4) \\ E[Y_i|S_i = 0] &= \alpha_i + E[u_i|S_i = 0] && (4.5) \end{aligned}$$

In the presence of observational data where individuals are not randomly selected, we generally estimate a naive estimator of the ATE which can be decomposed as follows[4]:

$$\begin{aligned} \text{NATE} &= E[Y_i|S_i = 1] - E[Y_i|S_i = 0] \\ &= \underbrace{E[\beta_i]}_{ATE} + \underbrace{E[u_i|S_i = 1] - E[u_i|S_i = 0]}_{\text{Selection bias}} \\ &+ P(S_i = 0)\underbrace{\{E[\beta_i|S_i = 1] - E[\beta_i|S_i = 0]\}}_{\text{Return bias}} && (4.6) \end{aligned}$$

[4] See the full mathematical development of this formula (Roberts, 2009) in the Appendix.

The *selection bias* is the main econometric issue in the literature on returns to schooling. Indeed, the OLS estimator is biased when some components which influence both schooling and earnings (e.g., ability) are not included among the observed covariates (Griliches, 1977). If students with higher unobserved ability tend to acquire more schooling, the return to university education is upward biased. This potential non-random assignment into schooling explains why the main bulk of the literature resorts to an instrumental variable (IV) strategy to solve this selectivity bias (Kane and Rousse, 1993; Card, 1995; Harmon and Walker, 1995; Pons and Gonzalo, 2002). An alternative to control for unobserved ability consists in using family fixed effects when data on twins are available (Behrman et al., 1994; Miller et al., 1995; Ashenfelter and Rousse, 1998, Rantanen, 2009).

When the schooling choice results from a self-selection process where individuals act (partially) on their unobserved wage gains, the estimation procedure faces another econometric bias called the *return bias* which cannot be solved by traditional instrumental variables (IV) or fixed effects (FE) methods[5]. In this context, the traditional approach to account for the return bias is to rely on control functions which represent the conditional expectations of unobserved heterogeneity (Garen, 1984; Heckman and Robb, 1985; Heckman and Vytlacil, 1998; Deschênes, 2007). Generally, these functions are represented by the standard inverse Mills ratios from the normal selection model (Heckman, 1979)[6]. The inclusion of these functions in the wage regression allows to obtain selection corrected estimator for the return to schooling. However, the strong limitations faced by the traditional instrumental variables in this literature (see Checchi (2006) for some relevant examples) reduce considerably the possibility to estimate these control functions.

In the absence of randomization which solves both selection biases, propensity score matching (PSM) techniques emerge as the most interesting identification strategy when assuming selection on observables. By assuming conditional independence[7], the return to schooling can be formulated as fol-

[5]Instrumental variables may be independent of the unobserved wage gains in the overall population, but conditional on those who participate in university education, they may no longer be independent of the unobserved wage gains in this subgroup (Carneiro et al., 2001; Basu et al. 2007). FE methods control for pre-treatment heterogeneity but not for treatment effect heterogeneity (Xie et al., 2011).

[6]More recent studies which estimate the marginal treatment effect (MTE) parameter developed by Bjorklund and Moffitt (1987) but extended by Nobel Prize laureate James Heckman and his co-authors (2001, 2006, 2007) rely on more developed control functions where the conventional assumptions (i.e., linearity, normality and separability) are relaxed.

[7]The conditional independence assumption (CIA) also called unconfoundedness or ig-

lows:

$$E[\beta_i] = E[Y_i|S_i = 1] - E[Y_i|S_i = 0] \perp S_i|P(S_i = 1|X_i) \qquad (4.7)$$

Compared to the traditional OLS regression, PSM does not rely on a parametric assumption between the outcome and the covariates and considers a common support (or overlap condition) between treated and untreated units. Xie and Wu (2005), Brand and Xie (2010) and Xie et al. (2011) estimate heterogeneous returns to schooling by using hierarchical linear models based on PSM and focus on the association between the propensity score and the returns to schooling to determine the nature of the self-selection process. More precisely, they estimate the return to higher education for different subpopulations grouped according to their propensity to complete this degree. Compared to the traditional case in which heterogeneity in returns to schooling is determined through the interaction between education and specific covariates (e.g., gender or race) (Altonji and Dunn, 1996; Ashenfelter and Rouse, 1998), accounting for the propensity score is the best approach for solving the problems of variations by schooling participation.

4.2.2 Returns to schooling in Switzerland: a short literature review

Numerous studies have estimated the average return to years of education in Switzerland, especially by focusing on wages differentials by gender (Kugler, 1988; Diekmann and Engelhardt, 1995; Bonjour, 1997; Ferro-Luzzi and Silber, 1998). While there is no clear consensus on the magnitude of the return to schooling, the general picture reveals that men obtain higher returns than women. Most of studies, however, rely on the traditional OLS method to estimate the return to schooling, most of them correcting only for sample selection bias[8] (Kugler, 1998; Dieckmann and Engelhardt, 1995; Falter and Ferro-Luzzi, 2000). One exception in this literature is the study of Suter (2005) which uses smoking as instrumental variable to rule out selection bias. Focusing on the role played by individual skills in the return to education, he finds that 20% of the return to schooling are explained by personal aptitudes rather than by education itself.

norability assumes selection on observables, i.e., no unobserved variables may affect the treatment and outcome variables simultaneously.

[8]Sample selection bias refers to the fact that non-observed wage of those who do not participate in the labour market makes the sample non-random.

The problem when estimating the return to years of education is that credentials among educational degrees are not accounted for. Indeed, most scholars consider that it is not years of schooling *per se* which have an impact on the wage premium but the obtaining of a diploma (*sheepskin effect*). Credential models focus on the return to schooling across different education levels by considering discrete measures of schooling such as highest educational achievement (Sheldon, 1992; Suter, 2005; Wolter and Weber, 2005; CSRE, 2010). Based on a cost-benefit analysis, Wolter and Weber (2005) report that private returns to education are not homogeneous along the schooling path. By comparing university and academic matura graduates, they report annual rates of return for university education of 5.4% for men and of 2.2% for women, which are the lowest rates among all education levels. Such findings are supported by CSRE (2010) whose estimation procedure relies on the traditional Mincer wage equation. Using compulsory schooling as reference group, they find that the annual rate of return to university education is just above 7% for men and equal to 6.5% for women. Both analyses, however, cannot pretend to give a causal interpretation to their results given that they do not control directly for self-selection bias. Sheldon (1992) and Suter (2005) control for this endogeneity issue by accounting for the selection process into schooling. Their approach consists in estimating wages and selection equations simultaneously by maximum likelihood. Using data from the Swiss health survey project SOMIPOPS, Sheldon (1992) obtains a rate of return to university education of 23.4% for men (5 years - annual rate: 4.7%). However, the size of the database considered (less than 50 observations) may cast some doubts on the robustness of these findings. Relying on data from the Swiss labour Force Survey (1991 to 2003), Suter (2005) obtains an average return to university education for men of 11.6% (5 years - annual rate: 2.3%) after correcting for selection bias.

All studies mentioned above, however, do not tackle the issue of heterogeneity in returns to schooling among individuals with similar educational attainment. A first attempt applied to the Swiss case is the study from Pereira and Silva-Martins (2004). Considering OECD countries and using a quantile regression framework, they estimate the return to years of education along the wage distribution which serves as proxy for unobserved ability. For Switzerland, the authors report that men in the 9^{th} decile of the conditional wage distribution benefit more from one additional year of schooling than their counterparts in the 1^{st} decile. In other words, inherent ability and education are complements regarding wage increases. Such a result argues in favor of the positive selection hypothesis where well-endowed students benefit most from schooling according to the notion of comparative advantage.

However, their approach does not control for unobserved heterogeneity, an issue that has been accounted for by Balestra and Backes-Gellner (2013). Their study considers an instrumental variable quantile regression (IVQR) framework by using a compulsory education expansion resulting from a major reform in the Swiss education system during the seventies as an instrument for years of education. Their findings report that less able individuals profit most from one additional year of schooling. They also compare the returns between academic and vocational education by including a spline in the traditional Mincer regression and conclude that academic education brings higher returns, especially for individuals in the upper part of the wage distribution. Table 4.1 proposes a short overview of the studies discussed in this section.

4.3 Data

This study uses information gathered in the Swiss Household Panel which is a longitudinal survey ongoing since 1999. For the empirical analysis, I consider the thirteen waves (1999 to 2011) and a set of biographical data[9]. By combining both datasets, I have access to a number of relevant variables such as educational achievement, family background characteristics and labour market variables. The main advantage of the biographical data is the presence of numerous pre-treatment variables observed when the individual was 15 years old (i.e., before university entrance) which are of primary interest when estimating the predicted probability to complete higher education. I take advantage of the longitudinal structure of the data to select the last observation per individual across all the waves. By dropping individuals with missing values, I obtain a sample size of N=961 (N=443 for men and N=518 for women) for the OLS framework and of N=898 (N=403 for men and N=495 for women) for the PSM framework which also uses the biographical data.

The dependent variable is determined by the monthly gross labour income. As is common in the literature, I take the logarithm of this variable for the estimation procedure. The treatment variable is defined by a binary variable which takes the value 0 if the highest educational achievement is an academic matura degree and the value 1 if it is a university degree. Only individuals who are no more in formation are considered to avoid a comparison between individuals who are going to obtain a university degree and those

[9]The biographical data contains information on social origins for all individuals who were personally interviewed in any of the waves since 1999.

