UNIVERSITÀ DI PISA
DIPARTIMENTO DI ECONOMIA E MANAGEMENT

Ph.D. Thesis
Ph.D. Program in Economics (SECS P/01)

The Determinants of the Italian Wages

Candidate:
Cecilia Gabbriellini

Supervisor:
Prof. Davide Fiaschi

December 2015

To my mother

Acknowledgements

I am deeply grateful to my supervisor Prof. Davide Fiaschi for detailed and constructive comments and suggestions that have been precious for my research.

A particular thanks to Irene Brunetti and Angela Parenti for their helpful suggestions and comments.

I owe my deepest gratitude to my parents, to my sister, Barbara, and my brother, Giovanni, for their moral encouragement.

Finally, I owe a special thanks to my husband, Fabio, the most beautiful surprise during my Ph.D. study, who has given his essential moral support.

Contents

Introduction

Education is one of the most important components of individual human capital (Becker, 1993) thus a significant determinant of wages. Indeed, the estimation of economic return of schooling is a relevant parameter of interest in economics studies and in public policy design.

The evaluation of policies that promote education is a central research question. The increase in wages due to additional schooling, what is usually called the return to schooling, is a main component of the benefits of the proposed policies. In fact, to the policy maker perspective, it is crucial to understand if the higher wages observed for better educated people are determined only by their higher education level or if they reflect inherent ability differences that correlate with educational attainment. Therefore, treating schooling as a way to increase market productivity it is important to understand if any increase in public spending for education is meaningful for people.

The aim of this work is to estimate the determinants of the wage function in Italy, focusing on the crucial role of education, taking into consideration even the impact of years of experience (training on the job and learning by doing activities), controlling for individual characteristics, and sectorial and geographical variables.

The benchmark model for the development of empirical estimation of the returns to education is the relationship derived by Mincer (1974) between log hourly wages, schooling and experience. However, the empirical estimation of the causal returns is not an easy task because simple regressions between wage and schooling does not report causal returns to education (and produce biased estimates) as the schooling variable is likely to be endogenous due to omitted variable, namely ability. To overcome this problem, we apply instrumental variable regressions.

In all the chapters, the empirical analysis is carried out using a representative sample of Italian households, drawn from the Bank of Italy's Survey of Household Income and Wealth (SHIW), for the period 1995-2012.

In the first chapter, in line with previous literature, we find that ordinary least squares (OLS) under-estimate the return to schooling. Considering the endogeneity of schooling, the return to an additional year in school increases. In addition, in the period considered, the findings show that the returns to schooling have changed from 5.4 percent to 7.9 percent. The highest level is recorded in 2006 and the lowest in 2012 thus the advantage to invest in education is decreasing in Italy. Moreover, a relative convenience to work in the public sector emerges as well as an evidence of a gender pay gap, in favor of men for all the period considered.

Understanding how the returns estimates vary with the level of schooling attainment is important. Most empirical studies in this area assume log wages linear or quadratic in years of education and year of experience. In the second chapter, we remove the hypothesis of homogeneity of the return to education and we estimate the wage-schooling and wage-experience profile to take into account all the shape of these variable and their possible non-linearity. In addition, to analyze the effect of endogeneity on the non-monotonicity of the marginal rate of return to education, we use a control function approach for a semiparametric estimation, as suggested in Blundell and Powell (2003). Results show that the wage–schooling relationship is non-linear. This implies that returns to education depend on the level of schooling. In particular, increasing returns are evident for workers until 8 years of schooling (junior high school), from 1995 to 2004; however, in the following years they show a flat pattern. If we consider worker with almost 13 years of schooling (secondary school), the marginal effects across year continue to increase. On the other hand, decreasing returns are observed for workers with 18 years of schooling (tertiary education), from 2008 to 2012.

Several studies focused on the estimation of the average impact of schooling, experience and other variables on wage without investigate if they affect individuals differently over the wage distribution. The aim of third chapter is to understand if individuals in different quantile of the wage distribution are differently affected by these determinants. Indeed, if returns to schooling are heterogeneous along the wage distribution, schooling can have an impact upon wage inequality. In a simple human capital model, wage inequality can increase because returns to education and experience increase, or because residual or within-group inequality increases (Lemieux, 2008). In the case that returns are increasing from the lower to the higher end of the wage distribution, it can be interpreted as an indication that ability and education (or skills) are complement between them, and more able workers can benefit from additional investment in education.

