Swiss Leading House
Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 91

Heterogeneous Returns to Education Over the Wage Distribution: Who Profits the Most?

Simone Balestra and Uschi Backes-Gellner

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Leading House Working Paper No. 91

Heterogeneous Returns to Education Over the Wage Distribution: Who Profits the Most?

Simone Balestra and Uschi Backes-Gellner

July 2013

Published as: "Heterogeneous Returns to Education Over the Wage Distribution: Who Profits the Most?" *Labour Economics*, 44(2017): 89-105. By Simone Balestra and Uschi Backes-Gellner.

HETEROGENEOUS RETURNS TO EDUCATION OVER THE WAGE DISTRIBUTION: WHO PROFITS THE MOST ?

SIMONE BALESTRA [a]
Center for Education Policy Analysis, Stanford University
520 Galvez Mall, Stanford, CA 94305, USA

USCHI BACKES-GELLNER [b]
Department of Business Administration, University of Zurich
Plattenstrasse 14, 8032 Zurich, Switzerland

FIRST DRAFT – PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

This study presents evidence of heterogeneous returns to education over the wage distribution. The authors use instrumental variable quantile regression and data from the Swiss Labor Force Survey to identify the causal link between education and earnings at different quantiles of the conditional distribution of wages. The results provide evidence that there is no unique causal effect of schooling and that for each individual the effect may be above or below those extensively documented by OLS or TSLS. In particular, while ordinary quantile regression estimates increasing returns in the quantile index, once the endogeneity of schooling is taken into account the authors instead observe higher returns at lower quantiles of the wage distribution. Interpreting the quantile index as a measure of unobserved ability, the findings suggest that less able individuals profit more form one additional year of education. The authors also investigate the presence of heterogeneity between and within educational paths, comparing the returns academic education with the returns vocational education over the wage distribution. The results indicate that academic education brings a significant return premium in the upper part of the wage distribution. However, such premium vanishes around the third decile and becomes negative at the bottom of the distribution.

Keywords: Returns to education, IVQR , academic education, vocational education.
JEL Classification: I21, J24, C31, C36.

[a] Simone Balestra is a visiting scholar at the Center for Education Policy Analysis at Stanford University and a research associate at the Institute for Strategy and Business Economics at the University of Zurich, Switzerland. Corresponding author: simone.balestra@uzh.ch. Postal Address: Plattenstrasse 14, 8032 Zurich, Switzerland. Telephone: +4144 634 42 20.
[b] Uschi Backes-Gellner is full professor and deputy dean at the faculty of Business Economics and Informatics at the University of Zurich, Switzerland.
This study is funded by the Swiss National Science Foundation (project PBZHP1_147219) and by the Swiss Federal Office for Professional Education and Technology through its Leading House on the Economics of Education, Firm Behavior and Training Policies. We are very grateful to Eric Bettinger, David Card, Thomas Dee, Guido Imbens, Brian Jacob, Simon Janssen, Susanna Loeb, Christian Rupietta, and Eric Taylor for their constructive comments and valuable help throughout the development of this study. We are also grateful to all the participants in the seminar on Advanced Topics in Quantitative Policy Analysis at Stanford University for their useful feedback. We would like to thank the Swiss Federal Statistical Office for the data, Christian Zuend for student assistance, and Christian Hansen for software provision through his webpage.

1. Introduction

While a positive relationship between schooling and wage is beyond doubt (Dickson and Harmon, 2011), the question whether education affects individuals differently over the wage distribution is much less analyzed (Wang, 2013). Moreover, in this distributional setting, the existing literature has not investigated whether different types of education result in differing returns, and whether one type of education—vocational or academic—brings a return premium compared to another at some points of the wage distribution. This issue is particularly important, because a lack of information about educational tracks may yield costly decisions for the individual as well as for the government (Bettinger and Baker, 2011).

To fill these gaps, in this paper we first estimate the returns to education over the wage distribution. This analysis allows uncovering the heterogeneous effects of education on wages, and whether the returns are increasing, decreasing, or u-shaped across the quantiles. In a second step, we compare the returns to one extra year of academic education with the returns to one extra year of vocational education; to investigate whether one track brings a return premium at some points in the wage distribution. Such a comparison is missing in the literature, generally because most countries do not have an extensive vocational education and training system that allows acquiring the same quality of education and the same number of years as in the academic track, or because the academic track is more prestigious or preferred than the vocational one.[1] One notable exception is Switzerland,[2] a country with an extensive vocational education system that attracts two thirds of the individuals in a cohort (Federal Office for Professional Education and Technology, 2012). The Swiss educational system lets students achieve tertiary education degrees for both academic and vocational tracks. Therefore, a study with Swiss data permits to shed light on heterogeneous returns to different

[1] As for example in Colombia (Bettinger *et al*. 2010).

[2] Other countries with a similar vocational system are Denmark and Germany (Hanushek *et al*,. 2011).

types of education, and allows to answer the question of how academic and vocational education differ over the wage distribution.

The analyses we propose face two major issues that are common when estimating returns to education: endogeneity of education attainment (Harmon *et al*., 2003) and heterogeneity in the returns to education (Henderson *et al*., 2011). While theory considers both issues simultaneously (Card, 1999; Arias *et al*., 2001), empirical work often deals with only one issue at a time. To overcome the endogeneity problem, the most used approach is instrumental variable estimation (Angrist and Krueger, 1991; Dickson, forthcoming; Harmon and Walker, 2000; Trostel *et al*., 2002). Conversely, when dealing with the heterogeneity issue, the literature has not converged to a standard method to integrate it in the analysis (Lemieux, 2008). Therefore, researchers usually rely on different methods to account for heterogeneity in returns to education: sub-sample analysis (see for example Harmon *et al*., 2003), nonparametric estimation (Henderson *et al*., 2011), Bayesian hierarchical models (Koop and Tobias, 2004), and quantile regression (Fasih *et al*., 2012; Martins and Pereira, 2004). The first three methods focus mainly on the existence and the nature of heterogeneity, which are not the focus of this paper. Quantile regression is instead more appropriate to our research question, because quantile regressions estimate the returns to education over the wage distribution allowing for heterogeneity through quantile-specific intercepts and quantile-specific slopes.

The use of quantile regressions (QR) in returns to education studies has been hindered because the endogeneity problem in QR models could not be solved in the past. However, recent studies by Chernozhukov and Hansen (2006, 2008, 2013) propose an instrumental variable quantile regression approach (IVQR) that addresses both heterogeneity and endogeneity issues at the same time. The IVQR method has been applied in many research fields, but it is rather new to the returns to education literature. To our knowledge, only two studies implement IVQR to propose alternative instruments for schooling (Arabsheibani and Staneva, 2012) and to examine the inequality-reducing effect of education in China (Wang, 2013).

Exploiting a major reform of the education system that took place in Switzerland in the years following the federal concordat of 1970, in this paper we causally estimate the returns to education over the wage distribution by IVQR and we compare the results with standard quantile regression and ordinary least squares to see whether taking endogeneity into account changes results and conclusions. In a second step, we also distinguish between educational paths, to convey a new comparison between and within academic and vocational education. In this latter comparison we are interested especially in the presence of heterogeneity, and we therefore present conventional QR methods only. We nevertheless performed several instrumental variable (quantile) regressions that are available upon request.

The reminder of the paper proceeds as follows. Section 2 gives an overview of the theoretical background related to our research questions. Section 3 introduces the dataset and presents some descriptive statistics. Section 4 shows the econometric models in detail. Section 5 presents the results, and Section 6 concludes.

2. Background

In this Section, we briefly present some theoretical background and empirical evidence to explain the underlying mechanisms in the individual education choice and provide a structure for our empirical analysis. We borrow the theoretical model developed by Card (1999), whose most interesting feature is that it considers both heterogeneity in the returns and endogeneity of education attainment in the earnings equation at the same time.

Following Card, we assume that an individual chooses his level of education to maximize the following utility function defined over wage[3] and years of education:

$$U(w,S) = \ln(w) - f(S) = \ln[g(S)] - f(S)$$

Where $\ln[g(S)]$ and $f(S)$ are increasing convex functions that represent the benefits and costs of schooling, respectively. The condition $w = g(S)$ captures the observable relationship of wage to schooling, i.e. the level of wages available at each level of education. The first order condition for optimal education is:

$$\frac{g'(S)}{g(S)} = f'(S)$$

In the optimum, the marginal rate of return to education ($g'(S)/g(S)$) equals the marginal cost ($f'(S)$). Individual heterogeneity in the optimal education choice arises from two sources: differences in the costs of education, represented by the variation in $f(S)$; and differences in the monetary benefit of education, represented by the variation in $g'(S)/g(S)$.

To characterize the well-documented fact that (log) wage is a nearly linear function of schooling that may vary across individuals,[4] we impose the following functional form to the heterogeneity components:

$$\mathrm{MB}_i = \frac{g'(S)}{g(S)} = b_i - k_1 S_i \qquad \text{and} \qquad \mathrm{MC}_i = f'(S) = r_i + k_2 S_i$$

where b_i and r_i are random variables with some joint distribution across the population $i = 1,2,...$ and k_1 and k_2 are nonnegative constants. This specification implies that the optimal educational choice is linear in the individual-specific heterogeneity terms:

$$S_i^* = \frac{b_i - r_i}{k} \quad , \quad \text{with} \quad k = k_1 + k_2$$

[3] Card and Krueger (1992), Heckman and Polachek (1974), and Hungerford and Solon (1987) present evidence suggesting that earnings are nearly log-linear with respect to schooling.

[4] Park (1994) finds log-linearity to be a good approximation of the earnings-schooling relationship not only at the mean but also for several quantiles of the earnings distribution.

To derive an equation for the natural logarithm of wage, Card integrates the expression for the marginal rate of return to education with respect to S_i and obtains:

$$\ln(w_i) = a_i + b_i S_i - \frac{1}{2} k_1 S_i^2$$

where a_i is an individual-specific constant of integration.

