Who Benefits Most from University Education in Switzerland?

Lionel Perini[a]

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 labor 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.

---

a   University of Neuchâtel, Institute of Economic Research, Pierre-à-Mazel 7, CH-2000 Neuchâtel, Switzerland. Email: lionel.perini@unine.ch.
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 proposes other acceptance criteria (e.g., students with a professional matura are allowed to enter directly in these institutions).

Contrary to the traditional formulation of human capital where returns to schooling are implicitly assumed to be homogenous, 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, 2001). In other words, youths self-select into schooling on the basis of the principle of comparative advantage (Willis and Rosen, 1979; Heckman, Urzua, and Vytlacil, 2006; Carneiro, Heckman, and Vytlacil, 2007), 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 labor market opportunities may have stronger economic incentives to invest in 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 labor market positions even in the absence of higher education. Expressed differently, completing a university degree may augment more the labor 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

---

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).

transmission of educational attainment in Switzerland whose results suggest that schooling choice is essentially determined by parental education and family income (Falter, 2004; Bauer and Riphahn, 2007; Cattaneo, Hanslin, and Winkelmann, 2007).

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, Brand, and Jann (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 Labor 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.

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 labor 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 line 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 homogenous 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 earning 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

---

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.

when labor 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 subsitutes in the human capital accumulation. However, as the estimation procedure is based on selection on observables, results must be interpreted in terms of associations and not in a causal way.

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. The section six presents an auxiliary analysis based on a quantile regression framework. The last section is devoted to the discussion.

2. Theoretical Background and Literature Review

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:

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

where the two right-hand side terms of equation (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):

$$ATE = E[\beta_i] = E[Y_i \mid S_i = 1] - E[Y_i \mid S_i = 0] \tag{2}$$

Let us consider the following wage regression:

$$Y_i = \alpha_i + \beta_i S_i + u_i \tag{3}$$

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:

$$E[Y_i \mid S_i = 1] = \alpha_i + E[\beta_i \mid S_i = 1] + E[u_i \mid S_i = 1] \tag{4}$$

$$E[Y_i \mid S_i = 0] = \alpha_i + E[u_i \mid S_i = 0] \tag{5}$$

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]

$$NATE = E[Y_i \mid S_i = 1] - E[Y_i \mid S_i = 0]$$
$$= \underbrace{E[\beta_i]}_{ATE} + \underbrace{E[u_i \mid S_i = 1] - E[u_i \mid S_i = 0]}_{\text{Selection bias}}$$
$$+ P(S_i = 0) \underbrace{\{ E[\beta_i \mid S_i = 1] - E[\beta_i \mid S_i = 0] \}}_{\text{Return bias}} \tag{6}$$

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 (MILLER, MULVEY, and MART, 1995; ASHENFELTER and ROUSE, 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 IV or fixed effects (FE) methods.[5] In this context, the traditional approach

---

4 See the full mathematical development of this formula (ROBERTS, 2009) in the Appendix.
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, HECKMAN, and VYTLACIL, 2001; BASU et al., 2007). FE methods control for pre-treatment heterogeneity but not for treatment effect heterogeneity (XIE, BRAND, and JANN, 2011).

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; Deschenes, 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 or quasi-natural experiments, propensity score matching (PSM) techniques emerge as the most interesting identification strategy when assuming selection on observables. Results, however, cannot be longer considered as causal. By assuming conditional independence[7], the return to schooling can be formulated as follows:

$$E[\beta_i] = E[Y_i \mid S_i = 1] - E[Y_i \mid S_i = 0] \perp S_i \mid P(S_i = 1 \mid X_i) \qquad (7)$$

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) as well as Xie, Brand, and Jann (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.

---

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, 2007, 2011) 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 ignorability assumes selection on observables, i.e., no unobserved variables may affect the treatment and outcome variables simultaneously.

2.2 Returns to Schooling in Switzerland: A Short Literature Review

Numerous studies have estimated the average return to years of education in Switzerland by focusing on wage discrimination. Some of them analyzed wages differentials by gender (Kugler, 1988; Diekmann and Engelhardt, 1995; Ferro-Luzzi and Silber, 1998) while some others were interested in discrimination towards foreign workers (Golder, 1997; De Coulon, 1998). There is no clear consensus on the magnitude of the return to schooling but the results suggest that men obtain higher returns than women and that wage disparities vary strongly across the immigrant group considered (e.g., country of origin, first or second generation of immigrants). These studies, however, rely on the traditional OLS method to estimate the return to schooling, most of them correcting only for sample selection bias (Kugler, 1998; Dieckmann and Engelhardt, 1995; De Coulon, 1998).[8] One exception in this literature is the study of Suter (2006) 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.

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 obtention 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; Wolter and Weber, 2005; Suter, 2006; CSRE, 2010). Based on a cost-benefit analysis, Wolter and Weber (2005) report that private returns to education are not homogenous 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 of 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 (2006) control for this endogeneity issue by accounting

---

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

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 Labor Force Survey (1991 to 2003), Suter (2006) obtains an average return to university education for men of 11.6% (5 years – annual rate: 2.3%) after correcting for selection bias. However, these studies do not tackle the issue of heterogeneity in returns to schooling across university graduates.

A first attempt to account for heterogeneity in returns to schooling across individuals for Switzerland is 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 9th decile of the conditional wage distribution benefit more from education than their counterparts in the 1st 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 principle of comparative advantage. However, their approach does not use a discrete measure of education and does not account for selectivity biases.

The current study focuses simultaneously on both heterogeneity and self-selection issues by estimating hierarchical models based on propensity score matching (PSM). While the self-selection process is accounted for when estimating the propensity score to complete university education, the empirical model also estimates the trend in returns to schooling along the propensity score to determine who benefits most from this degree.