Table 4.1: Literature on returns to schooling in Switzerland: a selective review

Studies	**Data**	**Identification strategy**	**Returns to schooling**	
			Men	Women
Years of schooling				
Kugler (1988)	SOMIPOPS 1981/1982	OLS	5.5%	9.1%
		Heckman	6.0%	7.2%
Diekmann and Engelhardt (1995)	SLFS 1991	OLS	7.5%	8.0%
		Heckman		7.5%
Bonjour (1997)	SLFS 1993	OLS	7.0%	8.0%
		Heckman		7.5%
Ferro-Luzzi and Silber (1998)	SLFS 1995	OLS	9.1%	9.0%
Falter and Ferro-Luzzi (2000)	LSE 1996	OLS (public sector)	6.1%	6.0%
		Heckman (public sector)	2.7%	2.2%
		OLS (private sector)	7.2%	7.2%
		Heckman (private sector)	7.2%	7.4%
Pereira and Silva-Martins (2004)	SLFS 1995	QR	$\tau(0.1)$:8.7%	
			$\tau(0.9)$:10.6%	
Suter (2005)	Leu, Burri und Priester (1997)	IV (smoking)	11%	
Balestra and Backes-Gellner (2013)	SLFS 2000-2009	IV (educational reform)	9.9%	
		IVQR	$\tau(0.1)$:18.3%	
			$\tau(0.9)$:-0.4%	
Educational attainment (*credentials*)				
Sheldon (1992)	SOMIPOPS 1981/1982	OLS (university versus matura)	1.5%	
		MLE (university versus matura)	23.4%	
Wolter and Weber (2005)	SLFS 2004	CBA (university versus matura)	5.4%	2.2%
Suter (2005)	SLFS 1991-2003	MLE (university versus matura)	11.6%	
CSRE (2010)	SLFS 2007	OLS (university versus compulsory schooling)	7%	6.5%
Balestra and Backes-Gellner (2013)	SLFS 2000-2009	OLS (academic versus vocational education)	7%	

τ is the quantile index. SOMIPOPS: Soziomedizinisches Indikatorensystem der Population der Schweiz. SLFS: Swiss Labour Force Survey. LSE: Swiss wage structure survey. CBA: cost-benefit analysis.

who have already graduated.

To estimate the predicted probability to complete university education, I use a set of different parental background variables which can be separated into four main categories: financial, human and social capital as well as migration background. The financial capital essentially refers to the family's wealth or income. Although the dataset does not include a measure of parental income, it contains a variable indicating if the individual has suffered from financial problems during her adolescence. The human capital is measured by parental education. Social capital is defined by the social gains resulting from interactions between individuals (Bourdieu, 1977; Coleman, 1988). According to Putnam (2000), we can separate this concept into two subgroups, i.e., the *bonding* social capital (intra-family relations like parental involvement, closeness, or stability) and the *bridging* social capital (social networks outside the family's sphere like parental connections with work colleagues or neighbours). For the former, I control for the family structure with a dummy variable equal to 1 if both parents were living together when the individual was 15. The number of siblings is also accounted for. For the latter, I introduce an index of social stratification related to parents' jobs (i.e., the Treiman prestige scale) which serve as proxy for parental social class. Indeed, parents with high social positions may have strong resources or social networks for helping their offspring to find a job in the labour market. Finally, nationality of both parents and country of birth are accounted for by migration background.

Labour market variables included in the wage equation are experience (number of years spent in paid job), job tenure (change of job during the last year) and activity rate (working time in percentage). A main advantage of the SHP data compared to the traditional SLFS or other datasets lies on the fact that the variable measuring labour market experience accounts for career interruption. Consequently, I do not have to construct a variable capturing potential experience as it is generally the case in the literature. Two additional variables, i.e., marital status (married or not) and having children (yes-no), are used in a specific model accounting for selection into job market. The description of variables and summary statistics are presented in Table 4.2.

Table 4.2: Description of variables and summary statistics

Variables	Description	Matura		University	
		Men	**Women**	**Men**	**Women**
labour market income					
Wage	Monthly gross labour income in logarithm form	8.663	8.230	9.101	8.592
		(0.761)	(0.750)	(0.681)	(0.681)
labour market variables					
Experience	Number of years spent in paid job	25.956	20.933	23.764	18.832
		(12.336)	(10.940)	(11.213)	(10.066)
New job	=1 if individual changed of job during last year	0.138	0.109	0.113	0.126
		(0.346)	(0.312)	(0.317)	(0.333)
Activity rate	Working time (in percentage)	93.136	77.656	93.502	77.568
		(19.111)	(29.396)	(16.896)	(24.690)
Human capital					
Father education	=1 if father is university education graduated	0.169	0.163	0.239	0.318
		(0.376)	(0.370)	(0.427)	(0.466)
Mother education	=1 if mother is university education graduated	0.049	0.026	0.069	0.118
		(0.216)	(0.160)	(0.250)	(0.322)
Financial capital					
Financial problems	=1 if individual experienced financial problems at 15	0.208	0.149	0.134	0.147
		(0.407)	(0.357)	(0.341)	(0.355)
Social capital					
Parents living together	=1 if yes	0.654	0.754	0.711	0.765
		(0.477)	(0.431)	(0.453)	(0.424)
Siblings	Having siblings =1	0.795	0.885	0.829	0.889
		(0.405)	(0.319)	(0.377)	(0.314)
Treiman scale: mother job	Index of mother prestige position (min: 13, max: 78)	39.770	39.952	42.882	44.402
		(7.775)	(7.228)	(7.104)	(8.899)
Treiman scale: father job	Index of mother prestige position (min: 13, max: 78)	45.014	45.039	48.584	50.438
		(11.185)	(11.732)	(11.265)	(12.485)
Migration background					
Father Swiss	=1 if father is Swiss	0.733	0.723	0.721	0.627
		(0.444)	(0.448)	(0.449)	(0.484)
Mother Swiss	=1 if mother is Swiss	0.679	0.714	0.692	0.589
		(0.468)	(0.453)	(0.462)	(0.492)
Country of birth	=1 if Switzerland	0.781	0.770	0.808	0.722
		(0.414)	(0.421)	(0.395)	(0.448)
Other individual covariates					
Age	Age in years	44.238	46.983	47.237	44.167
		(14.757)	(13.201)	(11.860)	(11.184)
Married	=1 if married	0.556	0.608	0.676	0.560
		(0.498)	(0.489)	(0.469)	(0.497)
Children	=1 if yes	0.586	0.698	0.638	0.602
		(0.494)	(0.460)	(0.481)	(0.490)

4.4 Empirical framework

4.4.1 Homogeneous returns to education

I first assume that the return to higher education is homogeneous across individuals. Using a traditional augmented Mincer equation, equation (4.8) is estimated with OLS method:

$$lnY_i = \beta_0 + \beta_1 S_i + \beta_2 L_i + \epsilon_i \tag{4.8}$$

where lnY_i is the logarithm of the monthly gross labour income of individual i, S_i is a dummy variable for university achievement, L_i is a set of labour market variables and ϵ_i is an error term. The coefficient β_1 represents the return to university education supposed to be the same across individuals. However, it is likely that women who would receive low wage in the labour market choose not to work and this sample selection bias may overestimate their returns to schooling. To account for this endogeneity issue, I also consider a sample selection model for women that involves the two following equations:

$$lnY_i = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \eta_i \tag{4.9}$$

where lnY_i is observed if

$$D_i = \vartheta_0 + \vartheta_1 Z_i + \upsilon_i > 0 \tag{4.10}$$

where D_i is a dummy variable indicating if the woman is working or not and Z_i is a set of observed covariates influencing the propensity to work. We assume that $\eta_i \sim N(0, \sigma)$ and $\upsilon_i \sim N(0, 1)$. We can then write the coefficient of correlation between the two residuals as $\rho = \text{corr}(\eta_i, \upsilon_i)$. From equation (4.9), we have then:

$$E[lnY_i|D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + E[\eta_i|D_i = 1] \tag{4.11}$$

Under the joint normality assumption, we have:

$$E[lnY_i|D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \rho[E[\upsilon_i|D_i = 1] \tag{4.12}$$

Finally,

$$E[lnY_i|D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \rho\lambda_i \tag{4.13}$$

where λ_i corresponds to the inverse Mills ratio. If $\rho \neq 0$, it means that the OLS estimation suffers from a sample selection bias[10].

[10] Economists, however, are used to estimate the selectivity effect by focusing on $\lambda(= \rho\sigma)$. Moreover, the stata command "heckman" does not report any direct estimates of ρ.

4.4.2 Heterogeneous returns to education

Stratification-multilevel model (SM-HTE)

Heterogeneous returns to education with regards to parental background are first estimated with a stratification-multilevel method of estimating heterogeneous treatment effects (SM-HTE). The estimation procedure is composed of the following steps[11]:

1. First, I estimate the predicted probability to select into higher education $P_i = P(S_i = 1 \,|X_i)$, i.e., the propensity score to complete higher education for each individual, through a logistic regression. I have then:

$$\frac{P_i}{(1 - P_i)} = \exp(\alpha_0 + \alpha_1 PB_i) \tag{4.14}$$

   where PB_i is a set of parental background characteristics.

2. Then, I obtain balanced propensity score strata where both treated and untreated do not differ significantly in their predicted probabilities to be treated. The implicit idea is to create subpopulations composed of "statistical twins".

3. Next, I estimate the return to schooling within each propensity score stratum by considering three different specifications:

   - *First specification*: schooling estimates are obtained through a direct wage comparison between university and matura graduates within each stratum:

$$lnY_{ip} = \varphi_{0p} + \varphi_{1p} S_{ip} + \eta_{ip} \tag{4.15}$$

     where p corresponds to the propensity score stratum and φ_{1p} represents the return to schooling for individuals belonging to a given propensity score stratum p;

   - *Second specification*: I estimate a OLS wage regression including additional covariates within each stratum:

$$lnY_{ip} = \varphi_{0p} + \varphi_{1p} S_{ip} + \varphi_{2p} L_{ip} + \eta_{ip} \tag{4.16}$$

[11] The stata module "hte" developed by Jann, Brand and Xie (2008) has been used for these analyses.

- *Third specification:* A sample selection model is considered within each propensity score strata for women. The wage equation can be represented as follows:

$$lnY_{ip} = \tau_{0p} + \varphi_{1p}S_{ip} + \varphi_{2p}L_{ip} + \rho_p\lambda_{ip} \quad (4.17)$$

4. Finally, I examine the pattern in rates of return across the propensity score strata by using a variance-weighted least-squares regression where the strata-specific return to schooling is regressed on the propensity score strata rank R:

$$\varphi_{1p} = \phi_0 + \phi_1 R + \phi_p \quad (4.18)$$

   where ϕ_0 corresponds to the predicted value of higher education for individuals in the lowest propensity score strata and ϕ_1 determines the pattern in returns to schooling across propensity score strata. Consequently, this last step allows to determine whether the return to schooling is positively or negatively associated to the propensity score.