This last equation is a general version of the functional form adopted in Mincer (1974). However, the appealing feature of this model is that individual heterogeneity potentially affects both the intercept of the wage equation (through a_i) and the slope of the wage-education relation (through b_i). Such feature introduces three important issues in the empirical work. First, we should expect different returns to education for individuals with different levels of ability. More specifically, given that individuals acquire education up to the point where the marginal cost equals the marginal rate of return and that costs depend negatively on ability, we should observe the returns to education to be decreasing in ability. As indicated by Ashenfelter and Rouse (1998), abler individuals indeed acquire more schooling, but this is because they face lower marginal costs and not because of higher marginal benefits. This implies that higher ability individuals have on average higher wages, but the slope of their wage-education profile is flatter than that for lower ability individuals.

Second, we cannot assess the true impact of education on wage without solving the bias introduced by the endogeneity of schooling attainment, because otherwise cross-sectional estimates are (marginally) upward biased by an omitted ability variable (Heckman *et al*., 2006). Third, if we want to study how education affects different individuals, we should account for both heterogeneity and endogeneity issues simultaneously.

To incorporate these features in our analysis we use IVQR, which estimates the causal effect of education on conditional quantiles of the wage distribution, allowing for quantile-specific intercepts and quantile-specific slopes. Given that IVQR is a rather new method, the vast majority of the existing literature uses conventional QR to investigate the heterogeneous effects of education on wage (Fasih *et al*., 2012; Harmon *et al*., 2003; Hartog *et al*., 2001; Martins and Pereira, 2004). From these studies we conclude that returns to education vary substantially over the wage distribution, which means that average effects—while being of interest—lose some important distributional features of the return to education. These studies also suggest that returns to education are increasing in the quantiles of wage distribution. As we can interpret the quantile index as a measure of ability (Arias *et al*., 2001; Mwabu and Schultz, 1996), this finding contrasts with what we would theoretically expect. However, the implicit assumption of exogenous schooling in conventional QR studies may explain the discrepancy between theoretical expectation and empirical findings (Chernozhukov and Hansen, 2006).

The few studies applying IVQR in the return-to-education context present mixed results. Using spousal education as an instrument for education, Wang (2013) investigates the evolution of the returns in China over time, to examine the inequality-reducing effect of education. He estimates slightly decreasing returns to education over the wage distribution, ranging from 5.1 percent at the lowest quartile to 3.1 percent at the highest quartile. Proposing risky sexual behavior at early age as a new instrument for schooling, Arabsheibani and Staneva (2012) apply IVQR to Russian data and find increasing returns over the wage distribution. Specifically, they estimate a 5-percent return at the lowest decile and a 15-percent return at the highest decile. However, both approaches rely on a demand-side variation in schooling to estimate the causal return to education, which makes it difficult to defend the orthogonality between the instruments and the error term of the wage equation (Arcand *et al*., 2005).

Pushing the analysis further, researchers and policymakers are often interested in the return to different educational paths such as academic education and vocational education.

While most studies on returns to education do not consider the curriculum content of the variable "years of education," policymakers—as well as students and parents—may need more information than simply the average return to a year of education, especially when they have to take decisions about different types of educational investments. In this context, a typical question that has arisen is whether vocational education yields a lower or higher return in the labor market, as compared to academic education with the same number of years.

Generally, the existing literature suggests that academic degrees have larger benefits than vocational degrees. Dearden *et al*. (2002) provide evidence on the relative value of academic and vocational qualifications in the British labor market. Their results show that the wage premium associated with academic qualifications is on average higher than the premium associated with vocational qualifications at the same level. Similarly, Saniter (2012) examines the returns to education for different educational groups in Germany. He finds that the return to education is 8.5 percent for the entire sample, 2.3 percent for graduates from the basic school track (vocationally-oriented), and 11 percent for graduates from a higher school track (academically-oriented). Focusing on nonmonetary benefits of educational tracks, Hanushek *et al*. (2011) find that gains in youth employment from vocational education are offset by less adaptability and consequent diminished employment later in life. Thus, over the life-cycle, academic education seems to have larger nonmonetary benefits compared to vocational education.

However, all these studies do not actually analyze the return to one extra year of academic education with the return to one extra year of vocational education to investigate whether the one track brings a return premium, and they do not explore the possibility of heterogeneous effects between and within educational paths. They rather focus on qualifications and nonmonetary benefits, probably because many countries do not have an education system that allows acquiring up to tertiary degrees for both academic and vocational tracks, so years of education are typically very different in the two tracks. This is not the case in Switzerland, and we complement the discussion on academic versus vocational track by revealing the heterogeneous effects of the two educational paths and by analyzing whether—and at which point of the wage distribution—one track has higher returns compared to the other.

3. Data and Descriptive Statistics

Before introducing the data and providing descriptive statistics, we briefly present the current Swiss education system. The education system in Switzerland consists of parallel paths divided into vocational and academic education. After 9 years of compulsory schooling, about two thirds of a youth cohort choose to pursue vocational education and training (Federal Office for Professional Education and Technology, 2012), mostly within the so-called dual system of apprenticeship training. This kind of training generally consists of a curriculum-based on-the-job training component and a theoretical component taught at specialized vocational schools. After graduation, most of these apprentices work as skilled workers within their occupational fields. Alternatively, vocational graduates also have several options to continue their education. They may choose to go into higher vocational education and acquire a higher vocational education degree or a university of applied sciences degree.

Another possibility for students after compulsory education is to stay in the academic school system, attend high school, and obtain a Matura. This degree is a necessary and sufficient condition to access tertiary academic education, which consists of universities and federal institutes of technology. At these tertiary academic institutions, students can acquire a bachelor, a master, and/or a doctorate.

We base our analysis on data from the Swiss Labor Force Survey (SLFS), which is produced annually by the Swiss Federal Statistical Office. The data are collected by telephone interviews and the sample is representative for the adult population permanently living in Switzerland. The main purpose of the SLFS is to provide information on the structure of the labor force and employment behavior patterns. Strict adherence to international definitions allows Swiss data to be compared with OECD, European, and U.S. data. The SLFS was conducted for the first time in 1991 and it is based on a sample of some 105,000 interviews. We select the period 2000-2009 and we pool these cross sections to build our sample.[5]

To avoid special circumstances such as those that might arise from retirement, our sample takes into account only males aged 18-60. We also restrict the sample to individuals who are employed to avoid misspecifications due to people who are in school or not active in the labor force. Among the employed, we focus on fully employed workers to retain individuals with attachment to the labor market.[6] The wage variable of the SLFS comes from the Swiss Survey on Income and Living Conditions, a very precise data source for income resulting from labor activity. Among those individuals with no missing wage, we excluded 0.5 percent of each tail end of the wage distribution to attenuate the impact of outliers and remove implausible values. Wages are expressed in Swiss Francs (CHF) throughout the entire paper, deflated to the year 2010.[7]

In the SLFS, for each individual, we can observe the entire education path from compulsory education to doctorate, and we dichotomize the educational paths into academic and vocational according to the official definition of the Swiss State Secretariat for Education

[5] The SLFS is a rotating panel. We keep one observation per individual to prevent problems of nonrandom attrition and clustering.

[6] We use the official definition of the Swiss Federal Statistical Office, which considers an individual as fully employed if he has an employment of at least 90 percent.

[7] In 2010, 1 CHF = 1 USD.

and Research (Appendix A). After removing missing values, we are left with 34,744 observations in the sample. Table 1 provides descriptive statistics.[8]

{ Place Table 1 about here }

From the descriptive analysis on the full sample (Table 1, Panel A), we observe that the average worker earns an annual wage of CHF 81,868 and has acquired 13.16 years of education. In line with the statistics at the national level (Federal Office for Professional Education and Technology, 2012), in our sample 65 percent of the individuals followed a vocational path, whereas 23 percent obtained an academic degree. The rest of the sample (12 percent) has compulsory education as the highest educational level.

Comparing descriptive statistics of vocational graduates (Panel B, Table 1) with descriptive statistics of academic graduates (Panel C, Table 1), we observe that vocational graduates earn on average less than their academic counterpart (CHF 77,378 per year compared to CHF 108,592 per year). Furthermore, vocational graduates tend to acquire less years of education in total, and this difference is explained by less time spent in post-compulsory education (4.34 years compared to 7.94 years for academic graduates). However, unobserved heterogeneity is not taken into account in these figures; particularly ability differences are not considered. Therefore descriptive results give no indication of causal wage effects of different types of education.

[8] See Appendix B for the details on sample construction.

4. Methods

In the first part of this Section, we introduce the equations to be estimated. We use two different models: one to analyze the return to education and one to compare the academic track with the vocational track. In the second part of this Section, we briefly describe the estimation methods we apply, which consist of ordinary least squares, quantile regressions, and instrumental variable estimations. Given that quantile regression and instrumental variable quantile regression are not as common as ordinary least squares and two-stage least squares, we give a brief overview on these two methods following Koenker and Basset (1978) and Chernozhukov and Hansen (2008). In the third part of this Section, we describe and discuss the instrumental variables we use for the causal estimation of the returns to education.

4.1 The Wage Equations

To estimate the private monetary return to one additional year of education, we consider the following Mincer-like equation:

$$\ln(w_i) = \delta_0 + \beta_S S_i + \delta_1 Age_i + \delta_2 Age_i^2 + \gamma_t + u_i \tag{1}$$

In equation (1), w_i is the annual wage of individual i, S_i represents the years of education, Age_i is a proxy for labor market experience, is γ_t a set of time controls, and u_i is an error term. As frequent in the literature, we exclude various determinants of earnings such as tenure and industry sector, because such variables are potentially endogenous and determined by education itself (Angrist and Pischke, 2009; Wang, 2013). In model (1), the coefficient of interest is the one on the variable years of schooling β_S, which we expect to be positive and significant.