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 labor market variables. The

---

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

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.[10] 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 labor 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 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 collegues 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 labor

---

10   It is worth to point out that biographical data are based on retrospective information and may be affected by measurement error.

market. Finally, nationality of both parents and country of birth are accounted for by migration background.

Labor 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 labor 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 1.

Table 1: Description of Variables and Summary Statistics

Variables	Description	Matura		University	
		Men	Women	Men	Women
Labor market income					
Wage	Monthy gross labor income in logarithm form	8.663 (0.761)	8.230 (0.750)	9.101 (0.681)	8.592 (0.681)
Labor market variables					
Experience	Number of years spent in paid job	25.956 (12.336)	20.933 (10.940)	23.764 (11.213)	18.832 (10.066)
New job	=1 if individual changed of job during last year	0.138 (0.346)	0.109 (0.312)	0.113 (0.317)	0.126 (0.333)
Activity rate	Working time (in percentage)	93.136 (19.111)	77.656 (29.396)	93.502 (16.896)	77.568 (24.690)
Human capital					
Father education	=1 if father is university education graduated	0.169 (0.376)	0.163 (0.370)	0.239 (0.427)	0.318 (0.466)
Mother education	=1 if mother is university education graduated	0.049 (0.216)	0.026 (0.160)	0.069 (0.250)	0.118 (0.322)
Financial capital					
Financial problems	=1 if individual experienced financial problems at 15	0.208 (0.407)	0.149 (0.357)	0.134 (0.341)	0.147 (0.355)

Variables	Description	Matura		University	
		Men	Women	Men	Women
Social capital					
Parents living together	=1 if yes	0.654 (0.477)	0.754 (0.431)	0.711 (0.453)	0.765 (0.424)
Siblings	Having siblings =1	0.795 (0.405)	0.885 (0.319)	0.829 (0.377)	0.889 (0.314)
Treiman scale: mother job	Index of mother prestige position (min: 13, max: 78)	39.770 (7.775)	39.952 (7.228)	42.882 (7.104)	44.402 (8.899)
Treiman scale: father job	Index of mother prestige position (min: 13, max: 78)	45.014 (11.185)	45.039 (11.732)	48.584 (11.265)	50.438 (12.485)
Migration background					
Father Swiss	=1 if father is Swiss	0.733 (0.444)	0.723 (0.448)	0.721 (0.449)	0.627 (0.484)
Mother Swiss	=1 if mother is Swiss	0.679 (0.468)	0.714 (0.453)	0.692 (0.462)	0.589 (0.492)
Country of birth	=1 if Switzerland	0.781 (0.414)	0.770 (0.421)	0.808 (0.395)	0.722 (0.448)
Other individual covariates					
Age	Age in years	44.238 (14.757)	46.983 (13.201)	47.237 (11.860)	44.167 (11.184)
Married	=1 if married	0.556 (0.498)	0.608 (0.489)	0.676 (0.469)	0.560 (0.497)
Children	=1 if yes	0.586 (0.494)	0.698 (0.460)	0.638 (0.481)	0.602 (0.490)

4. Empirical Framework

4.1 Homogenous Returns to Education

I first assume that the return to higher education is homogenous across individuals. Using a traditional augmented Mincer equation, equation (8) is estimated with OLS method:

$$\ln Y_i = \beta_0 + \beta_1 S_i + \beta_2 L_i + \varepsilon_i \tag{8}$$

where $\ln Y_i$ is the logarithm of the monthly gross labor income of individual i, S_i is a dummy variable for university achievement, L_i is a set of labor 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 labor 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:

$$\ln Y_i = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \eta_i \tag{9}$$

where $\ln Y_i$ is observed if

$$D_i = \vartheta_0 + \vartheta_1 Z_i + \upsilon_i > 0 \tag{10}$$

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 $\rho = corr(\eta_i,\upsilon_i)$. From equation (9), we then have:

$$E[\ln Y_i \mid D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + E[\eta_i \mid D_i = 1] \tag{11}$$

Under the joint normality assumption, we have:

$$E[\ln Y_i \mid D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \rho[E[\upsilon_i \mid D_i = 1] \tag{12}$$

Finally,

$$E[\ln Y_i \mid D_i = 1] = \varphi_0 + \varphi_1 S_i + \varphi_2 L_i + \rho \lambda_i \tag{13}$$

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.[11]

4.2 Heterogeneous Returns to Education

4.2.1 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:[12]

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:

$$\frac{P_i}{(1 - P_i)} = \exp(\alpha_0 + \alpha_1 PB_i) \tag{14}$$

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:

$$\ln Y_{ip} = \varphi_{0p} + \varphi_{1p} S_{ip} + \eta_{ip} \tag{15}$$

---

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

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:

$$\ln Y_{ip} = \varphi_{0p} + \varphi_{1p} S_{ip} + \varphi_{2p} L_{ip} + \eta_{ip} \tag{16}$$

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

$$\ln Y_{ip} = \tau_{0p} + \varphi_{1p} S_{ip} + \varphi_{2p} L_{ip} + \rho_p \lambda_{ip} \tag{17}$$

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:

$$\varphi_{1p} = \phi_0 + \phi_1 R + \phi_p \tag{18}$$

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.

4.2.2 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, Brand, and Jann (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 (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):

$$ATT = \frac{1}{n_1} \sum_{i}^{n_i} \left(Y_{i,S_i=1} - \sum_{i(j)}^{i,j} w_{i(j)} Y_{i(j),S_i=0} \right) \tag{19}$$

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.[13]

5. Results

5.1 Homogeneity Assumption

The two first columns of Table 2 represent the baseline OLS specification (ref. equation (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.[14] Concerning the covariates,

---

13 The objective of the ATE is to evaluate what is 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 programme is actually intended (GRILLI and RAMPICHINI, 2011).