Matching-smoothing method (MS-HTE)

Although the stratification-multilevel model is easily interpretable and very intuitive, this approach has two main shortcomings. First, assuming that individuals within the propensity score strata have the same return to schooling (within-group homogeneity) may be questionable. Second, representing the pattern in returns to schooling through a linear form may be restrictive. For these reasons, the matching-smoothing method of estimating heterogeneous treatment effects (MS-HTE) proposed in Xie et al. (2011) and Brand and Simon-Thomas (2012) consists in fitting a nonparametric smoothed curve representing the evolution of the returns to schooling along a continuous representation of the propensity score.

After estimating the propensity score for all individuals (see equation (4.14)), the second step consists in matching treated and untreated persons on the propensity score with a traditional matching estimator. If the region of common support is broad, kernel matching is relevant on the grounds that it uses weighted averages of all untreated units to construct the counterfactual outcome. If not, the traditional nearest neighbour matching (with one or five neighbours) is preferred. The third step consists in representing the differences in outcomes between the matched pairs created (i.e., one-to-one/five matching for nearest neighbour and one-to-multiple matching for kernel) along a continuous representation of the propensity score. Finally, the last step estimates a kernel-weighted local polynomial regression to fit the variation in matched differences as a function of the propensity score. At the end, we obtain the average treatment effect on the treated (ATT):

$$ATT = \frac{1}{n_1} \sum_{i}^{n_i} \left[Y_{i,S_i=1} - \sum_{i(j)}^{i,j} w_{i(j)} Y_{i(j),S_i=0} \right] \tag{4.19}$$

where n_1 is the number of treated units, i is the index over treatment cases, j is the index over control cases, and $w_{i(j)}$ represents the scaled weight that measures the distance between each treated and control unit in the matched pair. In the current study, ATT corresponds to the return to schooling for individuals who completed university[12].

4.5 Results

4.5.1 Homogeneity assumption

The two first columns of Table 4.3 represent the baseline OLS specification (ref. equation (4.8)). Results show that the average return to university education is positive and significant for both men and women but higher for men, which is a common finding in this literature. More precisely, I obtain an annual rate of return of 11.9% for men and 6.2% for females by assuming that the average length of university education is four years[13]. Concerning the covariates, experience, squared experience, new job and activity rate follow the expected sign for both genders. The sample selection model for women is presented in the third column of Table 4.3. The selection equation reports that having children and being married influence negatively and significantly the propensity of women to enter into the labour market. The parameter λ is negative and significant, which indicates that OLS estimates are upward biased. The selection corrected average return to schooling for women is equal to 5.4%.

[12] The objective of the ATE is to evaluate the expected effect on the outcome if individuals were randomly assigned to the treatment while the objective of the ATT is to explicitly evaluate the effects on those for whom the program is actually intended (Grilli and Rampichini, 2011).

[13] SHP data does not allow for distinguishing university education by study programs such as Bachelor (3 years), License (4 years) and Master (5 years). Annual schooling returns estimates are computed by using $(e^{\text{coef}} - 1)/4$ (Halvorsen and Palmquist, 1980).

Table 4.3: Homogeneous returns to university education, OLS estimation

Models	**OLS**		**Sample selection**
Gender	**Men**	**Women**	**Women**
Explanatory variables	**Coefficients**	**Coefficients**	**Coefficients**
Education			
University	0.388***	0.222***	0.196***
	(0.058)	(0.045)	(0.045)
labour market variables			
Experience	0.046***	0.049***	0.042***
	(0.009)	(0.010)	(0.010)
Experience squared	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
New job	-0.225***	-0.244***	-0.254***
	(0.081)	(0.076)	(0.076)
Activity rate	0.014***	0.018***	0.017***
	(0.001)	(0.001)	(0.001)
Constant	7.021***	6.662***	6.972***
	(0.155)	(0.115)	(0.150)
Selection equation			
Age			0.289***
			(0.025)
Age squared			-0.003***
			(0.000)
Children			-0.402***
			(0.088)
Married			-0.342***
			(0.081)
Constant			-5.277***
			(0.543)
Inverse Mills ratio			
lambda			-0.245***
			(0.078)
Adjusted R-squared	0.372	0.497	
N	443	518	1772
Censored observations			1254
Uncensored observations			518

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

4.5.2 Heterogeneity assumption

Stratification-multilevel model (SM-HTE)

I turn now to the stratification-multilevel model (SM-HTE) which consists in estimating the returns to university education within homogeneous sub-populations ranked according to their propensity to attend and complete university education.

Results from logistic regressions (ref. equation (4.14)) are presented in Table 4.4. Surprisingly, parental education does not influence significantly the propensity to complete university education but the parental social class plays an important role for both genders to ensure an intergenerational transmission of socioeconomic status between parents and their offspring. For men, being born in Switzerland and having parents from high social classes increase significantly the probability of access to university education. For women, the variable related to financial problems during adolescence reports a positive and significant sign. This finding suggest that there is a negative selection into schooling given that women suffering from financial difficulties when aged 15 are more prone to enroll into university.

Table 4.4: Predicted probability to complete university education, Logistic regressions

Gender	**Men**	**Women**
Explanatory variables	**Coefficients**	**Coefficients**
Father education	0.540	-0.072
	(0.462)	(0.314)
Mother education	-0.265	0.531
	(0.656)	(0.536)
Financial problems	-0.377	0.738**
	(0.353)	(0.304)
Parents living together	0.285	-0.149
	(0.413)	(0.376)
Treiman scale: mother job	0.029**	0.040***
	(0.018)	(0.014)
Treiman scale: father job	0.024*	0.029***
	(0.013)	(0.009)
Siblings	0.005	0.327
	(0.347)	(0.345)
Father Swiss	-0.242	-0.215
	(0.440)	(0.385)
Mother Swiss	-0.124	-0.730*
	(0.405)	(0.370)
Country of birth	0.768*	0.282
	(0.451)	(0.354)
Age	0.221**	0.171**
	(0.097)	(0.078)
Age squared	-0.002**	-0.002**
	(0.001)	(0.001)
Constant	-7.169***	-5.592**
	(2.517)	(1.879)
LR chi(2)	43.65	72.68
Prob > Chi2	0.000	0.000
Pseudo R-squared	0.070	0.108
N	403	495

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Figure 4.1 relates the frequency distribution of the propensity score for matura and university graduates. Although we can see a relatively good overlap within each group, there are no sufficient observations at the extreme tails of the distributions, especially for men in the lowest part of the distribution. To conduct a reliable statistical analysis, I collapse the propensity score strata in the extreme tails of the respective distributions to ensure

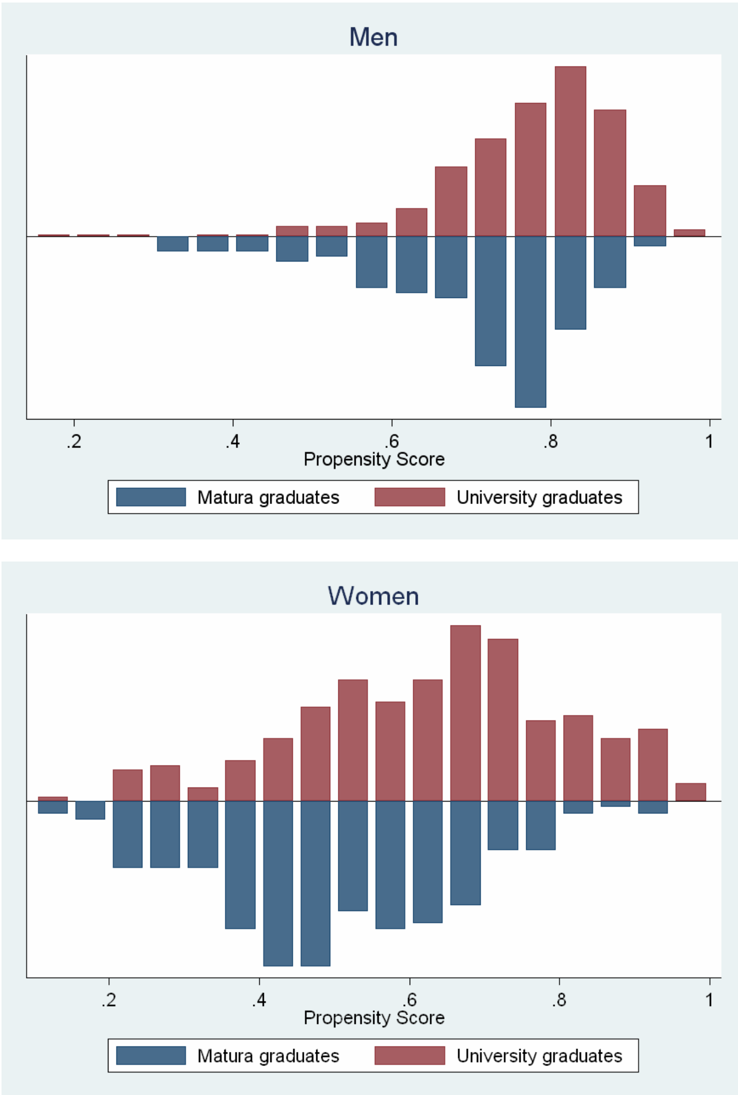


Figure 4.1: Distribution of the propensity score among matura and university graduates

at least 15 treated and untreated individuals within each stratum, which is the rule of thumb in this literature. As a result, I obtain three balanced propensity score strata for each gender[14].

Returns to university education by gender and by propensity score strata are reported in Table 4.5 whose columns (1)-(2) represent the baseline model which relies on wage comparison between university and matura graduates (ref. equation (4.15)), columns (3)-(4) refer to the OLS estimation with labour market variables (ref. equation (4.16)) while columns (5)-(6) report the sample selection model estimates for women (ref. equation (4.17))[15]. Figures 4.2 and 4.3 show the trends in returns to schooling along the propensity score strata for the first two specifications (ref. equation (4.18)). The patterns are reflected by the linear regression line where dots represent point estimates of the return to schooling within each stratum.