To compare the effect of one additional year of academic education to the effect of one additional year of vocational education, we develop a model similar to the one used by Hartog *et al*. (2001) and by Vandenbussche *et al*. (2006). Hartog *et al*. (2001) modify the classical

Mincer wage equation and include a spline in year of education at three categories of the school system, namely primary, secondary, and tertiary education. With this specification, they investigate the different effects of education on wages among different levels of education. With a similar specification, Vandenbussche *et al*. (2006) study the effect of tertiary education on countries' growth rate. They separate the effect of tertiary education from primary and secondary education to show that skilled labor has a higher growth-enhancing effect for countries closer to the technological frontier. In our case, we decompose the education variable as defined in model (1) in its three components: Compulsory education (C), vocational education (V), and academic education (A). Thus, we can rewrite equation (1) as follows:

$$\ln(w_i) = \delta_0 + \delta_C C_i + \delta_V V_i + \delta_A A_i + \delta_1 Age_i + \delta_2 Age_i^2 + \gamma_t + u_i \qquad (2)$$

In model (2), the parameters of interest are δ_V and δ_A. With this second specification, we want to compare the return premium of one additional year of vocational education with the premium of one additional year of academic education.[9] While it is reasonable to expect both parameters to be significant and positive, building expectations about the comparison between the two is not straightforward for two reasons. First, previous literature on the topic is rather scarce. The existing studies either compare higher tracks with lower tracks (Saniter, 2012), or focus on nonmonetary returns (Hanushek *et al*., 2011) and returns to qualifications (Dearden *et al*., 2002). Second, inserting this topic in a distributional framework brings additional challenge, because—as in the case of returns to education in general—the returns to the vocational (academic) path may be heterogeneous over the wage distribution.

[9] To test whether the two coefficients are different we perform an F-test, whose null hypothesis is: $\hat{\delta}_V - \hat{\delta}_A = 0$.

4.2 Quantile Regression (QR) and Instrumental Variable Quantile Regression (IVQR)

The vast majority of applied econometrics focuses on averages, and such focus partly reflects the fact that it is hard enough to produce credible average causal effects As long as the dependent variable is binary, the mean describes the entire distribution. However, many variables like earnings have continuous distributions, and these distributions can change in response to treatments in ways not fully revealed by averages. QR provides a straightforward, powerful tool for modeling distributional effects, even if the underlying mechanism is complex and multidimensional (Angrist and Pischke, 2009).

To allow for heterogeneous effects of education on wages, we consider the τ^{th} conditional quantile wage function hereafter:

$$Q_{\ln(w)}[\tau \mid X, S] = \alpha(\tau)X + \beta(\tau)S \qquad (3)$$

where X denotes all explanatory variables other than the educational variables $(1, Age_i, Age_i^2, \gamma_t)$, $\alpha(\tau)$ is the return to X at the τ^{th} quantile, $\beta(\tau)$ is the return to education at the τ^{th} quantile, and $\tau \in (0,1) \mapsto \alpha(\tau)X + \beta(\tau)S$ is strictly increasing in τ. In equation (3) the returns to education are function of τ, allowing for heterogeneous effects of education on earnings.

Assuming the error term in the wage equation to be independent of X and S, Koenker and Basset (1978) propose to find the best predictor of log-wage given X and S under the asymmetric least absolute deviation loss. This means estimating $\alpha(\tau)$ and $\beta(\tau)$ in equation (3) by solving the following minimization problem:

$$Q_{\ln(w)}[\tau \mid X, S] = \underset{\alpha(\tau), \beta(\tau)}{\arg\min}\, \mathrm{E}\left[\rho_\tau\left(\ln(w) - \alpha(\tau)X - \beta(\tau)S\right)\right]$$

where $\rho_\tau(u_i)$ is the "check function" defined as $\rho_\tau(u_i) = [\tau - 1(u_i \leq 0)]u_i$. In practice, the minimization problem is solved via linear programming and implemented in many statistical packages.

As already discussed, assuming independence between the education variable and the error term may be too stringent because of potential unobserved wage determinants (i.e., ability bias). To account for potential dependence between S and u in a distributional framework, we apply the IVQR method developed by Chernozhukov and Hansen (2006, 2008). As in the case of two-stage least squares, the identification of the IVQR approach relies on the existence of a vector of instrumental variables Z that is statistically related to S but independent of the error term u. Additionally, we have to assume that, given the information (X,Z), the distribution of the structural error does not vary across the endogenous state S ("rank similarity").[10] The structural error is responsible for heterogeneity of potential outcomes among individuals with the same observed characteristics, and this error term determines the relative ranking of observationally equivalent individuals in the distribution of potential outcomes given the individuals' observed characteristics. Chernozhukov and Hansen show that assuming rank similarity implies the following moment condition:

$$\mathrm{P}[\,\ln(w) \le Q_{\ln(w)}(\tau \mid X,Z) \mid X,Z\,] = \tau$$

and thus, in our case:

$$\mathrm{P}\left[\ln(w) - \alpha(\tau)X - \beta(\tau)S \le 0 \mid X,Z\right] = \tau \qquad (4)$$

The moment condition given in (4) provides a statistical restriction that can be used to estimate the parameters $\alpha(\tau)$ and $\beta(\tau)$. Pointing out that equation (4) is equivalent to the statement that 0 is the τ^{th} quantile of the random variable $\ln(w) - Q_{\ln(w)}(\tau \mid X,S)$ conditional on (X,Z), Chernozhukov and Hansen formulate the problem as finding $[\alpha(\tau),\beta(\tau)]$ such that zero is the solution to the standard quantile regression of $[\ln(w) - \alpha(\tau)X - \beta(\tau)S]$ on (X,Z):

---

[10] $u \mid X,Z \sim \mathit{Uniform}(0,1)$, which means that for each S and S' given (X,Z): $U_S \sim U_{S'}$.

$$0 = \underset{f \in F}{\arg\min}\, \mathrm{E}\left[\rho_\tau\left(\ln(w) - \alpha(\tau)X - \beta(\tau)S - f(X,Z)\right)\right] \qquad (5)$$

where F is the class of measurable functions of (X,Z). In empirical applications, F will be restricted either to the values of Z_i or to the predicted value from a least squares projection of S_i on X_i, Z_i. To obtain an estimate for $\beta(\tau)$, we look for a value β that makes the estimated coefficient on the instrumental variable $\hat{\gamma}(\tau,\beta)$ in equation (5) as close to zero as possible using conventional quantile regression.

In practice, the IVQR estimator consists of a two-step procedure: for a given value of $\beta^j(\tau)$, we first run the ordinary QR of $\ln(w_i) - \beta^j(\tau)S_i$ on X_i and Z_i to obtain the estimates $\hat{\alpha}(\beta^j(\tau),\tau), \hat{\gamma}(\beta^j(\tau),\tau)$. Second, we test $\hat{\gamma}(\beta^j(\tau),\tau) = 0$ and save the corresponding F-statistic, F_j. We then repeat these two steps for all the values in a pre-specified support for $\beta^j(\tau)$ and the value that minimizes the F-statistic is the IVQR estimator $\hat{\beta}(\tau)^{IVQR}$. Once we have $\hat{\beta}(\tau)^{IVQR}$, we retrieve the correspondent $\hat{\alpha}(\tau)$.[11]

Note that the IVQR approach allows interpretation of the $\hat{\beta}(\tau)^{IVQR}$ as actual effects on individuals having fixed their level of unobserved heterogeneity at a given quantile. Therefore, the effect is not only identified for the set of individuals whose treatment is altered by switching the instrument from zero to one as in the case of the IV quantile treatment estimator proposed by Abadie *et al*. (2002). Furthermore, the IVQR method puts no restriction of the form of the endogenous variables and instruments (i.e., they can be binary, discrete, or continuous).

[11] We use both the Stata command `ivqreg` and the Matlab function `inv_qr` to obtain the point estimates and standard errors, with almost no difference between the two approaches.

4.3 The Instruments

Given the widely acknowledged endogeneity of educational attainment in the wage equation, it is important to find valid instruments to control for this phenomenon. However, choosing suitable instruments is still a topic of great debate in the literature on returns to education (Arcand *et al*., 2005; Dickson, forthcoming; Heckman *et al*., 2006). In this setting, an ideal instrument should be correlated with educational attainment but uncorrelated with the unobserved determinants of the wage.

The literature on returns to education used several instruments for education: quarter of birth (Angrist and Krueger, 1991), early smoking habits (Evans and Montgomery, 1994), presence or sex of siblings (Angrist and Evans, 1998; Butcher and Case, 1994), college proximity (Card, 1994), parental education (Harmon and Walker, 2000), and spouse education (Trostel *et al*., 2002). In the last years, the literature has been investigating educational reforms as a source of exogenous variation in educational attainment.[12] In particular, changes in school leaving age (Dickson, forthcoming; Harmon and Walker, 1999) and compulsory education expansions (Brunello *et al*., 2013; Fang *et al*., 2012; Meghir and Palme, 1999) have been attracting researches' interest. Following this last strand of literature, we exploit a major reform in the Swiss educational system to build our instruments and estimate the true (causal) effect of education on earnings.

In Switzerland, the main responsibility for education and culture lies with the cantons, and they loosely coordinate their work at the national level. The 26 cantonal ministers of education together form a political body named the Swiss Conference of Cantonal Ministers of Education (EDK). The EDK is responsible for educational reforms, policies, and coordination at the national level. In 1970, the EDK produced an important educational concordat, whose aim was harmonizing the Swiss education from compulsory school to high

[12] For a recent study on the impact of educational reforms on educational attainment see Braga *et al*., 2013.

school. This concordat became official on October 29, 1970, and took a couple of years to be implemented in the first cantons. Before this concordat, cantons had different compulsory school duration (7, 8, or 9 years), different school year start (either spring of fall), and different duration from compulsory education to high school graduation (from 11.3 years to 13.5 years).

The concordat set 9 years of compulsory education for all cantons, and decided that the school year has to start in fall. Given that some cantons were already in line with this reform, only about one half of them had to change their education system. Furthermore, cantons did not introduce all reforms immediately after 1970. They instead had time to adapt their education systems in the years following the concordat and constantly feedback the EDK on the reform status. Thanks to that, we are able to keep track of the introduction of the reform in each canton. Additionally, we also contacted each canton's educational ministry that had to modify the educational system to double-check their reform status. Appendix C gives an overview of the reforms for each canton, when they introduced the reforms, and how the reform modified a canton's education system.