14 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 $(e^{coef} - 1)/4$ (HALVORSEN and PALMQUIST, 1980).

Table 2: Homogenous Returns to University Education, OLS Estimation

Models	OLS		Sample selection
Gender	Men	Women	Women
Explanatory variables		Coefficients	
Education			
University	0.388***	0.222***	0.196***
	(0.058)	(0.045)	(0.045)
Labor 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

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

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 2. The selection equation reports that having children and being married influence negatively and significantly the propensity of women to enter into the labor 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 of 5.4%.

5.2 Heterogeneity Assumption

5.2.1 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 homogenous subpopulations ranked according to their propensity to attend and complete university education.

Results from logistic regressions (ref. equation (14)) are presented in Table 3. 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. Tables 7 and 8 in the Appendix report the means for the covariates by treatment and control group, before and after matching.

Figure 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 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.[15]

---

15 Detailed information is presented in Table 9 in the Appendix.

Table 3: Predicted Probability to Complete University Education, Logistic Regressions

Gender	Men	Women
Explanatory variables	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

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

Figure 1: Distribution of the Propensity Score among Matura and University Graduates
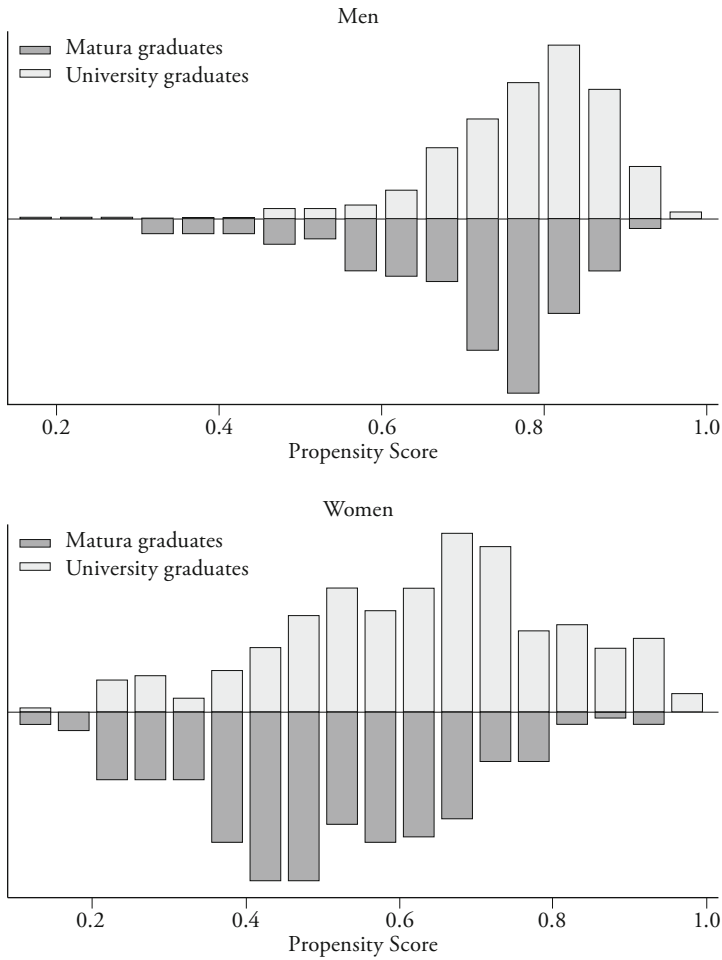
Table 4: Heterogeneous Returns to University Education, Stratification-Multilevel Model (SM-HTE)
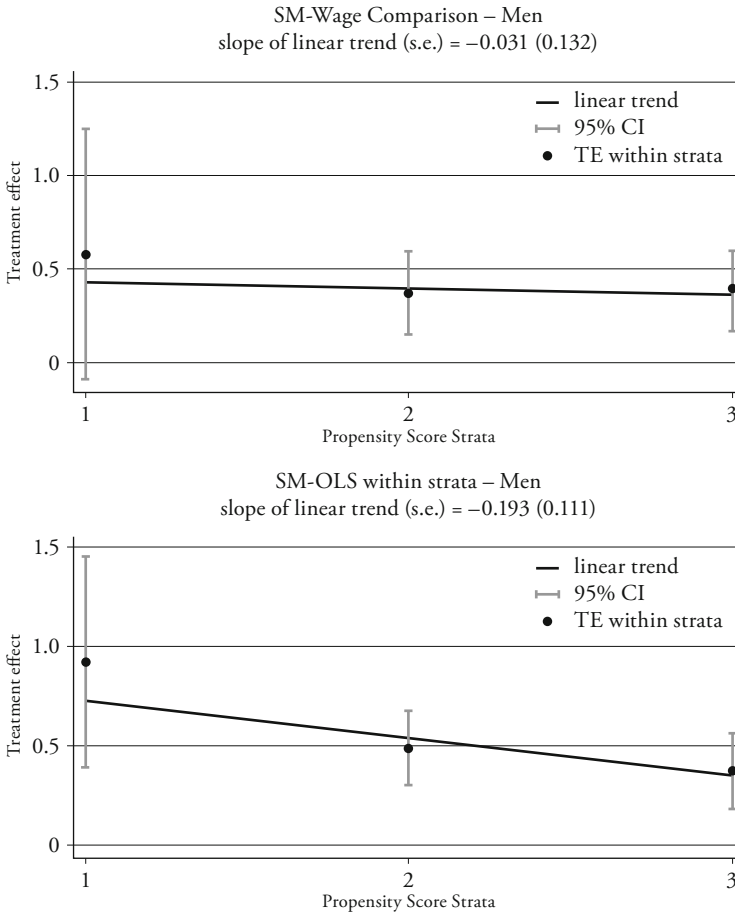
Models	Stratification-multilevel method (SM)					
Specifications	First specification		Second specification		Third specification	
	Outcomes' comparison		OLS		Sample selection	
	(1)	(2)	(3)	(4)	(5)	(6)
Gender	Men	Women	Men	Women	Women	
PS strata	Coefficients					λ
Strata 1	0.577*	0.307***	0.919**	0.222**	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)	

Note: *, **, 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.