Overall, results indicate that low propensity individuals benefit most from university education. However, the slope of the regression line across propensity score strata is only significant for men when labour market variables are considered, which indicates that completing a university degree reinforces the earnings capability of low propensity men when controlling for the number of years spent on the labour market. For women, results lead to a rejection of the heterogeneity assumption in terms of returns to schooling. Indeed, no empirical model reports a positive coefficient for the trend in returns to university education along the propensity score. Using Heckman selection models, the third specification reports that sample selection bias is still an issue, but only for women being in the middle of the propensity score distribution.

Matching-smoothing method (MS-HTE)

The second approach uses a matching-smoothing method (MS-HTE) which consists first in matching individuals on their predicted propensity score and second in estimating the returns to schooling nonparametrically at the matched group level. I consider both nearest-neighbour (one and five controls) and kernel matching algorithms while a kernel-weighted local polynomial regression is used to smooth the variation in matched differences along the propensity score. As mentioned before, the advantage of such a strategy is to relax the linear functional form used in the SM-HTE to detect patterns in returns to schooling and to consider heterogeneity at the matched group

[14]Detailed information is presented in Table 4.8 in the Appendix.
[15]Detailed results are given in Tables 4.9 to 4.11 in the Appendix.

Table 4.5: Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE)

Models	**Stratification-multilevel method (SM)**					
Specifications	**First specification**		**Second specification**		**Third specification**	
	Outcomes' comparison		**OLS**		**Sample selection**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Gender	**Men**	**Women**	**Men**	**Women**	**Women**	
PS strata	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**	**Coefficients**	λ
Strata 1	0.577*	0.307***	0.919**	0.222**	0.201**	-0.230
	(0.343)	(0.130)	(0.270)	(0.104)	(0.101)	(0.179)
Strata 2	0.370***	0.227**	0.488***	0.213***	0.206***	-0.522**
	(0.115)	(0.090)	(0.096)	(0.080)	(0.078)	(0.202)
Strata 3	0.387***	0.206***	0.369***	0.205**	0.212**	-0.551
	(0.111)	(0.096)	(0.098)	(0.088)	(0.086)	(0.358)
Additional covariates	No	No	Yes	Yes	Yes	Yes
Linear trend						
Slope	-0.031	-0.047	-0.193*	-0.008	0.006	
	(0.132)	(0.079)	(0.110)	(0.067)	(0.066)	
Intercept	0.463	0.337*	0.922***	0.230	0.195	
	(0.331)	(0.181)	(0.274)	(0.151)	(0.146)	

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively. Column (5): age, squared age, civil status and having children are considered as covariates in the selection equation.

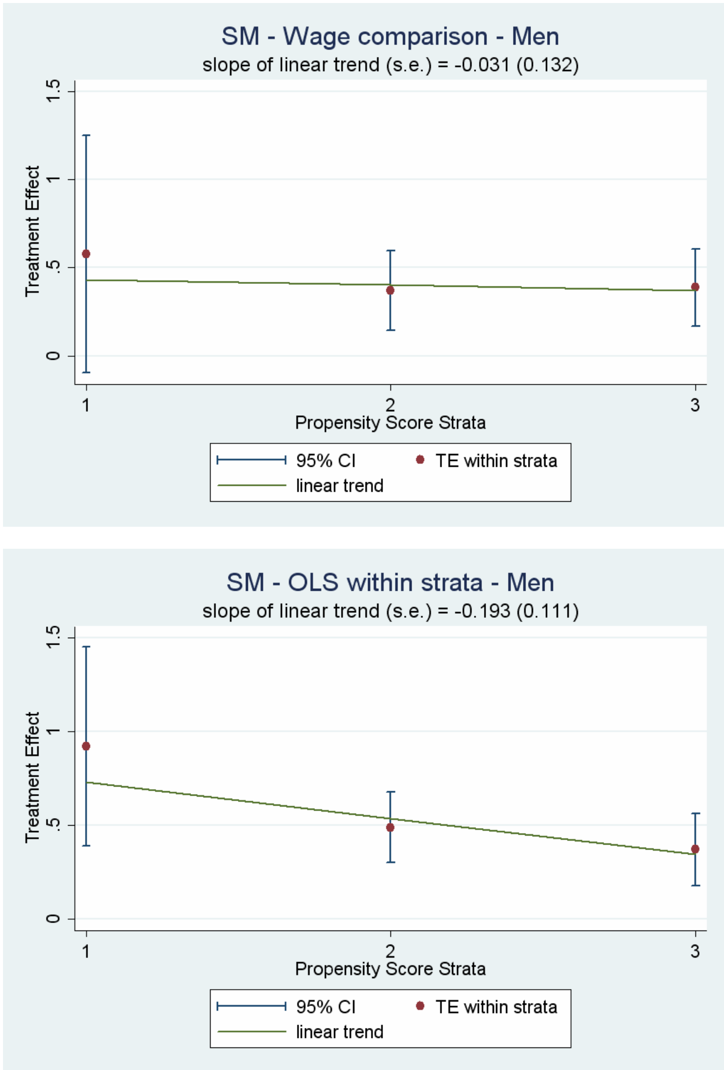


Figure 4.2: Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE), Men

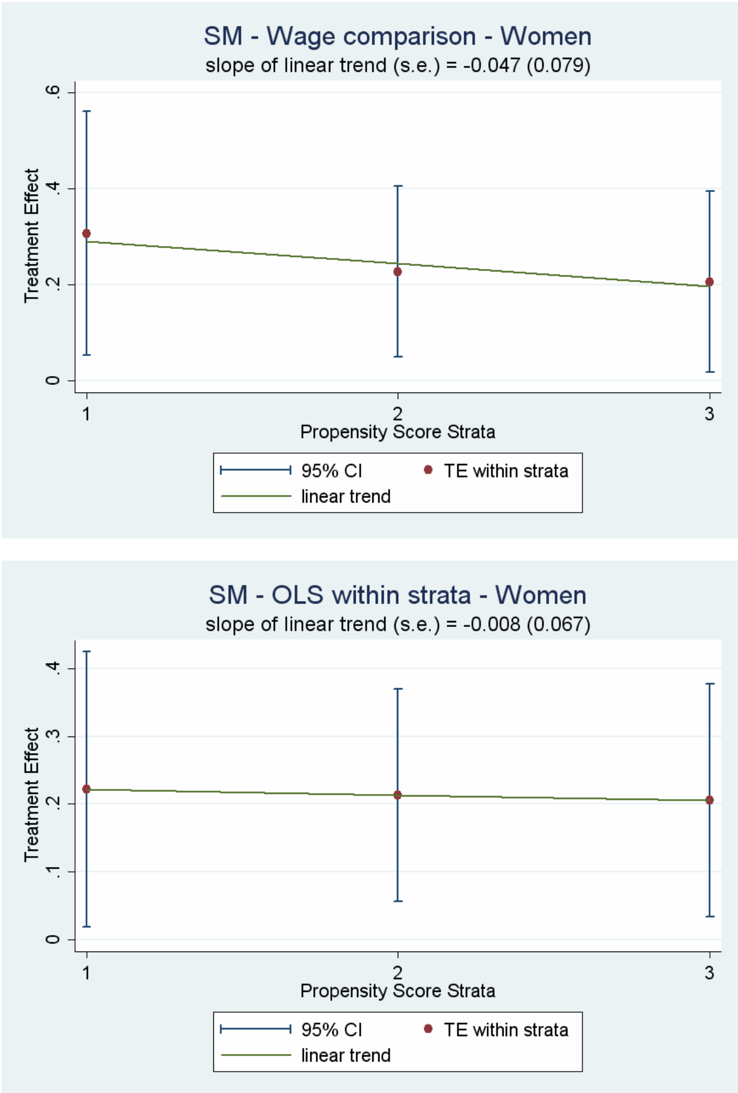


Figure 4.3: Heterogeneous returns to university education, Stratification-multilevel model (SM-HTE), Women

level (instead of assuming homogeneity within propensity score strata). This approach, however, does not allow to proceed to significance tests between the matched pairs.

Table 4.6: Heterogeneous returns to university education, Matching-smoothing estimates

Matching algorithms	**ATT**	
	Men	**Women**
Nearest neighbour (NN), 1 control	0.356***	0.176***
	(0.120)	(0.081)
Nearest neighbour (NN), 5 controls	0.364***	0.179***
	(0.093)	(0.010)
Kernel	0.426***	0.181***
	(0.075)	(0.052)

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.Wages in full-time equivalents.
Local polynomial smoothing (degree 3, bandwith 0.8).

Figure 4.4 illustrates that returns to schooling are rather homogeneous along the propensity score distribution. For men, however, the part of the distribution below 0.6 is hardly interpretable given the very low number of observations. In this context, we can observe a slightly decreasing trend in returns to schooling when considering the part of the propensity score distribution located between 0.6 and 1. In Table 4.6, ATT estimates are equal to 10.7% with NN matching, 11.0% with NN(5) matching and 13.3% with kernel matching, which are very close to the OLS estimate (11.9%)[16]. For women, the graphical analysis advocates clearly in favour of the homogeneity assumption. ATT estimates are 4.8% with NN matching, 4.9% with NN(5) matching and 5% with kernel matching, somewhat below the return to schooling estimated with OLS (6.2%) that is upward biased. Overall, the graphical analyses confirm - to a certain extent - the previous results obtained by considering propensity score strata.

4.6 Auxiliary analysis

This study estimates heterogeneous returns to university education along the propensity score distribution. The implicit objective is to analyze if parental

[16]The comparison between ATT and OLS parameters gives the nature of the selection bias. When ATT is higher than OLS, it means that OLS estimates are downward biased and conversely. Here, however, the comparison is altered given that PSM does not consider labour market variables as it is the case when estimating OLS regressions.