We use the compulsory education expansion as an instrument for years of education. The empirical literature suggests that postponing the allocation of pupils to tracks yields positive effects on average educational attainment, because students stay in school longer and drop out less (Braga *et al*., 2013). Similar to Brunello *et al*. (2013) and Fang *et al*. (2012), we exploit the series of natural experiments created by the staggered implementation of Switzerland's education reform as instrument for estimating each individual's completed years of schooling. This approach obviates the problem of endogeneity due to unobservable variables that are correlated with both education and earnings.

Given that the effective date of the shift in school-year start also constitutes a (small) exogenous change in years of education, we might be willing to use this change as a second instrument for years of education completed. In Appendices D and E we present TSLS and IVQR estimates for a series of overidentified models using both instruments, but we base our main analysis on the compulsory education expansion reform, because it is an instrument already known to the literature and because the results are qualitatively similar to the overidentified cases.

5. Results and Discussion

5.1 Returns to Education over the Wage Distribution

Table 2 shows the regression outputs for model (1), which focuses on the returns to education. Mean regression (column 1, Table 2) estimates a return to education of 6.73 percent, which indicates that earnings rise by 6.73 percent on average with each extra year of education. The effect is highly significant and not far from the few previous studies on returns to education in Switzerland, which estimate returns of about 7-8 percent (Weber and Wolter, 2006).

Allowing for heterogeneous effects of education on wage, an interesting picture emerges. QR estimates (columns 2-10, Table 2) show that returns to education increase over the quantiles of the wage distribution. The return to education is 3.88 percent at the bottom decile, increasing to 6.87 percent at the median ($\tau = 0.5$), and reaching 8.89 percent at the top decile of the wage distribution. These results underline that average effects may hide useful information about the rest of the distribution, and to further underline the heterogeneous effects of education on wage, Figure 1 reports the quantile-specific returns to education from $\tau = 0.1$ to $\tau = 0.9$ (with an interval of 0.05). Our estimated returns pattern over the wage distribution is very similar to those found by the existing literature for other countries (Fasih *et al*., 2012; Harmon *et al*., 2003; Hartog *et al*., 2001).

{ Place Table 2 about here }

{ Place Figure 1 about here }

Table 3 presents TSLS estimates of model (1). As an instrument for years of education we use the expansion in compulsory education that took place in some cantons after 1970. The returns to education estimated by TSLS are slightly higher than OLS estimates, with a point estimate (standard error) of 9.89 percent (0.019). This is a typical result in the literature on returns to education, and it is usually motivated by measurement error in the education variables (see for example Card, 2001) and local average treatment effects (Imbens, 2010). Using a different instrument and a combination of two instruments (see Appendix D) does not affect the estimated return largely, but there is a gain in precision of the education coefficient and R-squared when using multiple instruments.

The coefficient on the instrumental variable in both the reduced form and first stage has the expected sign and it is highly significant. In the first stage model (column 3 of Table 3); our instrument has a positive and significant effect on years of education. This is in line with the expectations discussed in Section 4.3, and consistent with the existing studies that used similar instruments (Braga *et al*., 2013; Brunello *et al*., 2013; Fang *et al*., 2012). In our specific case, the reforms increased educational attainment by 0.35 years on average, whereas previous studies estimated an effect of about half a year. The test for excluded instruments has an F-statistic of 40.75, which is well beyond the canonical 10 (Staiger and Stock, 1997). We are therefore confident about the strength of the instrumental variable. We also reject the null hypotheses of under-identification and weak identification for our instrument (Kleibergen-Paap statistic).

{ Place Table 3 about here }

Table 4 shows the IVQR estimates of model (1). With this regression analysis we causally estimate the impact of education on wage at a given quantile of the wage distribution. Similar to QR estimates, IVQR results also suggest that the returns to schooling vary substantially over the wage distribution, and this heterogeneity is most apparent in the IVQR estimates. While both QR and IVQR approaches indicate that returns to education are heterogeneous, the shapes of the estimated returns over the quantiles are very different. As several previous studies, QR estimates exhibit increasing returns to education, indicating that returns are higher at higher quantiles of the wage distribution. However, if education is endogenous to the wage equation of model (1), conventional QR does not consistently estimate the causal effect of education on wage. IVQR estimates are instead (asymptotically) consistent under endogeneity and show that returns are decreasing over the wage distribution.

In particular, the return to education estimated by IVQR is 18.31 percent at the first decile, decreasing to 9.61 percent at the median, and going down to an insignificant -0.38 percent at the last decile of the wage distribution. These results indicate that the largest gains to additional years of education accrue to individuals at the low end of the wage distribution. Figure 2 provides a graphical illustration of these results from $\tau = 0.1$ to $\tau = 0.9$ with an interval of 0.05, by comparing QR estimates with IVQR estimates. Interestingly, the reduced-form quantile IV approach produces qualitatively similar point estimates and distributional pattern to the structural IVQR approach, which indicates that our substantive results are not sensitive to the estimation procedure.

The IVQR estimates are consistent to the theoretical expectations we formulated previously. As the quantile index τ can be viewed as a measure of unobserved individual ability, IVQR results are in line with the argument that individuals acquire education up to the point where the cost equals the rate of return, and costs depend negatively on ability (Card, 1999). In this setting, we would expect the returns to education to be decreasing in ability, with the lower ability individuals having the highest return to education—which is exactly the pattern estimated by IVQR. Moreover, interpreting the quantile index as an ability measure is also consistent with the notion that individuals with higher ability would generate high earnings regardless of their educational level. On the other hand, individuals with lower unobserved ability would gain more from the training provided by formal education.

{ Place Table 4 about here }

{ Place Figure 2 about here }

5.2 Returns to Education over the Wage Distribution between Types of Education

We now focus on the comparison between educational paths. Table 5 gives an overview of the OLS and QR estimates of model (2). Column 1 of Table 5 presents OLS regressions, which estimate a return to vocational education of 6.82 percent and a return to academic education of 7.13 percent. These coefficients gather the effect of an extra year of vocational (academic) education on wage, filtering out the effect of compulsory schooling. By performing an F-test, we reject the null hypothesis of equal coefficients (p-value = 0.001). This means that, at the mean, the effect of one additional year of academic education on wage is larger than the effect on one additional year of vocational education. The question is whether modeling on average loses some important features of this comparison, and we thus bring the discussion into a distributional framework.

Columns 2-6 of Table 5 present the QR estimates for model (2) at various quantiles of the wage distribution. The first result is that, as in Table 2, returns to both vocational and academic educations are increasing in the quantiles of the wage distribution. However, the increasing pattern and the magnitude of the estimated effects are significantly different. At the lower quantiles of the wage distribution, vocational education has a statistically significant return premium in comparison to academic education. From the fourth decile on, the situation is reversed: academic education has higher returns for one additional year of schooling. Thus, in the upper part of the wage distribution, academic education brings a significant premium compared to vocational education.

In particular, at the bottom decile the return to one extra year of vocational education is 4.99 percent, whereas the return to one additional year of academic education is only 4.11 percent. We reject the null hypothesis of equal coefficients at each level of significance (p-value = 0.000). At the third decile the situation is different, with an estimated return of about 6.4 percent for both academic and vocational tracks (p-value = 0.645). At the median, the returns to vocational and academic educations are 6.91 percent and 7.32 percent, respectively. Similarly to OLS, at the median we reject the null hypothesis of equal coefficients, with a p-value of 0.000. At the top decile, academic education brings a return of 9.61 percent, while vocational education has an estimated return of 8.28 percent. The difference between the estimated coefficients is statistically significant (p-value = 0.000). Figure 3 provides graphical support complementing the results of Table 5 we just discussed, comparing OLS estimates with QR estimates across the entire wage distribution, estimated from $\tau = 0.1$ to $\tau = 0.9$ (with an interval of 0.05).

{ Place Table 5 about here }

{ Place Figure 3 about here }

To have a better understanding of the academic premium, we rewrite model (2) as a function of the difference between the two educational tracks, with vocational education as reference category. While doing so impedes us to see the pattern of vocational and academic educations separately, the transformation allows estimating confidence intervals for the academic premium. Figure 4 plots the academic premium over the wage distribution, along with its 95-percent confidence intervals.

{ Place Figure 4 about here }

One potential explanation behind these results is the skill formation of vocational and academic educations. In fact, while vocational education system provides a set of skills that are rather specific to the job the apprentices are learning (Busemeyer and Trampusch, 2012), in the case of academic education the exploitation of the acquired skills strongly depends on the type of job workers are doing (Dearden *et al*., 2002). Besides, vocational education is probably a better fit for students at the lower part of the wage distribution, because those students learn contents that better match and complement their innate abilities (Rosenbaum and Rosenbaum, 2013). As a consequence, at the lower quantiles of the wage distribution vocational education brings a return premium, because individuals with an academic education in this part of the distribution have a relative disadvantage in the job they are performing. Conversely, at some point in the wage distribution (in our case $\tau = 0.4$), academic education starts generating a return premium in comparison to vocational education because workers have the capacity of fully exploit their skills in the labor market.

Given that we are more interested in the presence of heterogeneity and because we did not find appropriate instrumental variables for both academic and vocational education, we do not claim that the estimated effects in the between-within path comparison are causal. We nevertheless performed some simple two-stage quantile regressions[13] that show lower returns to academic education at the bottom of the wage distribution and a return premium of the academic path in the upper part of the wage distribution.

6. Conclusions

In this paper we present evidence of heterogeneous returns to education over the wage distribution. We use instrumental variable quantile regression and data from the Swiss Labor Force Survey to isolate the causal link between education and earnings at different quantiles of the conditional distribution of wages. Our results provide significant evidence that there is no unique causal effect of schooling and that for each individual the effect may be above or below the estimates extensively documented using OLS or TSLS, depending on his position in the wage distribution and his unobservable wage determinants (ability). In particular, while ordinary quantile regression results indicate that returns to education are increasing in the quantile index, once the endogeneity of schooling is taken into account we instead observe higher returns at lower quantiles of the wage distribution. Interpreting the quantile index as a measure of unobserved ability, our findings suggest that less able individuals profit more form one additional year of education. Higher ability individuals have on average higher wages, but the slope of their wage-education profile is flatter than that for lower ability individuals. This would indicate, as discussed by Ashenfelter and Rouse (1998), that abler individuals acquire more schooling because they face lower marginal costs, and not because of higher marginal benefits.