Returns to university education by gender and by propensity score strata are reported in Table 4 whose columns (1)–(2) represent the baseline model which relies on wage comparison between university and matura graduates (ref. equation (15)), columns (3)–(4) refer to the OLS estimation with labor market variables (ref. equation (16)) while column (5)–(6) report the sample selection model estimates for women (ref. equation (17)).[16] Figures 2 and 3 show the trends in returns to schooling along the propensity score strata for the first

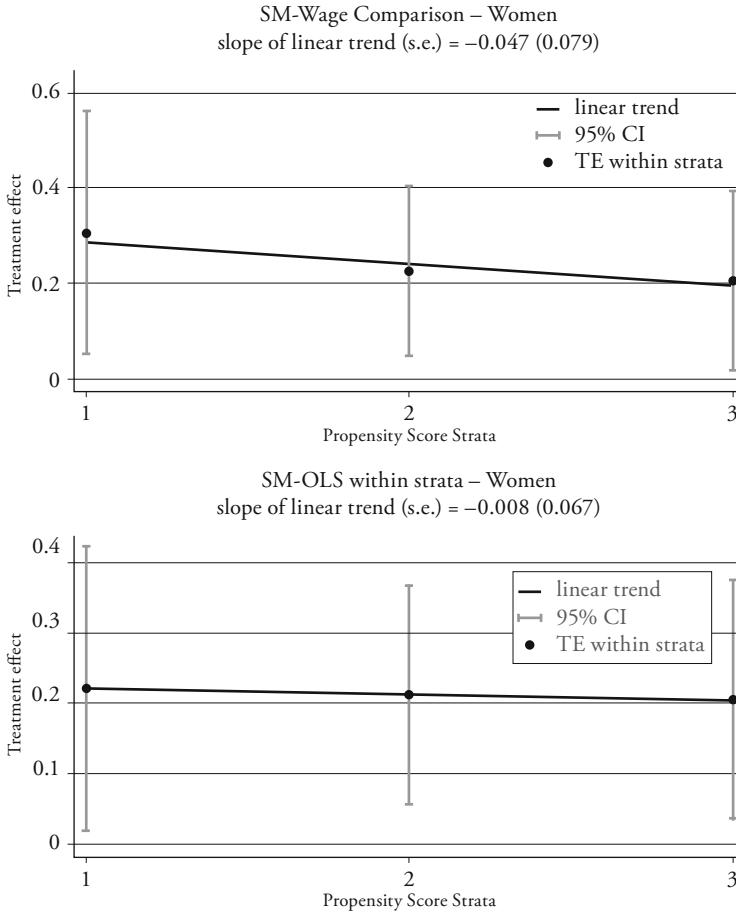16   Detailed results are given in Tables 10 to 12 in the Appendix.

**Figure 2: Heterogeneous Returns to University Education,
Stratification-Multilevel Model (SM-HTE), Men**

SM-Wage Comparison – Men
slope of linear trend (s.e.) = −0.031 (0.132)

SM-OLS within strata – Men
slope of linear trend (s.e.) = −0.193 (0.111)

two specifications (ref. equation (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 labor market variables are considered, which indicates that completing a university degree reinforces the earnings

Figure 3: Heterogeneous Returns to University Education, Stratification-Multilevel Model (SM-HTE), Women

SM-Wage Comparison – Women
slope of linear trend (s.e.) = −0.047 (0.079)

SM-OLS within strata – Women
slope of linear trend (s.e.) = −0.008 (0.067)

capability of low propensity men when controlling for the number of years spent on the labor 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.

5.2.2 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 to estimate the returns to schooling nonparametrically at the matched group level. I consider both nearest-neighbor (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 level (instead of assuming homogeneity within propensity score strata). This approach, however, does not allow to proceed to significance tests between the matched pairs.

Figure 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 difficulty 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 5, ATT estimates are of 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%).[17] For women, the graphical analysis prones clearly in favor of the homogeneity assumption. In Table 5, ATT estimates are of 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.

6. Auxiliary Analysis

This study estimates heterogeneous returns to university education along the propensity score distribution. The implicit objective is to analyze if parental background characteristics and education are complements or subsitutes in generating earning capability. The prior analysis, however, does not account for ability criteria because no variables measuring test score are available in the SHP data.

---

17 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 labor market variables as it is the case when estimating OLS regressions.