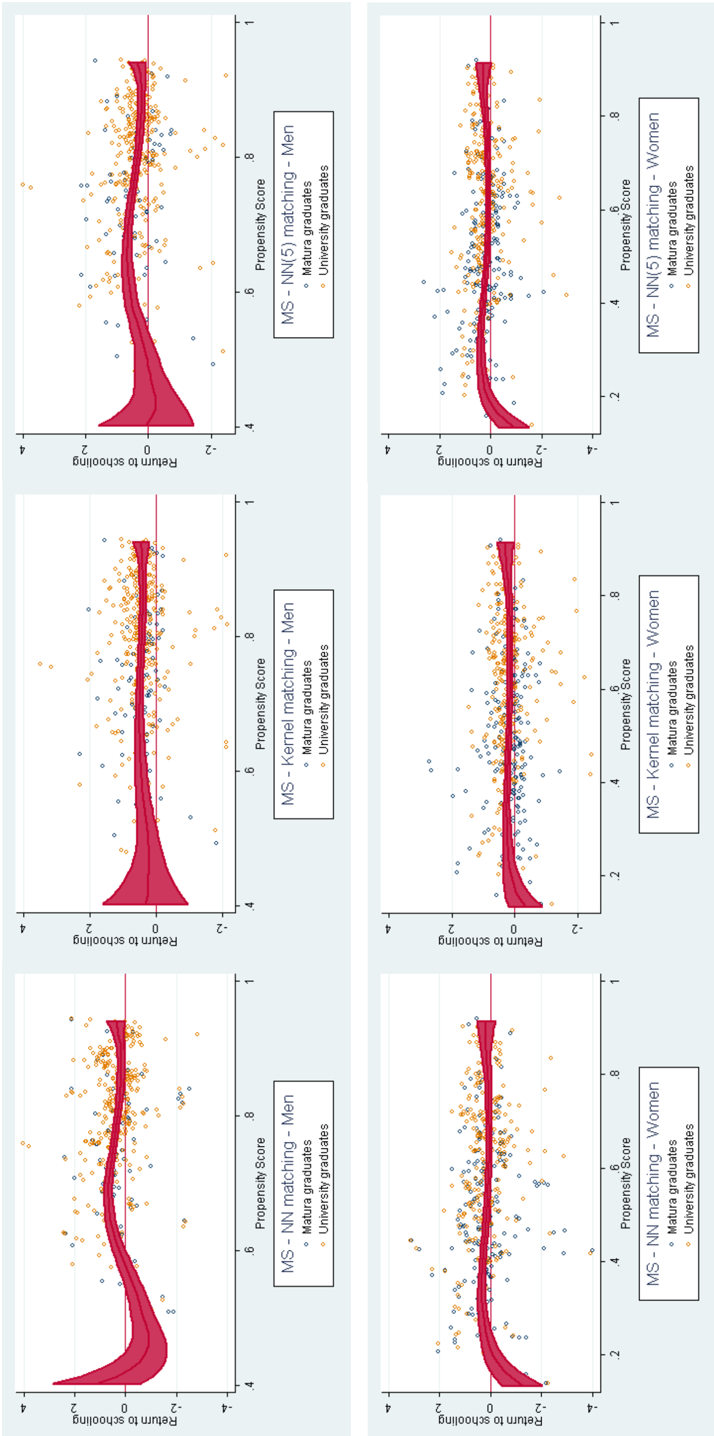


Figure 4.4: Heterogeneous returns to university education, Matching-smoothing model (MS-HTE)

background characteristics and education are complements or substitutes in generating earnings capability. The prior analysis, however, does not account for ability criteria because no variables measuring test score are available in the SHP data. One possibility to overcome this issue is to lead a quantile regression analysis by assuming that the conditional wage distribution reflects unobserved marketable factors which translate directly into higher earnings (Buchinsky, 1998; Arias et al., 2001; Staneva et al., 2010). More precisely, the relative positioning of individuals in the earnings distribution can be related to systematic differences in unobserved marketable attributes such as innate ability, motivation, interpersonal skills, persistence or communication skills. Consequently, the different quantiles represent groups of individuals with similar unobserved inherent abilities. One advantage of this approach is to focus on different types of abilities that may have an impact on earning potential while measures of test scores may be biased through parental background and prior education.

Results presented in Table 4.7 show that men at the bottom of the unobserved ability distribution benefit most from university education. Tests of equal slope coefficients indicate that there are significant differences between the effects of education on earnings along the conditional earnings distribution. More precisely, the schooling coefficient for men in the 15^{th} quantile is significantly higher (at the 5% level) than those obtained by their peers in the 25^{th}, 50^{th} and 75^{th} quantiles. Such a result supports the main finding obtained by Balestra and Backes-Gellner (2013). For women, however, returns to schooling are rather homogeneous along the unobserved ability distribution and coefficients do not differ significantly between the different quantiles. Overall, results suggest that inherent ability and education act as substitutes in generating earnings capability for men while both factors have no specific relation when considering women.

4.7 Conclusion

The objective of this research is to analyze the potential heterogeneity in wage gains after completing a university degree in Switzerland. Indeed, recent microeconometric studies focus on differences in treatment effects among different subgroups in the population on the grounds that the homogeneity assumption is not always appropriate. As valid instruments for education are rather scarce, PSM methods emerge as an interesting approach to account for heterogeneity and self-selection under the conditional independence assumption. In this context, this study relies on two different hierarchical models

Table 4.7: Quantile regressions, Men and Women

Quantile	0.15	0.25	0.50	0.75	0.85
Explanatory variables	**Coef.**	**Coef.**	**Coef.**	**Coef.**	**Coef.**
Men					
University	0.548***	0.470***	0.314***	0.334***	0.375***
	(0.141)	(0.118)	(0.053)	(0.059)	(0.057)
Experience	0.046***	0.038**	0.028***	0.042***	0.047***
	(0.018)	(0.010)	(0.011)	(0.013)	(0.013)
Experience squared	-0.001**	-0.001*	-0.000*	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
New job	-0.605***	-0.336**	-0.258**	-0.117	-0.046
	(0.201)	(0.141)	(0.111)	(0.085)	(0.078)
Activity rate	0.020***	0.017***	0.016***	0.010***	0.007***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Constant	6.095***	6.520***	7.084***	7.708***	8.003***
	(0.238)	(0.180)	(0.181)	(0.249)	(0.232)
Pseudo R-squared	0.339	0.276	0.202	0.132	127
Women					
University	0.180***	0.223***	0.220***	0.234***	0.221***
	(0.068)	(0.060)	(0.038)	(0.041)	(0.052)
Experience	0.085***	0.072***	0.042***	0.031***	0.034***
	(0.019)	(0.023)	(0.023)	(0.007)	(0.009)
Experience squared	-0.002***	-0.001***	-0.001***	-0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
New job	-0.465**	-0.354	-0.173*	-0.145	-0.122
	(0.212)	(0.223)	(0.102)	(0.094)	(0.081)
Activity rate	0.022***	0.020***	0.016***	0.015***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	5.578***	6.040***	6.904***	7.279***	7.534***
	(0.201)	(0.290)	(0.102)	(0.095)	(0.108)
Pseudo R-squared	0.371	0.344	0.334	0.300	0.271

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

based on propensity score matching: a stratification-multilevel model and a smoothing-matching approach.

The former approach suggests that low propensity students benefit most from a university degree. However, the positive association between the propensity score and the returns to education is only significant for men when considering a specification which also controls for labour market variables. This finding suggests that accounting for labour market experience increases the differences in returns to university education across the propensity score strata in favor of low propensity men. For women, the non-significant slope coefficient related to the trend in returns to education along the propensity score strata leads to the rejection of the heterogeneity assumption in terms of returns to schooling.

The latter approach - which plots matched differences in wages between matura and university graduates against a continuous representation of the propensity score - shows that the smoothed curve could also be well approximated by a flat horizontal line for both genders. However, the fitted line for men confirms the results obtained with the previous method if we focus only on the middle and upper parts of the propensity score distribution (i.e., between 0.6 and 1).

An auxiliary analysis based on quantile regressions reveals that men with low unobserved marketable skills also benefit most from university education while inherent ability and education act as two independent factors to determine the level of returns to education in the case of women. In summary, completing a university degree substitutes for a lack of inherent ability - and socioeconomic background - in generating earnings capability for men but has no heterogeneous impact for women.

The main conclusion of this study is that we cannot postulate in favor the comparative advantage hypothesis: individuals with the highest idiosyncratic returns to university education do not benefit most from this degree. Consequently, increasing the incentives for low propensity individuals to attend a university degree seems particularly well-adapted to reinforce both efficiency and equity in university education, especially for men. As Switzerland lacks sufficient tertiary education graduates to boost its economic growth, adapted educational policies should be adopted. In the short run, following OECD recommendations (OECD, 2013), government-sponsored loans to students with disadvantaged family background should be encouraged. Similarly, grant access should be facilitated. In this context, it is worth mentioning that an inter-cantonal agreement aiming at harmonizing grants and loans and increasing support measures entered into force in 2013, ratified by half of the Swiss cantons (OECD, 2013). In the long run, a focus on social policies is also of primary importance to reduce the family background's gap. Indeed, focusing on early childhood is very important given that differences in parental backgrounds have lingering consequences on student's educational path. For that purpose, the recent introduction of the concordat *Harmos* which makes pre-primary education compulsory for all children aged between 4 and 6 and the creation of numerous childcare facilities are very promising steps.

The main caveat of this study is that PSM only controls for observed selection bias and not for "hidden bias". In spite of this limitation, the three empirical models considered allow for a robust interpretation of the results. A second caveat lies in the fact that SHP data lack detailed information to analyze more precisely the self-selection mechanism into higher education. For example, available data suffer from the absence of questions related to the motivations for studying, parental support or ability tests. Future research should then focus on these issues to provide a clearer interpretation of the self-selection mechanisms into university education in Switzerland.

4.8 References

Altonji J., T. Elder and C. Taber, Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of Political Economy*, 113(1), pp. 151-184, 2005.

Arias O., K. Hallock and W. Sosa-Escudero, Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data, *Empirical Economics*, 26(1), pp. 7-40, 2001.

Ashenfelter O. and C. Rousse, Income, schooling, and ability: evidence from a new sample of identical twins, *Quarterly Journal of Economics*, Vol. 113, pp. 253-284, 1998.

Balestra S. and U. Backes-Gellner, Heterogeneous returns to education over the wage distribution: who profits the most?, Leading House Working Paper n°91, 2013.

Bauer P. and R. Riphahn, Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second-generation immigrants, *Journal of Population Economics*, Vol. 20(1), pp. 121-148, 2007.

Becker G., *Human capital: a theoretical and empirical analysis, with special reference to education*, New York: Columbia University Press, 1964.