Instrumental quantile regression methods is the appropriate tool to describe the impact of education on wages distribution, controlling for unobserved heterogeneity and endogeneity. Our results show that, while returns to education are positive everywhere, there exists a large degree of heterogeneity in returns to education across the wage distribution. In particular, gains are higher for individuals in the upper tail of the wages distribution than for those in the lower tail. This means that education have an inequality-increasing effect over time, because individuals with high ability, those at the upper quantile of the wage distribution, seem to benefit more from formal education. Therefore, the results suggest that the impact of education on the distribution of wages depends on the initial distribution of ability across population and, consequently, formal education does not compensate for differences in innate abilities and early life conditions.

Chapter 1

An Estimate of the Wage Function in Italy

Abstract. Education can be seen as an investment with returns incorporates in the future wages. The general model points out that higher individual education implies higher individual wages. Many studies have tested this relationship, in different countries. Using data come from the 1995 to 2012 waves of the Bank of Italy's Survey of Household Income and Wealth, we estimate the determinants of the wage function, focusing on the role of schooling and labor market experience.

The findings highlight the evidence of returns to schooling that have changed over the period considered and are between 5.4 percent and 7.9 percent, recording the highest level for 2006 and the lowest in 2012. Therefore, the advantage to invest in education is decreasing in Italy. Moreover, a relative convenience to work in the public sector emerges. Finally, there is evidence of a gender pay gap, in favor of men for all the period considered.

1.1 Introduction

Education is one of the most important components of individual human capital (Becker, 1993) thus a significant determinant of wages. The estimation of the economic return to education has been one of the predominant areas of analysis in applied economics for over 50 years, in both micro and macroeconomics.

The analysis of education has been driven by the concept of human capital, pioneered in the works of led economist such as Gary Becker, Jacob Mincer and Theodore Schultz. According the human capital theory, education is seen as an investment of current resources to get future returns.

The estimation of economic return of schooling is a relevant parameter of interest in economics studies and in public policy design. Indeed, a huge body of literature focus in the estimation of returns to education. This interest is due by the link between schooling and productivity growth (Lucas, 1988). Moreover, economists studying inequality and poverty seek to learn how schooling increases the incomes of the poor. Therefore, the evaluation of policies that promote education is a central research question.

The increase in wages due to additional schooling, what is usually called the return to schooling, is a main component of the benefits of the proposed policies. In fact, to the policy maker perspective, it is crucial to understand if the higher wages observed for better educated people are determined only by their higher education level or if they reflect inherent ability differences that correlate with educational attainment. Therefore, treating schooling as a way to increase market productivity it is important to understand if any increase in public spending for education is meaningful for people.

The benchmark model for the development of empirical estimation of the returns to education is the relationship derived by Mincer (1974) between log hourly wages, schooling and experience. The original Mincer equation assumes linear effect on wages of each year of education regardless of the attainment level.

Since the pioneer work of Mincer (1974) who has written the methodological foundation to estimate wage equations, a huge body of works were dedicated to finding the causal return to education. The causal return to education is the extra amount of wage that a randomly selected worker receives from an additional year of education. As explained before, knowing the causal return is important for policy makers, because it directly informs about the utility of educational programs in terms of monetary payoffs for its beneficiaries and for the economic system at all. However, the empirical estimation of the causal returns is not an easy task i.e. the simple regressions between wage and schooling does not report causal returns to education (and produce biased estimates) as the schooling variable is likely to be endogenous due to omitted variable, namely ability.

One well-established route to circumvent the endogeneity problem is to use instrument variable (IV) methods. These methods, while theoretically appealing, are not easy to implement in practice as they rely on the availability of valid and significant instruments.

The research question of this chapter is to investigate how years of education, experience and other variables affect wages in Italy. In addition, we want understand if the impact of these variables on wages vary over time.

Since education can be seen as a private decision to invest in human capital, we calculate the internal rate of return to this private investment. Moreover, we take into account differential effects of different educational level: vocational, upper-secondary and tertiary education. The data come from the Survey of Household Income and Wealth (SHIW) carried out by the Bank of Italy, covering the period from 1995 to 2012 where information about education, wage and demographics characteristics are collected at individual and household level.

Our results shows that returns to education have changed over the considered period, varying between 5.4 percent and 7.9 percent. Considering different sector of employment, a relative convenience to work in the public sector emerges. In addition, there is an evidence of a gender pay gap, in favor of men for all the period considered. When the kind of school attended is taken into consideration, the returns to education increase with higher levels of educational attainment.