**Figure 4: Heterogeneous Returns to University Education,
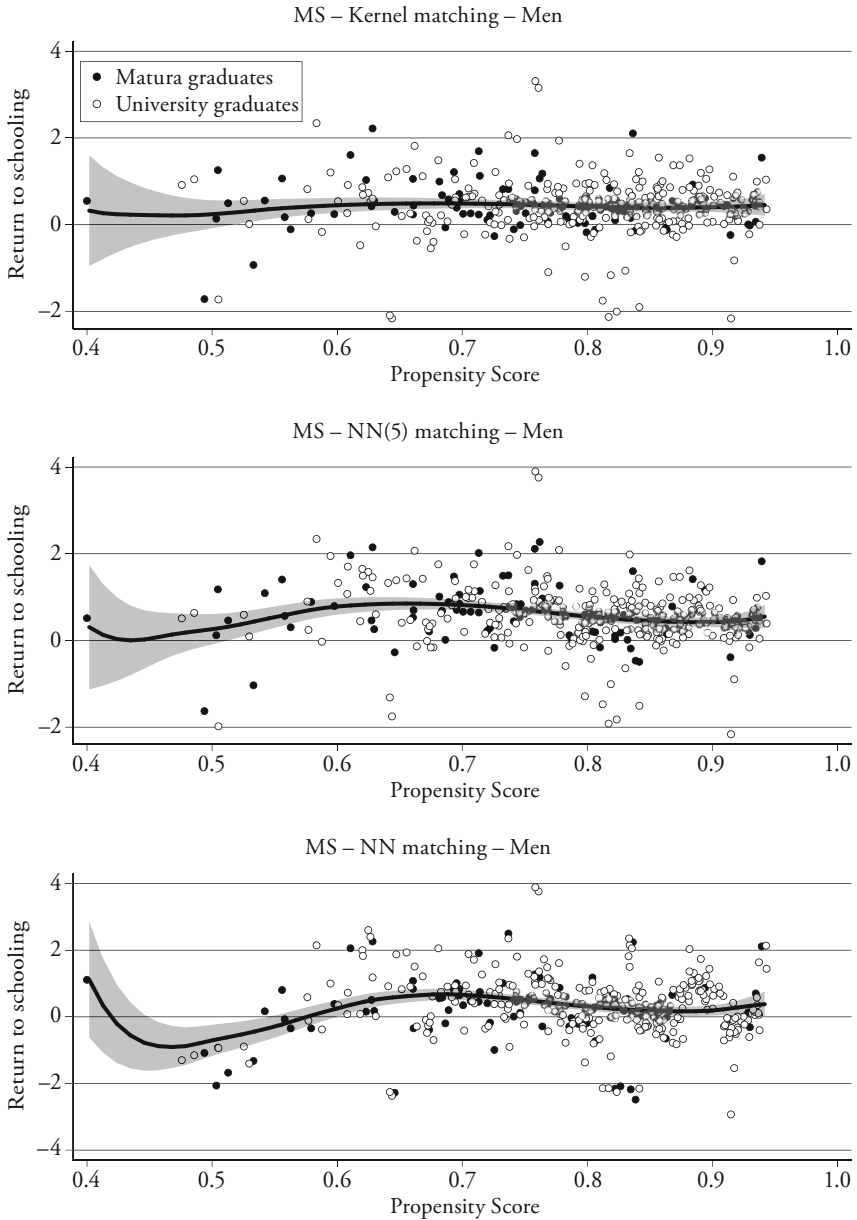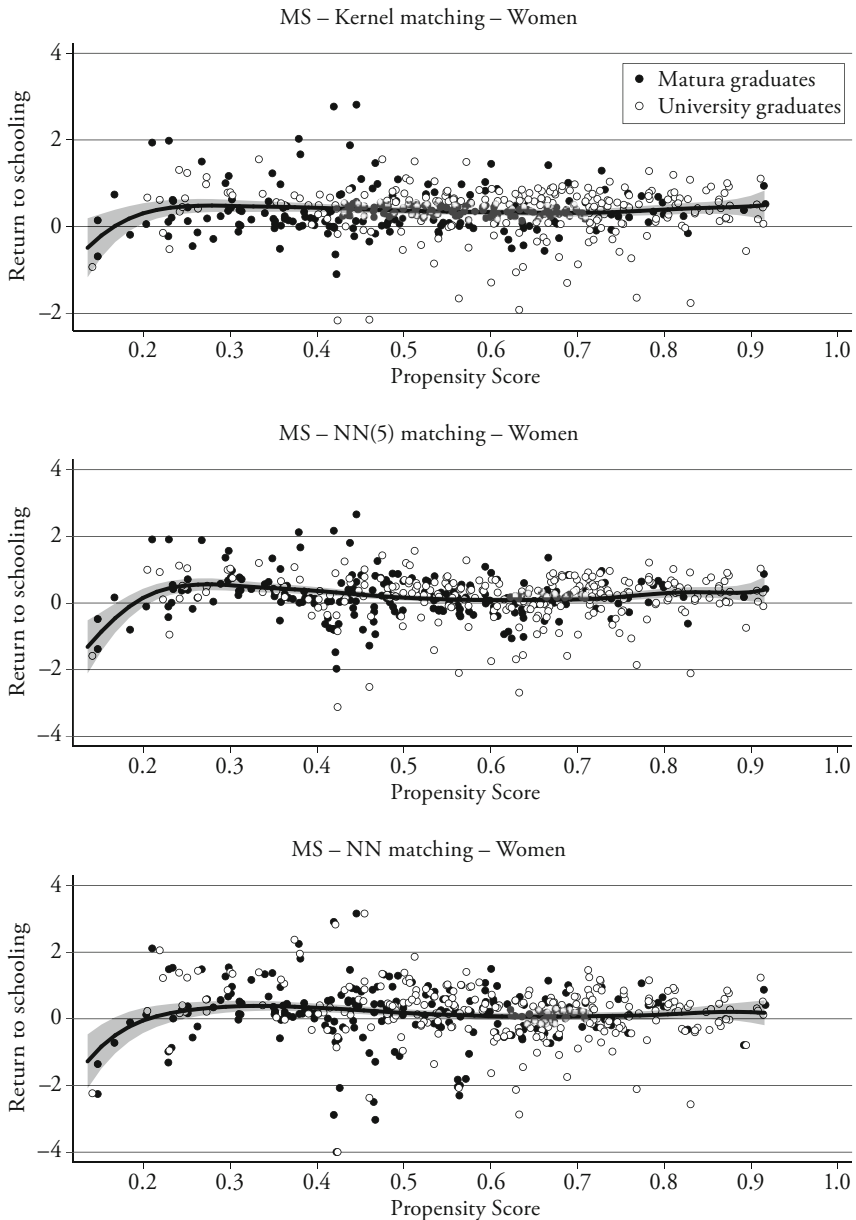Matching-Smoothing Model (MS-HTE)**

MS – Kernel matching – Men

MS – NN(5) matching – Men

MS – NN matching – Men

**Figure 4
continued**

MS – Kernel matching – Women

MS – NN(5) matching – Women

MS – NN matching – Women

**Table 5: Heterogeneous Returns to University Education,
Matching-Smoothing Estimates**

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

Notes: *, **, 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).