Behrman T., M. Rosenzweig and P. Taubman, Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment, *Journal of Political Economy*, Vol. 102, pp. 1131-1174, 1994.

Bjorklund, A. and R. Moffitt, The estimation of wage gains and welfare gains in self-selection models, *Review of Economics and Statistics*, Vol. 69, pp. 42-49, 1987.

Bourdieu P., Cultural reproduction and social reproduction, in J. Karabel and A. H. Halsey (eds.), *Power and Ideology in Education*, New York: Oxford University Press, 1977.

Brand J. and Y. Xie, Who benefits most from college? Evidence for negative selection in heterogeneous returns to higher education, *American Sociological Review*, 75(2), 2010.

Brand J. and J. Simon-Thomas, Causal effect heterogeneity, California Center for Population Research, Working paper 2011-012, 2012.

Buchinski M., Recent advances in quantile regression models, *Journal of Human Resources*, Vol. 33, pp. 88-126, 1998.

Card D., Using geographic variation in college proximity to estimate the return to schooling, in L. Christofides, E. Kenneth Grant and R. Swidinsky (eds.), *Aspects of Labour Market behaviour: Essays in Honour of John Vanderkamp*, pp. 201-222, University of Toronto Press, Toronto, Canada, 1995.

Card D., The causal effect of education on earnings, in O. Ashenfelter and D. Card (eds.), Vol. 3A, *Handbook of labour Economics*, Amsterdam: Elsevier, pp. 1801-1863, 1999.

Card D., Estimating the return to schooling: progress on some persistent econometric problems, *Econometrica*, vol. 69, pp. 1127-1160, 2001.

Carneiro P., J. Heckman, and E. Vytlacil, Estimating the return to education when it varies among individuals, unpublished manuscript, Department of Economics, University College London, 2001.

Carneiro P., J. Heckman, and E. Vytlacil, Estimating marginal and average returns to education, unpublished manuscript, Department of Economics, University College London, 2007.

Cattaneo A., S. Hanslin and R. Winkelmann, The Apple falls increasingly far: parent-child correlation in schooling and the growth of post-secondary education in Switzerland, Swiss Journal of Economics and Statistics, 143(2), pp. 133-153, 2007.

Checchi D., *The economics of education: human capital, family background and inequality*, Cambridge University Press, 2006.

Coleman J., Social capital and the creation of human capital, *American Journal of Sociology*, Vol. 94, pp. 95-120, 1988.

Swiss Coordination Centre for Research in Education (CSRE), Swiss education report 2010, Aarau, 2010.

Deschenes O., Estimating the effects of family background on the return to schooling, *Journal of Business and Economic Statistics*, 25(3), 2007.

De Coulon A., Wage differentials between ethnic groups in Swizterland, *Labour*, 15(1), pp. 111-132, 1998.

Diekmann A. and H. Engelhardt, Einkommensungleichheit zwischen Frauen and Männern, Eine ökonometrische Analyse der Schweizer Arbeitskräfteerhebung, *Swiss Journal of Economics and Statistics*, Vol. 1, pp. 57-83, 1995.

Falter J.-M. and Ferro-Luzzi G., Public-private sector wage differentials in Switzerland, Swiss Journal of Economics and Statistics, 136(3), pp. 319-340, 2000.

Falter J.-M., Educational attainment in Switzerland: the role of observable and non-observable family factors, unpublished manuscript, 2005.

Ferro-Luzzi G. and J. Silber, On gender differences in wages in Switzerland, *International Journal of Manpower*, 19(7), pp. 475-485, 1998.

Garen J., The returns to schooling: a selectivity bias approach with a continuous choice variable, *Econometrica*, 52(5), pp. 1199-1218, 1984.

Golder S., Die Arbeitsmarktperformance von Einwanderern in der Schweiz. Eine theorische und empirische Analyse im Rahmen eines internationalen Vergleichs, WWZ-Forschungsbericht, Universtität Basel, 1997.

Griliches Z., Estimating the returns to schooling: some econometric problems, *Econometrica*, Vol. 45, pp. 1-22, 1977.

Grilli L. and C. Rampichini, Propensity scores for the estimation of average treatment effects in observational studies, Training Sessions on Causal Inference, Bristol, June 2011.

Halvorsen R. and R. Palmquist, The interpretation of dummy variables in semilogarithmic equations, *American Economic Review*, 70(3), pp. 474-475, 1980.

Harmon C. and I. Walker, Estimates of the economic return to schooling for the United Kingdom, *American Economic Review*, 85(5), pp. 1278-1286, 1995.

Heckman J., Sample selection bias as a specification error, *Econometrica*, Vol. 47, pp. 153-161, 1979.

Heckman J. and R. Robb, Alternative methods for evaluating the impact of interventions, in J. Heckman and B. Singer (eds.), *Longitudinal Analysis of labour Market Data*, New York: Cambridge University Press, pp. 156-245, 1985.

Heckman, J. and E. Vytlacil, Instrumental variables methods for the correlated random coefficient model: Estimating the rate of return to schooling when the return is correlated with schooling, *Journal of Human Resources*, Vol. 23, pp. 974987, 1998.

Heckman J., J. Tobias and E. Vytlacil, Four parameters of interest in the evaluation of social programs, *Southern Economic Journal*, Vol. 68, pp. 210-223, 2001.

Heckman J., S. Urzua and E. Vytlacil, Understanding instrumental variables in models with essential heterogeneity, *Review of Economics and Statistics*, 88(3), 2006.

Heckman J. and E. Vytlacil, Econometric evaluation of social programs, part II: using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments, in J. Heckman and E. Leamer (eds.), Vol. 6B, *Handbook of Econometrics*, Chapter 71, pp. 4875-5144, Amsterdam: Elsevier, 2007.

Holland P., Statistics and causal inference, American Statistical Association, 81(396), pp. 945-960, 1986.

Jann B., J. Brand and Y. Xie, hte Stata module to perform heterogeneous treatment effect analysis, available from (http://ideas.repec.org/), 2008.

Kane T. and C.Rouse, labour-market returns to two- and four-year college, *American Economic Review*, Vol. 85, pp.600-614, 1995.

Kugler P., Lohndisckriminierung in der Schweiz: Evidenz von Mikrodaten, *Swiss Journal of Economics and Statistics*, N° 1, pp.23-47, 1988.

Lee D., The early socioeconomic effects of teenage childbearing: A propensity score matching approach, *Demographic Research*, Vol. 23, pp. 697-736, 2010.

Leu E., S. Burri and T. Priester, Lebensqualität und Armut in der Schweiz, Bern: Paul Haupt., 1997.

Miller P., C. Mulvey and N. Martin, What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. *American Economic Review*, 85(3), pp. 586-599, 1995.

Mincer J., *Schooling, experience, and earnings*, New York: Columbia University Press, 1974.

Morgan S., *On the edge of commitment: educational attainment and race in the United States*, Stanford, CA: Stanford University Press, 2005.

OECD, OECD Economic Surveys: Switzerland 2009, OECD Publishing, 2009.

OECD, OECD Economic Surveys: Switzerland 2013, OECD Publishing, 2013.

Pereira P. and P. Martins, Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11(3), pp. 355-371, 2004.

Pons E. and M. Gonzalo, Returns to schooling in Spain: how reliable are instrumental variable estimates?, *Labour*, 16(4), pp.747-770, 2002.

Putnam R., *Bowling alone: the collapse and revival of American community*, New York, Simon and Schuster, 2000.

Rantanen M., Returns to education using the Finnish twin data, unpublished manuscript, 2009.

Roberts M., Causality and Experiments, University of Pennsylvania, 2009.

Scheldon G., Selbstselektion und Bildungsrenditen - Ökonometrische Untersuchungen an einem Mikro-Datensatz für die Schweiz, in: Sadowski D. und A. Timmesfeld (eds), Ökonomie und Politik beruflicher Bildung - Europäische Entwicklungen, Duncker und Humblot, Berlin, 1992.

Schultz T., Investment in human capital, *American Economic Review*, 51(1), March 1961.

Smith H. and B. Powell, Great expectations: variations in income expectations among college seniors, *Sociology of Education*, Vol. 63, pp. 194-207, 1990.

Staneva A., G.R. Arabsheibani and P. Murphy, *Returns to education in four transition countries: quantile regression approach*, IZA Discussion Paper N° 5210, 2010.

Suter S., Bildung oder Begabung? Eine ökonometrische Untersuchung zu den Ursachen individueller Lohnunterschiede, PNR FNS 43, 2005.

Tsai S. and Y. Xie, Changes in earnings returns to higher education in Taiwan since the 1990s, *Population Review*, Vol. 47, pp. 1-20, 2008.

Vellacott M.C. and S. Wolter, *Equity in the Swiss education system: dimensions, causes and policy responses*, National report from Switzerland contributing to the OECD's review of "Equity in Education", CSRE, 2004.

Willis R. and S. Rosen, Education and self-Selection, *Journal of Political Economy*, Vol. 87, pp.7-36, 1979.

Willis R., Wage determinants: A survey and reinterpretation of human capital earnings functions, in O. Ashenfelter and R. Layard (eds.), Vol. 1,*Handbook of labour Economics*, Chapter 10, 1987.

Wolter S. and B. Weber, La rentabilité de la formation: un indicateur économique essentiel à l'enseignement, *La Vie Economique*, Octobre 2005.

Xie Y. and X. Wu, Market premium, social process, and statisticism, *American Sociological Review*, Vol. 70, pp. 865-870, 2005.

Xie Y., J. Brand and B. Jann, Estimating heterogeneous treatment effects with observational data, Population Studies Center Research Report 11-729, 2011.

4.9 Appendices

4.9.1 NATE parameter

The mathematical development of this Appendix is taken from Roberts (2009). Let us consider the following regression model where returns to schooling β_i may vary across individuals:

$$Y_i = \alpha_i + \beta_i S_i + u_i \tag{4.20}$$

If I take the conditional expectations of potential wages, I have then:

$$E[Y_i|S_i = 1] = \alpha_i + E[\beta_i|S_i = 1] + E[u_i|S_i = 1] \tag{4.21}$$
$$E[Y_i|S_i = 0] = \alpha_i + E[u_i|S_i = 0] \tag{4.22}$$

By subtracting the above equations from each other, I obtain:

$$\begin{aligned} E[Y_i|S_i = 1] - E[Y_i|S_i = 0] &= \underbrace{E[\beta_i|S_i = 1]}_{ATT} \\ &+ \underbrace{(E[u_i|S_i = 1] - E[u_i|S_i = 0])}_{\text{Selection bias}} \end{aligned} \tag{4.23}$$

where the ATT parameter ($E[\beta_i|S_i = 1]$) is defined as the average return to schooling for individuals who select in university education.