From a methodological point of view, a noteworthy result of our analysis is that a reduced-form quantile IV approach, akin to TSLS, produces qualitatively similar point

[13] Results available upon request, we do not report them because the instruments we use are only *arguably* exogenous. We used the two-stage quantile regression approach of Chen and Portnoy (1996), based on the early work by Powell (1983).

estimates to the structural IVQR approach, which is based on stronger assumptions. Reassuringly, the comparability of these estimates indicates that our core results are not sensitive to the estimation procedure.

In this study we also investigate the potential heterogeneity in the returns within and between different educational paths. Exploiting the unique feature of the Swiss educational system that allows students achieve tertiary education degrees for both academic and vocational tracks, we complement the existing literature by confirming that, at the mean, academic education brings higher returns. However, if we examine the returns over the wage distribution, we observe two relevant—and until now unknown—features of the returns to vocational and academic education. First, we uncover significant heterogeneity within each educational path, with both vocational and academic educations presenting increasing returns over the wage distribution. Second, a comparison between the two tracks reveals that academic education not always yields higher returns. While in the upper part of the wage distribution individuals with an academic background have a return premium compared to individuals with a vocational background, at lower quantiles vocational education has higher returns than academic education. These results imply that answering the question whether academic education yields higher returns in the labor market compared to academic education is not as easy as it might have seemed in the past from descriptive statistics or mean regression. In fact, the answer depends on the individual's position in the conditional wage distribution.

There are a number of ways in which our work can be extended. First, it would be interesting to analyze the evolution over time of the quantile returns to education, and what impact the returns have on the structure of wages. According to our results, education should have an inequality-reducing effect over time, because individuals with lower ability (i.e., those at the lower quantiles of the wage distribution) seem to profit more form formal education. While appealing, such inquiry is complicated by the fact that the endogeneity and measurement error biases are likely to change over time. Second, researchers and policy makers might be interested in a cross-country comparison to study how the causal returns to education change with different wage distributions and education systems.

Third, one could explore the potential non-linear relationship between education and wages by allowing the returns to differ not only between educational paths but also across education levels, as for example in Buchinsky (1994) and Hartog *et al*. (2001). A last compelling extension to our work could be evaluating the impact of changes in the distribution of education on quantiles of the unconditional (marginal) distribution of wages. This would help to estimate the effect of one additional year of schooling on the entire wage distribution and not at a given quantile only. However, to shed light on this topic, we would need an adaptation of the unconditional quantile regression approach (Firpo *et al*., 2009) to instrumental variables—an adaptation that is currently not available.

To conclude, this study shows that typical estimates of the mean return to education provide a rather incomplete characterization of the impact of education on labor market outcomes and are thus a weak guide for public policy. Similarly, distributional analyses using ordinary quantile regression are also an inappropriate tool to describe the true impact of education on wages, because they do not control for unobserved heterogeneity. Our results suggest that the net impact of education on the long-run distribution of income does not necessarily depend on the initial distribution of ability across the population, and we support empirically the argument that formal education partially compensates for differences in innate abilities and early life conditions.

REFERENCES

Abadie, A., Angrist, J.D., Imbens, G., 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. Econometrica 70, 91-117.

Angrist, J.D., Evans, W.N., 1998. Children and their parents' labor supply: evidence from exogenous variation in family size. American Economic Review 88, 450-77.

Angrist, J.D., Krueger, A. B., 1991. Does compulsory school attendance affect schooling and earnings? Quarterly Journal of Economics 106, 979-1014.

Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton.

Arabsheibani, G. R., Staneva, A., 2012. Returns to education in Russia: where there is risky sexual behavior there is also an instrument. IZA Discussion Paper Series, Working Paper No. 6726.

Arcand, J., D'Hombres, B., Gyselinck, P., 2005. Instrument choice and the returns to education: new evidence from Vietnam. Labor and Demography, 0510011, EconWPA.

Arias, O., Hallock, K.F., Sosa-Escudero, W., 2001. Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twins data. Empirical Economics 26, 7-40.

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

Bettinger, E. P., Kremer, M., Saavedra, J. E., 2010. Are educational vouchers only redistributive? The Economic Journal 120, 204-228.

Bettinger, E. P., Baker, R., 2011. The effects of student coaching in college: an evaluation of a randomized experiment in student mentoring. National Bureau of Economic Research Working Paper Series, Working Paper No. 16881.

Braga, M., Checchi, D., Meschi, E., 2013. Educational policies in a long-run perspective. Educational Policies January 2013, 45-100.

Brunello, G., Fabbri, D., Fort, M., 2013. The causal effect of education on body mass: evidence from Europe. Journal of Labor Economics 31, 195-223.

Buchinsky, M., 1994. Changes in the U.S. wage structure 1963-1987: an application of quantile regression. Econometrica 62, 405-458.

Busemeyer, M. R., Trampusch, C., 2012. The Political Economy of Collective Skill Formation. Oxford University Press, New York.

Butcher, K., Case, A., 1994. The effect of sibling composition on women's education and earnings. Quarterly Journal of Economics 109, 531–563.

Card, D., 1994. Earnings, schooling, and ability revisited. National Bureau of Economic Research Working Paper Series, Working Paper No. 4832.

Card, D., 1999. The causal effect of education on earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, Elsevier, Rotterdam, pp. 1801-1863.

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

Card, D., Krueger, A.B., 1992. Does school quality matter? Returns to education and the characteristics of public schools in the United States. Journal of Political Economy 100, 1-40.

Chen, L., Portnoy, S., 1996. Two-staged regression quantiles and two-staged trimmed least-squares estimators for structural equation models. Communication in Statistics: Theory and Methods 25, 1005-1032.

Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. Journal of Econometrics 132, 491-525.

Chernozhukov, V., Hansen, C., 2008. Instrumental quantile regression: a robust inference approach. Journal of Econometrics 142, 379-398.

Chernozhukov, V., Hansen, C., 2013. Quantile models with endogeneity. Annual Review of Economics, DOI: 10.1146/annurev-economics-080511-110952.

Dearden, L., McIntosh, S., Myck, M., Vignoles, A., 2002. The returns to academic and vocational qualifications in Britain. Bulletin of Economic Research 54, 249-274.

Dickson, M., forthcoming. The causal effect of education on wages revisited. Oxford Bulletin of Economics and Statistics, DOI: 10.1111/j. 1468.0084.2012.00708.x.

Dickson, M., Harmon, C., 2011. Economics returns to education: what we know, what we don't know, and where are going—some brief pointers. Economics of Education Review 30, 1118-1122.

Evans, W. N., Montgomery, E., 1994. Education and health: where there's smoke there's an instrument. National Bureau of Economic Research Working Paper Series, Working Paper No. 4949.

Fang, H., Eggleston, K.N., Rizzo, J.A., Rozelle, S., Zeckhauser, R.J., 2012. The returns to education in China: evidence from the 1986 compulsory leducation law. National Bureau of Economic Research Working Paper Series, Working Paper No. 18189.

Fasih, T., Kingdon, G., Patrinos, H.A., Sakellariou, C., Soderbom, M., 2012. Heterogeneous returns to education in the labor market. The World Bank Human Development Network, Working Paper No. 6170.

Federal Office for Professional Education and Technology, 2012. Vocational and professional education and training in Switzerland. Federal Office for Professional Education and Technology. Federal Department of Economic Affairs, Bern.

Hanushek, E. A., Woessmann, L., Zhang, L., 2011. General education, vocational education, and labor-market outcomes over the life cycle. National Bureau of Economic Research Working Paper Series, Working Paper No. 17504.

Harmon, C., Walker, I., 1999. The marginal and average return to schooling in the UK. European Economic Review 43, 879–887.

Harmon, C., Walker I., 2000. Returns to the quantity and quality of education: evidence for men in England and Wales. Economica 67, 19-35.

Harmon, C., Oosterbeek, H., Walker, I., 2003. The returns to education: microeconomics. Journal of Economic Surveys 17, 115-141.

Hartog, J., Pereira, P. T., Vieira, J. A. C., 2001. Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators. Applied Economics 33, 1021-1037.

Heckman, J.J., Polachek, S., 1974. Empirical evidence on the functional form of the earnings-schooling relationship. Journal of the American Statistical Association 69, 350-354.

Heckman, J.J., Lochner, L.J., Todd, P.E., (2006). Earnings functions, rates of return and treatment effects. In: Hanushek, E., Welch, F. (Eds.), Handbook of the Economics of Education, Elsevier, Amsterdam, pp. 307-458.

Henderson, D. J., Polachek, S. W., Wang, L., 2011. Heterogeneity in schooling rates of return. Economics of Education Review 30, 1202-1214.

Hungerford, T., Solon, G., 1987. Sheepskin effects in the return to education. Review of Economics and Statistics 69, 175-177.

Imbens, G.W., 2010. Better LATE than nothing: some comments on Deaton (2009) and Heckman and Urzua (2009). Journal of Economic Literature 48, 399–423.

Koenker, R., Bassett, G.S., 1978. Regression quantiles. Econometrica 46, 33–50.

Koop, G., Tobias, J. L., 2004. Learning about heterogeneity in returns to schooling. Journal of Applied Econometrics 19, 827-849.

Lemieux, T., 2008. The changing nature of wage inequality. Journal of Population Economics 21, 21-48.

Martins, P. S., Pereira, P. T., 2004 Does Education reduce wage inequality? Quantile regression evidence from 16 countries. Labour Economics 11, 355-371.

Meghir, C., Palme, M., 1999. Assessing the effect of schooling on earnings using a social experiment. IFS Working Paper Series, Working Paper No. W99/10.

Mincer, J.A., 1974. Schooling, Experience, and Earnings. Columbia University Press, New York.

Mwabu, G., Schultz, P.A., 1996. Education returns across quantiles of the wage function: alternative explanations for returns to education by race in South Africa. American Economic Review 86, 335-339.