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, Hallock, and Sosa-Escudero, 2001). 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 ability. 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 6 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 15th quantile is significantly higher (at the 5% level) than those obtained by their peers in the 25th, 50th and 75th quantiles. 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 earning capability for men while both factors have no specific relation when considering women.

Table 6: Quantile Regressions, Men and Women

Quantile	0.15	0.25	0.50	0.75	0.85
Explanatory variables	Coefficients				
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

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

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 micro-econometric 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 differents hierarchical models based on propensity score matching: a stratification-multilevel model and a smoothing-matching approach.

The current work tests two opposite hypothesis about self-selection into university education. The former comes from the selection explanation of human capital theory which assumes that individuals self-select into schooling on the basis of their expected returns to schooling which in turn leads to higher earnings in the labor market. The latter stipulates that high propensity students who have more ascribed attributes that are valuable in the labor market may still obtain high earnings even if they do not well educationally. Consequently, it is possible that, under certain conditions, low propensity students benefit most from a university degree.

The stratification-multilevel method (SM-HTE) 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 labor market variables. This finding suggests that accounting for labor 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 matching-smoothing method (MS-HTE) – 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 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 the generation of earning capability for men but has not 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. In this context, increasing the incentives of 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 sufficiently tertiary education graduates to boost its economic growth, social policies focusing on the family background's gap may be of primary importance. Indeed, focusing on early childhood is very important given that differences in parental backgrounds have far-reaching 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 is a very promising step.

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 the SHP data lacks some information to analyze more precisely the self-selection mechanism into higher education. For example, data available 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.

8. Appendices

8.1 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:

$$Y_i = \alpha_i + \beta_i S_i + u_i \tag{20}$$

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

$$E[Y_i \mid S_i = 1] = \alpha_i + E[\beta_i \mid S_i = 1] + E[u_i \mid S_i = 1] \tag{21}$$

$$E[Y_i \mid S_i = 0] = \alpha_i + E[u_i \mid S_i = 0] \tag{22}$$

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

$$E[Y_i \mid S_i = 1] - E[Y_i \mid S_i = 0]$$
$$= \underbrace{E[\beta_i \mid S_i = 1]}_{ATT} + \underbrace{(E[u_i \mid S_i = 1] - E[u_i \mid S_i = 0])}_{\text{Selection bias}} \tag{23}$$

where the ATT parameter ($E[\beta_i \mid S_i = 1]$) 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:

$$E[\beta_i] = P(S_i = 0)E(\beta_i \mid S_i = 0) + P(S_i = 1)E(\beta_i \mid S_i = 1)$$
$$= P(S_i = 0)E(\beta_i \mid S_i = 0) + (1 - P(S_i = 0))E(\beta_i \mid S_i = 1)$$
$$= P(S_i = 0)[E(\beta_i \mid S_i = 0) - E(\beta_i \mid S_i = 1)] + E[\beta_i \mid S_i = 1]$$

I have then:

$$E[\beta_i \mid S_i = 1] = E[\beta_i] - P(S_i = 0)[E(\beta_i \mid S_i = 0) - E(\beta_i \mid S_i = 1)] \tag{24}$$

By pluging equation (24) into equation (23), I obtain the naive estimator of the average treatment effect (NATE) which can be written as

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

8.2 Additional Tables

Table 7: Propensity Score Matching and Covariates Balance, Men

Variable	Sample	$\bar{X}_{Treated}$	$\bar{X}_{Untreated}$	Bias(%)	% reduc $|Bias|$
Father education	Unmatched	0.239	0.106	35.7	20.5
	Matched	0.213	0.107	28.4	
Mother education	Unmatched	0.075	0.047	11.8	37.1
	Matched	0.066	0.048	7.4	
Financial problems	Unmatched	0.113	0.2	−24.0	8.6
	Matched	0.115	0.194	−21.9	
Parents living together	Unmatched	0.906	0.871	11.1	5.1
	Matched	0.908	0.875	10.5	
Treiman scale: mother job	Unmatched	42.461	39.225	39.0	16.6
	Matched	42.141	39.442	32.5	
Treiman scale: father job	Unmatched	48.217	42.775	44.6	16.6
	Matched	47.518	42.980	37.2	
Siblings	Unmatched	0.840	0.812	7.3	10.5
	Matched	0.843	0.812	6.5	
Father Swiss	Unmatched	0.774	0.741	7.5	31.1
	Matched	0.767	0.745	5.2	
Mother Swiss	Unmatched	0.723	0.682	8.9	6.5
	Matched	0.725	0.686	8.4	
Country of birth	Unmatched	0.846	0.788	14.9	12.0
	Matched	0.843	0.792	13.1	
Age	Unmatched	47.330	47.965	−5.7	54.8
	Matched	47.59	47.877	−2.6	
Age squared	Unmatched	2348.100	2434.600	−8.3	38.5
	Matched	2372.500	2425.700	−5.1	