The reminder of the paper is organized as follows. Section 2 presents the theoretical background of the wages equation to be estimated. Section 3 describes the dataset used in the empirical estimation and the characteristics of the sample. Section 4 reports the estimates of the effect of schooling, experience and other variables on wages. Finally, section 5 summarizes and concludes.

1.2 Theoretical Framework for the Empirical Analysis

The theoretical framework underlying most empirical studies on the determinants of wages, and the related estimate of the return to schooling, is the model of accumulation of human capital developed by Schultz (1961), Becker (1962) and Mincer (1958, 1974). In particular, Mincer (1974) focuses on the life-cycle dynamics of earnings and on the relationship between (observed) earnings, earnings capacity (proportional to the individual stock of human capital) and investment in earnings capacity (human capital); such investments can regard both formal schooling and on-the-job training (learning by doing).

Earnings will be a function of earnings capacity net of the costs of investment in earnings capacity. In particular, let E_t be the earnings capacity at time *t*. Earnings capacity can be increased by investment in human capital. To maintain as simple as possible the analysis, investments are expressed as a fraction of earnings capacity:

$$C_t = k_t E_t, \tag{1.1}$$

where k_t is the fraction of earnings capacity invested at time *t*. Let ρ_t be the return on investments made at time *t*. Then:

$$E_{t+1} = E_t + C_t\rho_t - \delta_t E_t = E_t(1 + k_t\rho_t) - \delta_t E_t, \tag{1.2}$$

where δ_t is the depreciation on obsolescence of earnings capacity (see Rosen, 1974). Recursive substitution yields:

$$E_t = \prod_{j=0}^{t-1}\left(1 + \rho_j k_j - \delta_t\right)E_0, \tag{1.3}$$

where E_0 is the earnings capacity, independent of schooling and experience.

Formal schooling is defined as the numbers of years spent in full-time investment ($k_t = 1$). Assume that the rate of return on formal schooling of length s is constant for all years of schooling and equal to ρ_s and that formal schooling takes place at the beginning of life, i.e. $\rho_t = \rho_s \; \forall \; t = 0, \dots, s$. Therefore, assume that the rate of return to post-school investment is constant over time and equals ρ_{ps}, i.e. $\rho_t = \rho_{ps} \; \forall \; t \geq s$. Then, we can write:

$$\ln E_t = \ln E_0 + s\ln(1 + \rho_s) + \sum_{j=s}^{t-1}\ln\left(1 + \rho_{ps}k_j - \delta_t\right), \tag{1.4}$$

which yields the approximate relationship (for small ρ_s and ρ_{ps})[1]:

$$\ln E_t \approx \ln E_0 + \rho_s s + \rho_{ps}\sum_{j=s}^{t-1} k_j - \sum_{j=s}^{t-1}\delta_j. \tag{1.5}$$

To establish a relationship between earnings capacity and years of experience, Mincer (1974) approximates the Ben-Porath (1967) model and further assumes a linearly declining rate of post-school investment in human capital:

$$k_{s+x} = \kappa\left(1 - \frac{x}{T}\right), \tag{1.6}$$

where $\kappa > 0$ is a scale parameter, $x = t - s \geq 0$ is the amount of work experience as of age t. The length of working life, T, is assumed to be independent of years of schooling[2]. Given Equation (1.6), the relationship between earnings capacity, schooling and experience is given by:

[1] See pag.19 of Mincer (1974).
[2] This means that educated workers retire after not educated workers.

$$\ln E_{x+s} \approx \ln E_0 - \kappa\rho_{ps} + \rho_s s + \rho_{ps}\kappa x\left(1+\frac{1}{2T}\right) - \frac{\rho_{ps}\kappa}{2T}x^2 - x\delta, \qquad (1.7)$$

under the assumption that $\delta_j = \delta \ \forall j$. Observed earnings are equal to earnings capacity less investment costs, i.e. $w(s,x) = (1-k_{s+x})E_{x+s}$. Therefore:

$$\ln w(s,x) \approx \ln E_{x+s} - \kappa\left(1-\frac{x}{T}\right)=$$

$$= \ln E_0 - \kappa\rho_{ps} - \kappa + \rho_s s + \left[\kappa\left(\rho_{ps} + \frac{\rho_{ps}}{2T} + \frac{1}{T}\right) - \delta\right]x - \frac{\rho_{ps}\kappa}{2T}x^2=$$

$$= \alpha_0 + \rho_s s + \beta_0 x + \beta_1 x^2, \qquad (1.8)$$

where $\alpha_0 = \ln E_0 - \kappa(1+\rho_{ps})$, $\beta_0 = \kappa\left[\rho_{ps}\left(1+\frac{1}{2T}\right)+\frac{1}{T}\right] - \delta$, $\beta_1 = -\frac{\rho_{ps}\kappa}{2T}$.