Powell, J., 1983. The asymptotic normality of two-stage least absolute deviations estimators. Eonometrica 51, 1569-1575.

Rosenbaum, J.E., Rosenbaum, J., 2013. Beyond BA blinders: lessons from occupational colleges and certificate programs for nontraditional students. Journal of Economic Perspectives 27, 153–172.

Saniter, N., 2012. Estimating heterogeneous returns to education in Germany via conditional heteroskedasticity. DIW Discussion Papers, Discussion Paper No. 1213.

Staiger, D., Stock, J. H.,1997. Instrumental variables regression with weak instruments. Econometrica 65, 557–586.

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

Vandenbussche, J., Aghion, P., Meghir, C., 2006. Growth, distance to frontier and composition of human capital. Journal of Economic Growth 11, 97-127.

Wang, L., 2013. How does education affect the earnings distribution in urban China 1995-2002? Oxford Bulletin of Economics and Statistics 75, 435-454.

Weber, B., Wolter, S., 2006. Wages and human capital: evidence from Switzerland. In: Asplund, R., Pereira, P. (Eds.), Returns to Human Capital in Europe—A Literature Review, ETLA, Helsinki, pp. 325-350.

Tables and Figures

Table 1. Descriptive Statistics

Panel A: Full sample	Mean	Std. dev.	Min	Max
Annual wage	81,868	41,824	12,816	390,292
Ln (annual wage)	11.21	0.43	9.46	12.87
Years of education	13.16	2.88	7.00	21.00
Compulsory education	0.12	0.33	0.00	1.00
Vocational education	0.65	0.48	0.00	1.00
Academic education	0.23	0.42	0.00	1.00
Age	40.23	10.09	18.00	60.00
$Age^2/100$	17.20	8.30	3.24	36.00
N	34,744			

	Panel B: Vocational education sample				***Panel C: Academic education sample***			
	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max
Annual wage	77,378	33,677	12,816	390,096	108,592	56,013	12,965	390,292
Ln (annual wage)	11.18	0.38	9.46	12.87	11.47	0.50	9.47	12.87
Years of education	12.85	1.45	10.50	18.00	16.53	2.70	10.00	21.00
Years of compulsory education	8.51	0.58	7.00	9.00	8.59	0.54	7.00	9.00
Years of vocational education	4.34	1.37	3.50	9.00	0.00	0.00	0.00	0.00
Years of academic education	0.00	0.00	0.00	0.00	7.94	2.67	2.33	12.50
Age	40.05	10.27	18.00	60.00	40.44	9.28	18.00	60.00
$Age^2/100$	17.09	8.41	3.24	36.00	17.22	7.80	3.24	36.00
N	22,620				7,877			

Source: Swiss Labor Force Survey, Authors' calculations.

Table 2. Returns to Education, OLS and QR Estimates

	OLS (1)	$\tau = 0.1$ (2)	$\tau = 0.3$ (3)	$\tau = 0.5$ (4)	$\tau = 0.7$ (5)	$\tau = 0.9$ (6)
Years of education	0.0673***	0.0388***	0.0596***	0.0687***	0.0764***	0.0889***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.053***	0.037***	0.042***	0.048***	0.054***	0.069***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
$Age^2/100$	-0.051***	-0.039***	-0.042***	-0.046***	-0.051***	-0.064***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Constant	9.118***	9.537***	9.338***	9.200***	9.083***	8.803***
	(0.029)	(0.054)	(0.029)	(0.027)	(0.029)	(0.044)
Year dummies	YES	YES	YES	YES	YES	YES
(Pseudo) R-squared	0.292	0.060	0.149	0.206	0.239	0.232
N	34,744	34,744	34,744	34,744	34,744	34,744

* P < 0.10, ** P < 0.05, *** P < 0.01

Standard errors in parenthesis (robust standard errors for OLS, bootstrapped standard errors for QR).

The dependent variable is the natural logarithm of the annual wage.

Source: Swiss Labor Force Survey, Authors' calculations.

Table 3. Returns to Education, TSLS Estimates

	OLS (1)	Reduced Form (2)	First Stage (3)	Second Stage (4)
Years of education	0.0673***			0.0989***
	(0.001)			(0.019)
Age	0.053***	0.065***	0.174***	0.048***
	(0.001)	(0.002)	(0.011)	(0.003)
$Age^2/100$	-0.051***	-0.066	-0.211***	-0.046***
	(0.002)	(0.002)	(0.014)	(0.004)
Constant	9.118***	9.748***	10.192***	8.725***
	(0.029)	(0.033)	(0.228)	(0.214)
IV – Education expansion		0.034***	0.346***	
		(0.007)	(0.054)	
Year dummies	YES	YES	YES	YES
R-squared	0.292	0.095	0.018	0.249
N	34,744	34,744	34,744	34,744
Test for excluded instruments				
F-statistic			40.75***	
Under-identification test				
Kleibergen-Paap rk LM statistic				39.90***
Weak identification test				
Kleibergen-Paap rk F statistic				40.75***

* P < 0.10, ** P < 0.05, *** P < 0.01

Robust standard errors in parenthesis.

In columns (1), (2) and (4) the dependent variable is the natural logarithm of the annual wage, in column (3) the dependent variable is years of education.

Source: Swiss Labor Force Survey, Authors' calculations.

Table 4. Returns to Education, IVQR Estimates

	TSLS (1)	$\tau = 0.1$ (2)	$\tau = 0.3$ (3)	$\tau = 0.5$ (4)	$\tau = 0.7$ (5)	$\tau = 0.9$ (6)
Years of education	0.0989***	0.1831***	0.1690***	0.0961***	0.0656***	-0.0038
	(0.019)	(0.017)	(0.013)	(0.010)	(0.010)	(0.015)
Age	0.048***	0.005	0.026***	0.046***	0.055***	0.107***
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
$Age^2/100$	-0.046***	0.005	-0.022***	-0.045***	-0.052***	-0.105***
	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)	(0.004)
Constant	8.725***	7.781***	8.048***	8.828***	9.178***	9.331***
	(0.214)	(0.076)	(0.055)	(0.042)	(0.043)	(0.067)
Year dummies	YES	YES	YES	YES	YES	YES
Reduced Form Estimates	0.034***	0.063***	0.046***	0.037***	0.035***	-0.001
	(0.007)	(0.012)	(0.007)	(0.009)	(0.011)	(0.017)
N	34,744	34,744	34,744	34,744	34,744	34,744

* $P < 0.10$, ** $P < 0.05$, *** $P < 0.01$

Standard errors in parenthesis. The dependent variable is the natural logarithm of the annual wage.

Source: Swiss Labor Force Survey, Authors' calculations.

Table 5. Vocational Education Compared to Academic Education, OLS and QR Estimates

	OLS (1)	$\tau = 0.1$ (2)	$\tau = 0.3$ (3)	$\tau = 0.5$ (4)	$\tau = 0.7$ (5)	$\tau = 0.9$ (6)
Compulsory education	-0.032***	-0.062***	-0.038***	-0.028***	-0.018***	-0.015**
	(0.003)	(0.006)	(0.003)	(0.003)	(0.004)	(0.007)
Vocational education	0.0682***	0.0499***	0.0639***	0.0691***	0.0750***	0.0828***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Academic education	0.0713***	0.0411***	0.0643***	0.0732***	0.0823***	0.0961***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.045***	0.029***	0.035***	0.041***	0.046***	0.058***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.003)
$Age^2/100$	-0.043***	-0.031***	-0.035***	-0.039***	-0.043***	-0.053***
	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)
Constant	10.133***	10.529***	10.302***	10.177***	10.062***	9.941***
	(0.045)	(0.082)	(0.043)	(0.041)	(0.046)	(0.082)
Year dummies	YES	YES	YES	YES	YES	YES
(Pseudo) R-squared	0.309	0.072	0.163	0.220	0.250	0.243
F-statistic $\hat{\delta}_V = \hat{\delta}_A$	11.71***	26.71***	0.21	22.87***	60.40***	58.19***
N	34,744	34,744	34,744	34,744	34,744	34,744

* P < 0.10, ** P < 0.05, *** P < 0.01
Standard errors in parenthesis (robust standard errors for OLS, bootstrapped standard errors for QR).
The dependent variable is the natural logarithm of the annual wage.
Source: Swiss Labor Force Survey, Authors' calculations.