Table 8: Propensity Score Matching and Covariates Balance, Women

Variable	Sample	$\bar{X}_{Treated}$	$\bar{X}_{Untreated}$	Bias(%)	% reduc $|Bias|$
Father education	Unmatched	0.298	0.148	33.7	22.6
	Matched	0.270	0.154	28.4	
Mother education	Unmatched	0.123	0.029	36.1	47.5
	Matched	0.078	0.028	19.0	
Financial problems	Unmatched	0.154	0.110	13.3	24.8
	Matched	0.148	0.114	10.0	
Parents living together	Unmatched	0.902	0.924	−7.8	7.5
	Matched	0.900	0.924	−8.4	
Treiman scale: mother job	Unmatched	44.091	40.000	46.4	31.9
	Matched	43.038	40.253	31.6	
Treiman scale: father job	Unmatched	50.220	44.315	46.9	24.6
	Matched	49.176	44.724	35.4	
Siblings	Unmatched	0.912	0.890	7.3	58.8
	Matched	0.907	0.898	3.0	
Father Swiss	Unmatched	0.695	0.814	−28.0	19.1
	Matched	0.711	0.808	−22.7	
Mother Swiss	Unmatched	0.646	0.790	−32.6	18.7
	Matched	0.667	0.784	−26.5	
Country of birth	Unmatched	0.786	0.848	−16.0	13.7
	Matched	0.793	0.846	−13.8	
Age	Unmatched	44.877	47.538	−26.3	19.8
	Matched	45.004	47.138	−21.1	
Age squared	Unmatched	2103.400	2373.8	−29.7	19.6
	Matched	2115.000	2332.5	−23.9	

Table 9: 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 10: Stratification-Multilevel Model, OLS Estimation, Men

Propensity score strata	Strata 1	Strata 2	Strata 3
Explanatory variables		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

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

Table 11: Stratification-Multilevel Model, OLS Estimation, Women

Propensity score strata	Strata 1	Strata 2	Strata 3
Explanatory variables		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

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

Table 12: Stratification-Multilevel Model, Sample Selection Specification, Women

Propensity score strata	Strata 1	Strata 2	Strata 3
Explanatory variables		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

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

9. References

ALTONJI, JOSEPH G., and THOMAS A. DUNN (1996),"The Effects of Family Characteristics on the Return to Education", *Review of Economics and Statistics*, 78(4), pp. 692–704.

ARIAS, OMAR, KEVIN F. HALLOCK, and WALTER SOSA-ESCUDERO (2001), "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data", *Empirical Economics*, 26(1), pp. 7–40.

ASHENFELTER, ORLEY, and CECILIA E. ROUSE (1998), "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics*, 113, pp. 253–284.

BASU, ANIRBAN, JAMES J. HECKMAN, SALVADOR NAVARRO-LOZANO, and SERGIO URZUA (2007), "Use of Instrumental Variables in the Presence of Heterogeneity and Self-Selection: An Application in Breast Cancer Patients", *Health Economics*, 16(11), pp. 1133–1157.

BAUER, PHILIPP C., and REGINA T. RIPHAHN (2007), "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.

BECKER, GARY S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: Columbia University Press.

BJORKLUND, ANDERS, and ROBERT MOFFITT (1987), "The Estimation of Wage Gains and Welfare Gains in Self-Selection Models", *Review of Economics and Statistics*, 69, pp. 42–49.

BOURDIEU, PIERRE (1977), Cultural Reproduction and Social Reproduction, in *Power and Ideology in Education*, J. Karabel and A. H. Halsey, eds., New York: Oxford University Press.

BRAND, JENNIE, and YU XIE (2010), "Who Benefits most from College? Evidence for Negative Selection in Heterogeneous Returns to Higher Education", *American Sociological Review*, 75(2).

BRAND, JENNIE, and JULI SIMON-THOMAS (2012), "Causal Effect Heterogeneity", California Center for Population Research, Working paper 2011–2012.

BUCHINSKI, MOSHE (1998), "Recent Advances in Quantile Regression Models", *Journal of Human Resources*, 33, pp. 88–126.

CARD, DAVID (1995), "Using Geographic Variation in College Proximity to Estimate the Return to Schooling", in *Aspects of Labour Market Behavior: Essays in Honour of John Vanderkamp*, L. Christofides, E. Kenneth Grant and R. Swidinsky, eds., pp. 201–222, University of Toronto Press, Toronto, Canada.

Card, David (1999), "The Causal Effect of Education on Earnings", in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, eds., Vol. 3A, Amsterdam: Elsevier, pp. 1801–1863.

Card, David (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, vol. 69, pp. 1127–1160.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil (2001), "Estimating the Return to Education when It Varies Among Individuals", Department of Economics, University of Chicago, mimeo.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil (2011), "Estimating Marginal Returns to Education", *American Economic Review*, 101(6), pp. 2754–2781.

Cattaneo, Alejandra, Sandra Hanslin, and Rainer Winkelmann (2007), "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.

Checchi, Daniele (2006), *The Economics of Education: Human Capital, Family Background and Inequality*, Cambridge University Press.

Coleman, James S. (1988), "Social Capital in the Creation of Human Capital", *American Journal of Sociology*, Vol. 94, pp. 95–120.

(CSRE) Centre Suisse pour la Recherche en Education (2010), Swiss Education Report 2010, Aarau.

Deschenes, Olivier (2007), "Estimating the Effects of Family Background on the Return to Schooling", *Journal of Business and Economic Statistics*, 25(3).

De Coulon, Augustin (1998), "Wage Differentials between Ethnic Groups in Switzerland", *Labour*, 15(1), pp. 111–132.

Diekmann, Andreas, and Henriette Engelhardt (1955), „Einkommensungleichheit zwischen Frauen und Mannern, Eine ökonometrische Analyse der Schweizer Arbeitskräfteerhebung", *Swiss Journal of Economics and Statistics*, 1, pp. 57–83.

Falter, Jean-Marc (2004), Educational Attainment in Switzerland: The Role of Observable and Non-Observable Family Factors, mimeo.

Ferro-Luzzi, Giovanni G., and Jacques Silber (1998), "On Gender Differences in Wages in Switzerland", *International Journal of Manpower*, 19(7), pp. 475–485.

Garen, John (1984), "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable", *Econometrica*, 52(5), pp. 1199–1218.