The objective is now to recover the ATE ($E[\beta_i]$) to define the second endogeneity bias, i.e., the return bias. For that purpose, I first decompose the ATE parameter as follows:

$$\begin{aligned} E[\beta_i] &= P(S_i = 0)E(\beta_i|S_i = 0) + P(S_i = 1)E(\beta_i|S_i = 1) \\ &= P(S_i = 0)E(\beta_i|S_i = 0) + (1 - P(S_i = 0))E(\beta_i|S_i = 1) \\ &= P(S_i = 0)[E(\beta_i|S_i = 0) - E(\beta_i|S_i = 1)] + E[\beta_i|S_i = 1] \end{aligned}$$

I have then:

$$E[\beta_i|S_i = 1] = E[\beta_i] - P(S_i = 0)[E(\beta_i|S_i = 0) - E(\beta_i|S_i = 1)] \tag{4.24}$$

By pluging equation (4.24) into equation (4.23), I obtain the naive estimator of the average treatment effect (NATE) which can be written as

$$\begin{aligned} E[Y_i|S_i = 1] - E[Y_i|S_i = 0] &= \underbrace{E[\beta_i]}_{ATE} + P(S_i = 0)\underbrace{[E(\beta_i|S_i = 1) - E(\beta_i|S_i = 0)]}_{\text{Return bias}} \\ &+ \underbrace{(E[u_i|S_i = 1] - E[u_i|S_i = 0])}_{\text{Selection bias}} \end{aligned}$$

4.9.2 Additional tables

Table 4.8: Detailed information on the propensity score strata

Gender	**Men**		**Women**	
Common support				
Min	0.178		0.088	
Max	0.977		0.964	
Propensity score strata	Size of the strata			
Strata 1	[0.0-0.6]		[0.0-0.4]	
Strata 2	[0.6-0.8]		[0.4-0.6]	
Strata 3	[0.8-1.0]		[0.6-1.0]	
Propensity score strata	N within strata			
Strata 1	33		95	
Matura		*15*		*62*
University		*18*		*33*
Strata 2	183		199	
Matura		*50*		*103*
University		*133*		*96*
Strata 3	187		201	
Matura		*21*		*45*
University		*166*		*156*
Total	403		495	

Table 4.9: Stratification-multilevel model, OLS estimation, Men

Propensity score strata	**Strata 1**	**Strata 2**	**Strata 3**
Explanatory variables	**Coefficients**	**Coefficients**	**Coefficients**
Education			
University	0.919***	0.488***	0.369***
	(0.270)	(0.096)	(0.098)
Individual covariates			
Experience	0.016	0.055***	0.040**
	(0.048)	(0.017)	(0.016)
Experience squared	-0.003	-0.001***	-0.001**
	(0.001)	(0.000)	(0.000)
New job	-0.716	-0.266*	-0.372***
	(0.452)	(0.147)	(0.137)
Activity rate	0.008	0.013***	0.018***
	(0.004)	(0.002)	(0.001)
Constant	7.895***	6.939***	6.796***
	(0.835)	(0.294)	(0.249)
Adjusted R-squared	0.395	0.323	0.430
N	29	166	182

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 4.10: Stratification-multilevel model, OLS estimation, Women

Propensity score strata	**Strata 1**	**Strata 2**	**Strata 3**
Explanatory variables	**Coefficients**	**Coefficients**	**Coefficients**
Education			
University	0.222**	0.213**	0.205**
	(0.104)	(0.080)	(0.088)
Individual covariates			
Experience	0.085***	0.067***	0.035*
	(0.022)	(0.023)	(0.019)
Experience squared	-0.001***	-0.001***	-0.001
	(0.000)	(0.000)	(0.001)
New job	0.053	-0.303**	-0.296**
	(0.241)	(0.137)	(0.120)
Activity rate	0.022***	0.017***	0.019***
	(0.002)	(0.001)	(0.002)
Constant	5.934***	6.587***	6.762***
	(0.311)	(0.249)	(0.213)
Adjusted R-squared	0.658	0.475	0.501
N	86	187	178

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Table 4.11: Stratification-multilevel model, Sample selection specification, Women

Propensity score strata	**Strata 1**	**Strata 2**	**Strata 3**
Explanatory variables	**Coefficients**	**Coefficients**	**Coefficients**
Wage equation			
University	0.201**	0.206***	0.212**
	(0.101)	(0.078)	(0.086)
Experience	0.081***	0.048**	0.028
	(0.022)	(0.024)	(0.020)
Experience squared	-0.001*	-0.001	-0.000
	(0.000)	(0.001)	(0.001)
New job	0.049	-0.311**	-0.324***
	(0.232)	(0.134)	(0.122)
Activity rate	0.021***	0.017***	0.018***
	(0.002)	(0.002)	(0.002)
Constant	6.107***	6.881***	7.026***
	(0.332)	(0.272)	(0.274)
Selection equation			
Age	0.372***	0.352***	0.136
	(0.108)	(0.091)	(0.107)
Age squared	-0.004***	-0.004***	-0.002
	(0.001)	(0.001)	(0.001)
Child	-0.560*	-0.030	-0.128
	(0.290)	(0.244)	(0.251)
Married	-0.171	-0.045	-0.467**
	(0.267)	(0.223)	(0.230)
Constant	-5.909**	-5.753***	-1.263
	(2.575)	(2.107)	(2.325)
Inverse Mills ratio			
lambda	-0.230	-0.522**	-0.551
	(0.179)	(0.202)	(0.358)
Prob > Chi2	0.000	0.000	0.000
N	149	244	230
Censored observations	63	57	52
Uncensored observations	86	187	178

*, **and*** indicate a statistical significance at 10%, 5% and 1% level, respectively.

Chapter 5

Conclusions

Education is a key factor when discussing Switzerland's competitiveness at the international level. The quality of Switzerland's research institutions, the numerous on-the-job training opportunities, the high degree of collaboration between the academic and business sectors or the smooth transition from the education system to the labour market are some elements explaining Switzerland's economic success. As Switzerland has a comparative advantage in human capital, it is of primary importance for policy-makers to preserve and ensure efficiency, equity and quality in the education system. This task, however, faces several challenges given that cultural, socioeconomic and behavioural factors (e.g., cultural pluralism, norms, values, health behaviour or social interactions) may have an impact on how the education system works. This dissertation focuses on such issues by analyzing the impact of peer effects (between classmates), risky health behaviours (through cannabis consumption) and parental background (which may affect schooling choice and learning process) on different educational outcomes. The three contributions explore three different but successive levels in post-primary education and rely on sophisticated econometric methods to establish causal, or at least robust, relationships between treatment and outcomes of interest.

The decentralized structure of compulsory education in Switzerland allows to deal with differences in cultures and school traditions which characterize this multilingual country. However, since 2006 and the acceptance of new constitutional articles by Swiss people, some rules and procedures in compulsory education had to be standardized. In this context, the Swiss Conference of Cantonal Ministers of Education (EDK) approved in June 2007 the intercantonal *Harmos* agreement which has for objective to harmonize the school starting age, compulsory school attendance, the duration and objectives of the levels of education, and the transitions between the levels of education. This concordat came into force in April 2009 to be implemented over six years. One consequence of the introduction of the concordat Harmos is the postponement of first tracking at age 13. The implicit idea is to improve equality of opportunity because early tracking is associated with intergenerational immobility of human capital given that parental background plays an important role at early stages of the educational process. However, the harmonization in terms of schooling placement policies - comprehensive versus tracking - was not discussed in the concordat, essentially because cantonal and communal authorities attach great importance to the local and historical characteristics of their own school systems.

The first contribution is integrated in the debate on school organizational design by analyzing if forming mixed-ability classes could enhance both efficiency and equality of opportunity at the lower secondary level. As mentioned before, this question arises on the grounds that ability tracking lacks effectiveness and generates risks for the progression of less-endowed children. The study is positioned in the peer effects literature given that considering the magnitude and nature of social interactions among classmates is an interesting and relevant way to determine to what extent and how pupils are affected by their classmates. The peer group quality is measured through the average parental background characteristics of the student's classmates. As peer group formation is non-random due to selectivity issues, the study introduces ability track fixed effects in the identification strategy. Quantile regressions are estimated to consider non-linearity in peer effects along the conditional ability distribution.

The main findings report that peer coefficients are positive, small but significant in reading and sciences while the peer coefficient in mathematics is not significant. When accounting for non-linear peer effects through quantile regressions, we observe decreasing returns in peer effects in reading and sciences while we witness to increasing returns for mathematics. Results also report that peer heterogeneity does not decrease the school performances within the classroom but reduces the family background effect on school performances, whatever the field considered. To summarize, detracking could lead to equity and efficiency gains in reading and sciences but only to equity gains in mathematics given that students are better off with peers of their own ability level.

The empirical analysis, however, suffers from some important limitations. First, I cannot disentangle between exogenous and endogenous peer effects. While the former refers to spillovers generated by peer background characteristics, the latter results directly from peers' behaviours. Although this distinction is important to draw policy recommendations, it is very difficult in practice to distinguish between both effects. Second, peer group effects are not only subject to unobserved school or class characteristics but also to individual or teacher unobserved characteristics. Due to data limitations, I cannot include individual and teacher fixed effects in the empirical framework. Moreover, I cannot control for time-varying variables given that I work with cross-sectional data. Finally, the discussion on students' reallocation does not account for the additional formation costs related to the reinforcement of teachers' skills to teach in a heterogeneous environment. An interesting point for further research would be to consider a longer perspective, i.e., analyzing the impact of peer effects on school choices and/or labour market outcomes.