Figure 1. Returns to Education over the Quantiles, QR Estimates

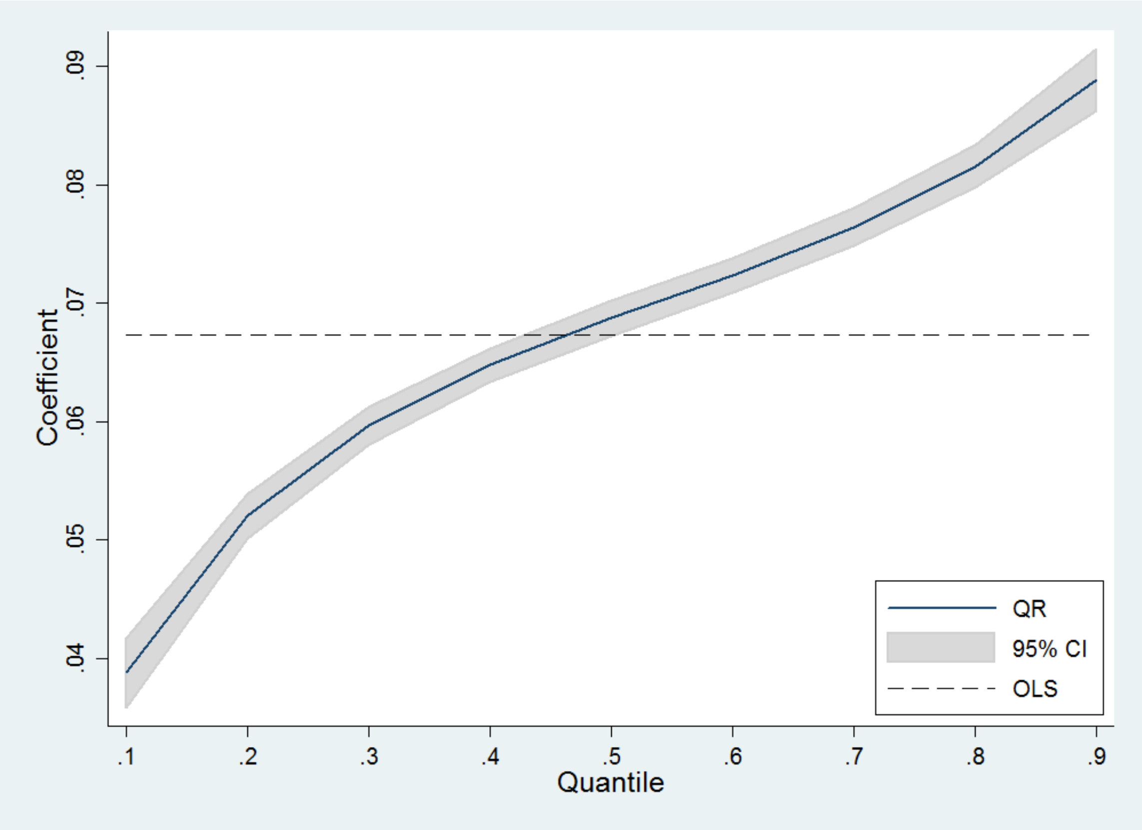


Figure 2. Returns to Education over the Quantiles, IVQR Estimates

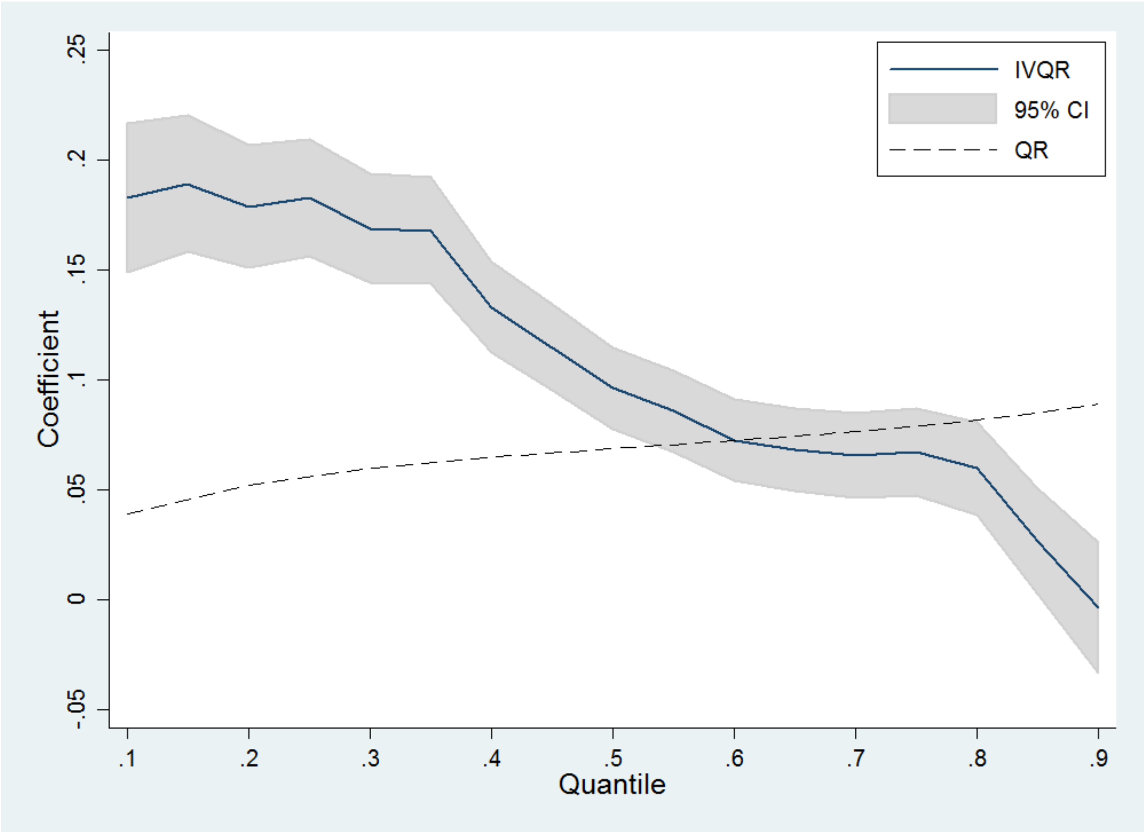


Figure 3. Returns to Vocational and Academic Educations over the Quantiles, QR Estimates

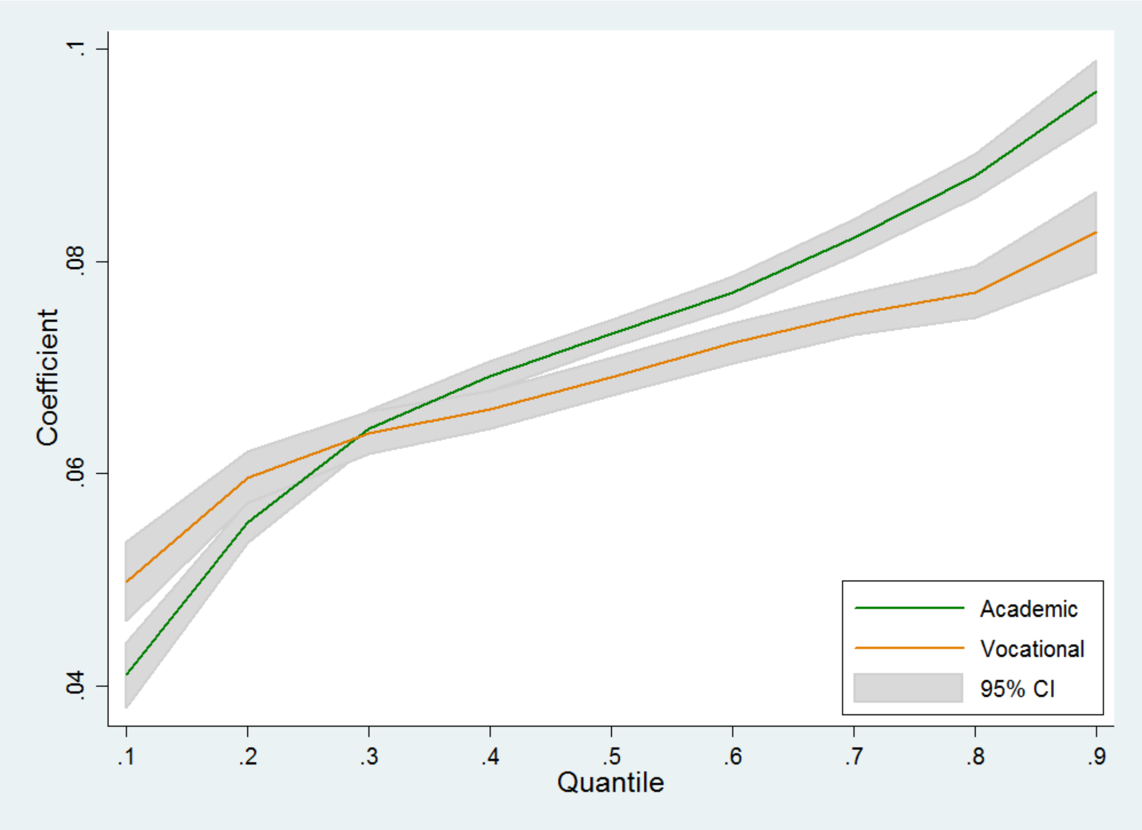
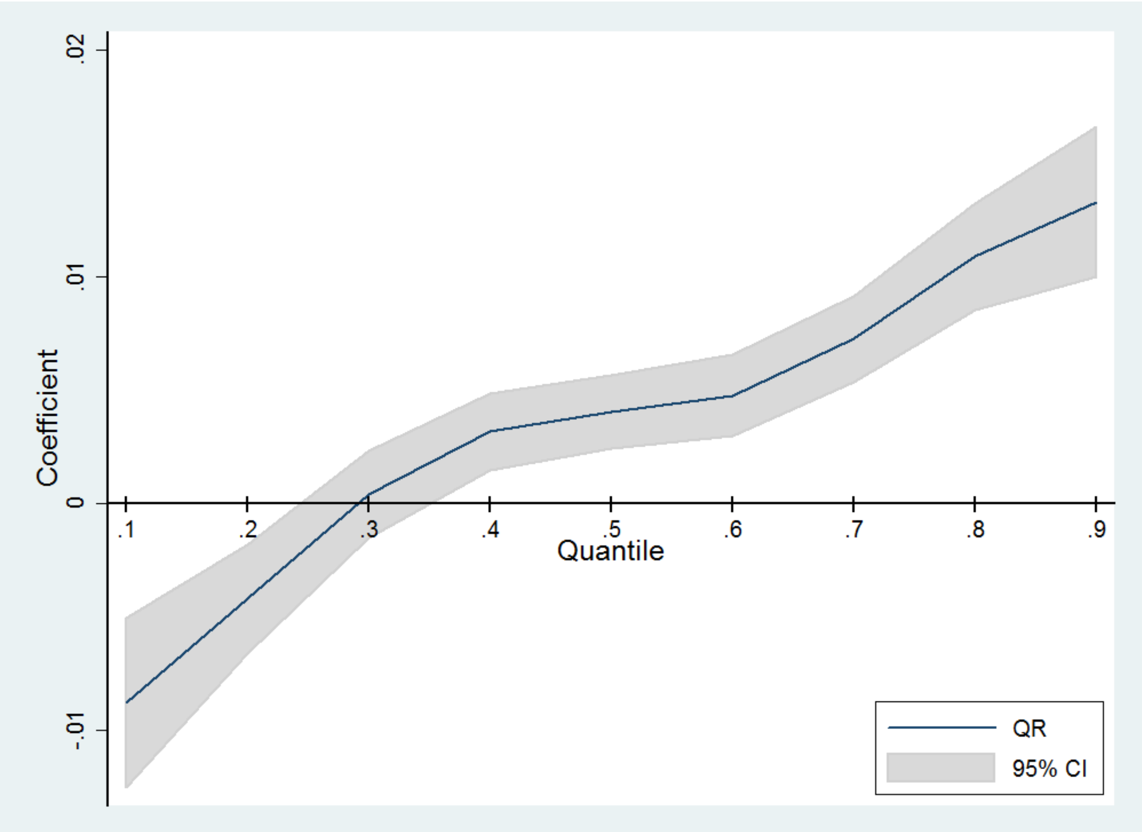
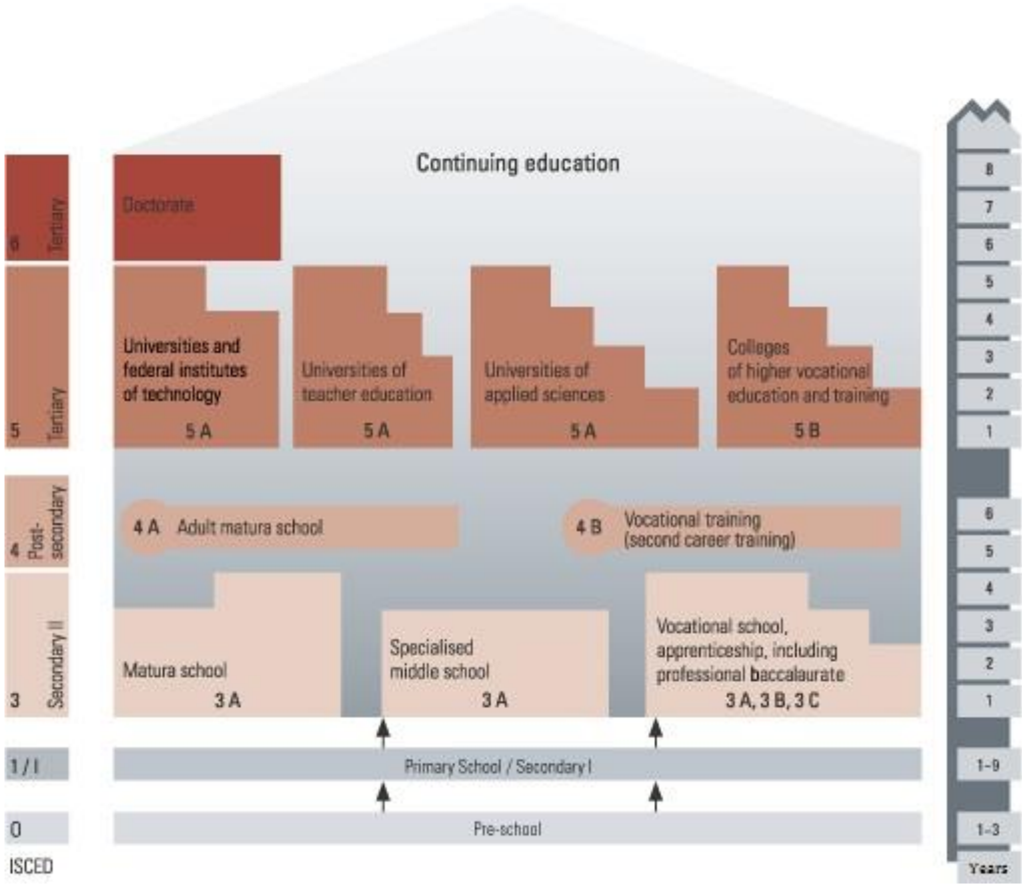