Golder, Stefan M. (1997), „Die Arbeitsmarktperformance von Einwanderern in der Schweiz. Eine theorische und empirische Analyse im Rahmen eines internationalen Vergleichs", WWZ-Forschungsbericht, Universtität Basel.

GRILICHES, ZVI (1977), "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, 45, pp. 1–22.

GRILLI, LEONARDO, and CARLA RAMPICHINI (2011), "Propensity Scores for the Estimation of Average Treatment Effects in Observational Studies", Training Sessions on Causal Inference, Bristol.

HALVORSEN, ROBERT, and RAYMOND PALMQUIST (1980), "The Interpretation of Dummy Variables in Semilogarithmic Equations", *American Economic Review*, 70(3), pp. 474–475.

HARMON, COLM, and IAN WALKER (1995), "Estimates of the Economic Return to Schooling for the United Kingdom", *American Economic Review*, 85(5), pp. 1278–1286.

HECKMAN, JAMES J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, pp. 153–161.

HECKMAN, JAMES J., and RICHARD ROBB (1985), "Alternative Methods for Evaluating the Impact of Interventions", in *Longitudinal Analysis of Labor Market Data*, J. Heckman and B. Singer, eds., New York: Cambridge University Press, pp. 156–245.

HECKMAN, JAMES J., and EDWARD VYTLACIL (1998), "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. 974–987.

HECKMAN, JAMES J., JUSTIN L. TOBIAS, and EDWARD VYTLACIL (2001), "Four Parameters of Interest in the Evaluation of Social Programs", *Southern Economic Journal*, Vol. 68, pp. 210–223.

HECKMAN, JAMES J., SERGIO URZUA, and EDWARD VYTLACIL (2006), "Understanding Instrumental Variables in Models with Essential Heterogeneity", *Review of Economics and Statistics*, 88(3), pp. 389–432.

HECKMAN, JAMES J., and EDWARD VYTLACIL (2007), "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 *Handbook of Econometrics*, J. Heckman and E. Leamer, eds., Vol. 6B, Chapter 71, pp. 4875–5144, Amsterdam: Elsevier.

HOLLAND, PAUL W. (1986), "Statistics and Causal Inference", *American Statistical Association*, 81(396), pp. 945–960.

JANN, BEN, JENNIE BRAND, and YU XIE (2008), -hte- Stata Module to Perform Heterogeneous Treatment Effect Analysis.

KANE, THOMAS J., and CECILIA E. ROUSE (1995), "Labor-Market Returns to Two- and Four-Year College", *American Economic Review*, Vol. 85, pp. 600–614.

KUGLER, PETER (1988), „Lohndiskriminierung in der Schweiz: Evidenz von Mikrodaten", *Swiss Journal of Economics and Statistics*, 124(1), pp. 23–47.

MILLER, PAUL, CHARLES MULVEY, and NICK MART (1995), "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.

MINCER, JACOB A. (1974), *Schooling, Experience, and Earnings*, New York: Columbia University Press.

MORGAN, STEPHEN L. (2005), *On the Edge of Commitment: Educational Attainment and Race in the United States*, Stanford, CA: Stanford University Press.

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

PEREIRA, PEDRO T., and PEDRO SILVA-MARTINS (2004), "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 countries", *Labour Economics*, 11(3), pp. 355–371.

PONS, EMPAR, and MARIA T. GONZALO (2002), "Returns to Schooling in Spain: How Reliable Are Instrumental Variable Estimates?", *Labour*, 16(4), pp. 747–770.

PUTNAM, ROBERT (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York, Simon and Schuster.

RANTANEN M. (2009), "Returns to Education Using the Finnish Twin Data", mimeo.

ROBERTS, MICHAEL R.(2009), "Causality and Experiments", University of Pennsylvania.

SHELDON, GEORGE (1992), „Selbstselektion und Bildungsrenditen – ökonometrische Untersuchungen an einem Mikro-Datensatz für die Schweiz", in *Ökonomie und Politik beruflicher Bildung – Europäische Entwicklungen*, Sadowski D. und A. Timmesfeld, eds., Duncker und Humblot, Berlin.

SMITH, HERBERT L., and BRIAN POWELL (1990), "Great Expectations: Variations in Income Expectations among College Seniors", *Sociology of Education*, 63, pp. 194–207.

SUTER, STEPHAN (2006), *Bildung oder Begabung? Eine ökonometrische Untersuchung zu den Ursachen individueller Lohnunterschiede*, Schriftenreihe volkswirtschaftliche Forschungsergebnisse, 113, Hamburg.

TSAI, SHU-LING, and YU XIE (2008), "Changes in Earnings Returns to Higher Education in Taiwan since the 1990s", *Population Review*, 47, pp. 1–20.

VELLACOTT, MAJA C., and STEFAN C. WOLTER (2004), *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.

Willis, Robert J., and Sherwin Rosen (1979), "Education and Self-Selection", *Journal of Political Economy*, 87, pp. 7–36.

Willis, Robert J. (1987), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", in *Handbook of Labor Economics*, Vol. 1, O. Ashenfelter and R. Layard, eds., Chapter 10.

Wolter, Stefan C., and Bernhard A. Weber (2005), «La rentabilité de la formation: un indicateur économique essentiel à l'enseignement», *La Vie Economique*.

Xie, Yu, Jennie Brand, and Ben Jann (2011), "Estimating Heterogeneous Treatment Effects with Observational Data", Population Studies Center Research Report 11-729.

SUMMARY

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 labor market variables – benefit most from a university degree while returns for women are rather homogenous along the propensity score distribution. This finding suggests that completing university increases more the earning capability of men with disadvantaged family backgrounds than that of men with more favorable background, refuting the hypothesis of comparative advantage. An auxiliary analysis focusing on the relationship between returns to education and inherent ability within a quantile regression framework leads to similar conclusions.