The difference between mathematics, on the one hand, and reading and sciences, on the other hand, merits special attention. One potential explanation may lie in the fact that disparities in mathematical skills which occur during primary education can be hardly compensated by peer quality effects at the lower secondary level. For reading and sciences, however, peer quality effects seem to have the potential to raise educational performances for less-able students. Even if peer effects in mathematics might be strengthened, the practical implementation of a mixed-ability system would remain difficult given that cantons share different viewpoints concerning schooling practices and the fact that some efforts in this direction (e.g., postponement of ability tracking, development of mixed-ability classes with level grouping) have recently been done. In any case, detracking would imply a reinforcement of teacher training programs so that teachers can adapt their teaching to heterogeneous classes. This strategy has been applied in Finland - a country which imposes a uniform academic curriculum until the end of lower secondary school - and the excellent learning outcomes obtained in the PISA test scores confirm the effectiveness of the Finnish teacher training programs. In short, my findings suggest that a more in-depth debate on the practice of ability tracking is necessary, as mentioned in the latest OECD report on Switzerland. Further experiments on mixed-ability classes at the national level should be included in the future agenda for educational reforms, associated with a careful examination of teacher training programs to facilitate teaching in a heterogeneous environment.

While the first contribution focuses on parental background and educational peer effects to explain students' test scores, several other factors may influence educational outcomes. For instance, peer pressure or time preferences considerations may have an impact on student's health behaviour whose consequences can also affect schooling outcomes. Focusing on risky health behaviour, the most worrisome issue in Switzerland is related to cannabis consumption whose lifetime consumption rate during adolescence is significantly higher than in any other OECD countries. According to the neuroscience literature, cannabis consumption is associated with deterioration in cognitive functioning and addiction, two issues that can play an important role in the student's educational pathway. However, analyzing the relationship between substance use and educational outcomes face strong endogeneity problems (e.g., selectivity or reverse causality), which explains the difficulty to find strong evidence in this area.

In this setting, the second contribution analyzes the relationship between cannabis use and a set of short-term academic outcomes at the upper secondary level. Six different outcomes are considered: school difficulties, absenteeism, poor grades, lack of engagement, lack of motivation and concentration problems. In terms of cannabis consumption, we make a distinction between any and frequent use in order to investigate if frequency of consumption leads to different results. To obtain consistent estimates, we rely on two identification strategies, i.e., fixed effects and propensity score matching methods.

Both specifications show that frequent cannabis users are more likely to obtain poor grades than non-users and that cannabis consumption has a negative impact on school attendance, namely that it increases the number of school days skipped. An auxiliary analysis based on Rosenbaum bounds determines to which extent our matching estimates may be affected by unobserved heterogeneity. Results confirm that establishing causal relationships in this literature is challenging given that a small departure from the conditional independence assumption may invalidate some of our coefficients.

Some caveats can be mentioned. First, data on cannabis consumption and educational outcomes rely on self-reported measures which are subject to many biases (e.g., lie, poor memory, cognitive bias, perception of the construct). Second, we cannot explain through which channels cannabis use may influence the probability to obtain poor grades. For instance, is it through a reduction of cognitive functioning or due to an increase in school day skipped? Finally, only frequency of consumption is observed whereas a variable capturing the intensity of use would have been of primary interest for this study.

In terms of policy recommendations, a coordination between health and education policies is strongly recommended in order to raise knowledge about the short-term damages of cannabis consumption on some educational outcomes. The main issue, however, lies in the fact that the political messages concerning the harmful effects of cannabis use are unclear and therefore leave the population in uncertainty about the real risks of this drug. The recent liberalizations of cannabis in Colorado (United States) and Uruguay reinforce this feeling. Further academic research in this area should be strengthened to obtain consistent information for the political debate. In parallel, public authorities should look carefully at the effects of cannabis liberalizations which take place abroad in order to benefit from their experiences.

The two first contributions analyze the impact of peer quality and substance use on educational outcomes but in a short-term perspective, i.e., they consider educational outcomes in a strict sense (e.g., test scores, absenteeism, school difficulties). It is then interesting to adopt a longer term perspective and to consider also labour market outcomes. In this context, focusing on parental background is relevant given that inequality of opportunity may affect the student's educational pathway and, therefore, the situation on the labour market. This statement is especially adapted to the Swiss case given that different reports (OECD or Swiss Coordination Center for Research in Education) confirm that students with disadvantaged family backgrounds face important barriers along the entire educational path.

The third essay seeks to analyze if returns to university education vary across individuals with different family background characteristics. The objective is to determine the relationship between university education and parental background in generating earnings capability. For that purpose, the study uses an identification strategy based on a two-step estimation procedure. The first step consists in estimating the predicted probability to complete university education by focusing on different parental background characteristics. The second step estimates returns to schooling across propensity score strata and along the propensity distribution (discrete versus continuous approaches). Both analyses are conducted for men and women to determine if the distributions of returns to university education differ between gender. The nature of the relationship between the propensity score and the return to schooling - positive or negative - allows to test if self-selection into university education is based on the principle of comparative advantage, i.e., if individuals with higher expected returns to university education benefit most from this degree, or if graduation leads to higher returns for students with disadvantaged background attributes.

Results from both OLS and PSM methods report higher returns for men than women. Such findings confirm the gender wage gap underlined in prior literature, although we witness during the last years a reduction in the gender gap in Switzerland regarding educational outcomes (e.g., equal expected years of schooling and same enrollment rates at university). Focusing on the impact of family background on returns to schooling, empirical evidence reports a negative association between the predicted probability to complete a university education and the returns associated with completing this degree. This negative relationship, however, is only significant for men when labour market experience is accounted for, indicating that the number of years spent on the labour market reinforces more strongly the individual earnings capability of low propensity men. For women, the absence of a significant relation between the two variables of interest means that the impact of university education on wages along the propensity score distribution is homogeneous. Both results, however, converge to the rejection of positive wage sorting into schooling.

This study entails two main limitations. First, propensity score matching models rely on the untestable conditional independence assumption. However, the convergence of the results, based on three empirical models, argues in favour of a robust interpretation of the coefficients. Second, some important variables not available in SHP data - e.g., ability measures or questions related to the motivation for studying or school choice - could bring further insights when discussing self-selection mechanisms. Two different aspects, not explored in the current study, could be of interest for further investigations. First, analyzing some non-monetary outcomes, e.g., social or health outcomes, can broaden the scope of the debate on this topic. Second, the current study focuses only on matura and university graduates. Other educational attainments could be of great interest. For example, a comparison within universities (between bachelor, licence or master students) or between universities and other tertiary education levels could bring relevant information in the discussion.

The main findings of this study allow to address policy advices for the gender wage gap as well as for the role played by family background on labour market outcomes. For the former, we acknowledge that women are still overrepresented in part-time work and underrepresented in managing or leading positions. To remedy this issue, as recommended by the OECD, questions related to childcare costs, tax rates on second earners or cultural hurdles in the society need to be (again) addressed by political bodies. For the latter, results suggest that facilitating access to university education for low propensity individuals would not only improve equality of opportunity but also would lead to efficiency gains given that these students obtain the highest returns to schooling. This statement indicates that educational policies should be reinforced or extended in different ways. First, facilitating access to grants or loans for students with disadvantaged family backgrounds is of primary importance to encourage them to pursue studies without having the obligation to find an employment alongside their studies. In this context, the intercantonal agreement on grants which entered in force in 2013 has for objective to reduce the unequal prospects to obtain such financial resources. Second, reducing the impact of parental background on education performances implies policy interventions since early childhood. The most significant measure taken during the recent years comes from the concordat *Harmos* which established a compulsory pre-school program in most Swiss cantons. This policy can contribute significantly to raise educational outcomes for children, especially for those with low socioeconomic parental backgrounds. The recent increase in childcare provision through the development of a national system of accreditation and the federal support attributed to special-needs children go in the same direction. An alternative solution would consist in introducing a national voucher scheme. A school voucher is a funding certificate from the government which would allow the parents to compensate for their schoolchild's tuition fees. Proponents of such a system consider that the benefits go directly to the concerned persons while opponents consider that introducing these educational vouchers would threaten the public education funding. Until now, no specific voucher scheme has been considered in Switzerland.

More generally, the main results of this doctoral thesis can be commented as follows. First, equality of opportunity is not necessarily in contradiction with efficiency gains. The first and third contributions show that increasing equity at school, for example through the adoption of a comprehensive school system at the lower secondary level or the introduction of measures which aim at facilitating access to university education (e.g., loans or grants), could lead to some improvements in terms of academic performances or return to education investment. As equality of opportunity is a topic of current debate at the international level, contributions in this area of research are particularly welcome. This issue is even more relevant for the Swiss case given that background characteristics still play a prominent role in explaining school choice, school performances, educational achievement and labour market outcomes.

Second, substance use has negative effects on short-term educational outcomes such as absenteeism or the probability of having poor grades at school, which is a relevant finding given the current debates regarding cannabis liberalization in Switzerland. The results advocate for in-depth reflections about cannabis use and the potential risks associated with the trivialization of this product. Further research in this field is crucial to determine if cannabis consumption affects educational outcomes in a causal way, as argued in the corresponding contribution, or if if both factors are commonly affected by other unobserved variables.

Last but not least, the empirical results of this dissertation confirm the importance of controlling for unobserved heterogeneity. The two first contributions illustrate perfectly this issue. In the first one, the magnitude of the average peer effect is significantly reduced - and for some specifications even no more significant - once accounting for ability track fixed effects. In the second one, results show significant differences between the OLS and FE estimation. An auxiliary analysis based on Rosenbaum bounds also reports that the propensity score matching coefficients are very sensitive to unobserved heterogeneity and that even a small departure from the conditional independence assumption may invalidate some of the results.

To conclude, it is worth mentioning that Switzerland has operated a lot of changes during the past years in terms of educational policies, e.g., adoption of partial tracking policies, postponement of the age of first tracking, recognition of diplomas across the country, harmonization of the duration of compulsory education or enforcement of the intercantonal agreement on grants. In my opinion, all these new policies address key issues to strengthen equity and efficiency in the Swiss school system and to preserve or improve the comparative advantage of Switzerland in the field of education. As mentioned previously, however, the society through norms, behaviours or expectations may affect the schooling system in different ways. Consequently, it is of primary importance for academic research to account carefully for the socioeconomic context around the school environment when discussing or drawing policy recommendations.