Figure 4. The Academic Premium over the Quantiles, QR Estimates



Appendix A. The Swiss Education System



Source: Swiss State Secretariat for Education and Research.

Appendix B. Sample Construction

Initial sample (SLFS, 2000-2009)	160,925
Sample restrictions	
Males	74,871
Fully employed	47,347
Age between 18 and 60	44,670
Not in education or gap year	42,612
Wage not missing	35,095
99% the wage distribution	34,744
Analytic sample	34,744

Source: Swiss Labor Force Survey, Authors' calculations.

Appendix C. Summary of the Concordat of 1970

		Expansion in compulsory education					*Change in school year start*				
Canton	School-entry age	Reform adopted	Actual year of reform	Years before reform	Years after reform	First cohort affected	Reform adopted	Actual year of reform	Semester start before	Semester start after	First cohort affected
Zurich	6	Yes	1977	8	9	1971	Yes	1989	Spring	Fall	1974
Bern	6	No		9	9		Yes	1989	Spring	Fall	1974
Luzern	6	Yes	1985	8	9	1979	No		Fall	Fall	
Uri	7	Yes	1977	7	9	1970	No		Fall	Fall	
Schwyz	7	Yes	1992	7	9	1985	Yes	1989	Spring	Fall	1975
Obwalden	7	Yes	1992	7	9	1985	No		Fall	Fall	
Nidwalden	6	Yes	1992	7	9	1986	No		Fall	Fall	
Glarus	6	Yes	1983	8	9	1977	Yes	1989	Spring	Fall	1975
Zug	7	Yes	1990	8	9	1983	Yes	1973	Spring	Fall	1958
Fribourg	7	No		9	9		No		Fall	Fall	
Solothurn	7	Yes	1970	8	9	1963	Yes	1989	Spring	Fall	1973
Basel-Stadt	6	No		9	9		Yes	1989	Spring	Fall	1974
Basel-Land	6	Yes	1980	8	9	1974	Yes	1989	Spring	Fall	1975
Schaffhausen	6	Yes	1982	8	9	1976	Yes	1989	Spring	Fall	1975
Appenzell A.	6	No	1981	8	9	1975	Yes	1989	Spring	Fall	1975
Appenzell I.	6	Yes	1984	7	9	1978	Yes	1989	Spring	Fall	1976
St. Gallen	6	Yes	1983	8	9	1977	Yes	1989	Spring	Fall	1975
Graubünden	7	No		9	9		No		Fall	Fall	
Aargau	7	Yes	1982	8	9	1975	Yes	1989	Spring	Fall	1974
Thurgau	6	Yes	1980	8	9	1974	Yes	1989	Spring	Fall	1974
Ticino	6	No		9	9		No		Fall	Fall	
Vaud	7	No		9	9		Yes	1973	Spring	Fall	1957
Valais	7	Yes	1987	8	9	1980	No		Fall	Fall	
Neuchâtel	6	No		9	9		Yes	1973	Spring	Fall	1958
Genève	6	No		9	9		No		Fall	Fall	
Jura	6	No		9	9		Yes	1989	Spring	Fall	1974

Source: Authors' research and calculations.

Appendix D. Alternative Instrument and Overidentified TSLS

	First Stage (1)	Second Stage (2)	First Stage (3)	Second Stage (4)	First Stage (5)	Second Stage (6)	First Stage (7)	Second Stage (8)
Years of education		0.0989***		0.1042***		0.1011***		0.1017***
		(0.019)		(0.026)		(0.017)		(0.014)
Age	0.174***	0.048***	0.166***	0.047***	0.185***	0.048***	0.179***	0.048***
	(0.011)	(0.003)	(0.011)	(0.004)	(0.011)	(0.003)	(0.011)	(0.003)
$Age^2/100$	-0.211***	-0.046***	-0.201***	-0.045***	-0.221***	-0.045***	-0.212***	-0.045***
	(0.014)	(0.004)	(0.013)	(0.005)	(0.014)	(0.004)	(0.014)	(0.003)
Constant	10.192***	8.725***	10.348***	8.669***	9.876***	8.703***	9.978***	8.696***
	(0.228)	(0.214)	(0.223)	(0.288)	(0.240)	(0.188)	(0.241)	(0.159)
IV_1 – Education expansion	0.346***				0.302***		0.620***	
	(0.054)				(0.055)		(0.083)	
IV_2 – Shift in school-year start			0.221***		0.178***		0.267***	
			(0.043)		(0.044)		(0.047)	
$IV_1 \cdot IV_2$							-0.539***	
							(0.104)	
Year dummies	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.018	0.249	0.018	0.233	0.019	0.243	0.019	0.241
N	34,744	34,744	34,744	34,744	34,744	34,744	34,744	34,744
Test for excluded instruments								
F-statistic	40.75***		26.69***		28.37***		26.53***	
Under-identification test								
Kleibergen-Paap rk LM-statistic		39.90***		26.58***		55.91***		77.95***
Weak identification test								
Kleibergen-Paap rk F-statistic		40.75***		26.69***		28.37***		26.53***
Hansen J statistic								
P-value						0.860		0.982

* P < 0.10, ** P < 0.05, *** P < 0.01

Robust standard errors in parenthesis. In even columns the dependent variable is the natural logarithm of the annual wage, in odd columns the dependent variable is years of education.

Source: Swiss Labor Force Survey, Authors' calculations.

Appendix E. Overidentified IVQR

	$\tau = 0.25$ (1)	$\tau = 0.5$ (2)	$\tau = 0.75$ (3)	$\tau = 0.25$ (4)	$\tau = 0.5$ (5)	$\tau = 0.75$ (6)
Years of education	0.1779***	0.0901***	0.0818***	0.1639***	0.0850***	0.0921***
	(0.032)	(0.023)	(0.025)	(0.025)	(0.019)	(0.021)
Age	0.020***	0.047***	0.056***	0.023***	0.047***	0.055***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)
$Age^2/100$	-0.014**	-0.045***	-0.053***	-0.017***	-0.045***	-0.051***
	(0.007)	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)
Constant	8.143***	8.910***	8.979***	8.124***	8.972***	8.882***
	(0.338)	(0.250)	(0.270)	(0.281)	(0.207)	(0.228)
Year dummies	YES	YES	YES	YES	YES	YES
IV_1 – Education expansion	x	x	x	x	x	x
IV_2 – Shift in school-year start	x	x	x	x	x	x
$IV_1 \cdot IV_2$				x	x	x
N	34,744	34,744	34,744	34'744	34,744	34,744

* P < 0.10, ** P < 0.05, *** P < 0.01
Standard errors in parenthesis. The dependent variable is the natural logarithm of the annual wage.
Source: Swiss Labor Force Survey, Authors' calculations