Starting from this standard form of the Mincer wages model, it is possible to derive an econometrics model in order to estimate the parameters. Therefore, the log wages are regressed on a constant term, a linear term in years of schooling, and linear and quadratic term in years of labor market experience. In most of applications of the Mincer model, it is assumed that the intercept and slope coefficients are identical across persons. This implicitly assumes that E_0, κ, ρ_s , ρ_{ps} and δ are the same across workers and do not depend on the schooling level. However, Mincer formulates a more general model that allows for the possibility that $E_0, \kappa,\ \rho_s,\ \rho_{ps}$ and δ differ across workers, which produces a random coefficient model:

$$\ln w(s_i, x_i) = \alpha_{0i} + \rho_{si}s_i + \beta_{0i}x_i + \beta_{1i}x_i^2. \qquad (1.9)$$

Denoting $\alpha_0 = E(\alpha_{0i}), \rho_s = E(\rho_{si}), \beta_0 = E(\beta_{0i}), \beta_1 = E(\beta_{1i})$, we can rewrite Equation (1.9) as:

$$\ln w(s_i, x_i) = \rho_s s_i + \beta_0 x_i + \beta_1 x_i^2 + [\alpha_{0i} + (\rho_{si} - \rho_s)s_i + (\beta_{0i} - \beta_0)x_i + (\beta_{1i} - \beta_1)x_i^2], \qquad (1.10)$$

where the terms in brackets are part of the error. Mincer assumes that $\alpha_{0i}, (\rho_{si} - \rho_s), (\beta_{0i} - \beta_0), (\beta_{1i} - \beta_1)$ are independent of (s_i, x_i) which reduces Equation (1.10) to Equation (1.8) in terms of estimations with individual data, i.e:

$$\ln w(s_i, x_i) = \alpha_{0i} + \rho_s s_i + \beta_0 x_i + \beta_1 x_i^2 + \varepsilon_i. \tag{1.11}$$

That is the Mincerian wage equation where ε_i is a mean zero residual with $E(\varepsilon_i | s_i, x_i) = 0$

Mincer derives several implications from the accounting identity model under different assumptions about the relationship between formal schooling and post-school investment patterns. Under the assumption that post-school investment ρ_{ps} are identical across persons and do not depend on the schooling level s, we have that $\frac{\partial \ln w(s_i, x_i)}{\partial s_i \partial x_i} = 0$ and $\frac{\partial \ln w(s_i, x_i)}{\partial s_i \partial t} = \frac{\rho_{ps} \kappa}{T} > 0$. These two conditions imply:

(i) log-wages experience profiles are parallel across schooling levels;

(ii) log-wages age profile diverge with age across schooling levels.

Equation (1.10) highlights how error term ε_i captures unobservable individual effects, as unobserved ability; this also influences schooling decision *s*, and thus induces a correlation between schooling and the error term in the wages function. With endogeneity, the estimation of the return to schooling by ordinary least squares is biased. In literature, the problem has been addressed in different ways. The measures of ability have been incorporated with a proxy variable for unobserved effects, in order to control separately the effect of education and ability (Mendolicchio, 2006). Another solution is to apply within-twins differences in wages and education, assuming that unobserved effects are additive and common within twins so they can be differentiated out by regressing the wage difference within twins against their education differences (Bonjour et al., 2003).

An additional approach deals with the simultaneous relationship between schooling and wages by specifying a two-equation system, which is identified by exploiting instrumental variables that affect *s* but not *w* (Blundell, Dearden and Sianesi, 2001), where family background is used as instruments for schooling. The last approach is the most applied in the literature and will be our strategy to deal with endogeneity.

1.3 Data and Sources

The analysis is based on data drawn from the Bank of Italy's Survey of Household Income and Wealth (SHIW), which reports several socio-economic characteristics of Italian households.

The SHIW is a biannual survey on Italian families with a sample of approximately 8,000 household per year. From 1995 to 2012 observations from nine subsequent surveys are available. In particular, the SHIW contains information both on households (family composition) and on individuals. Moreover, it provides detailed information on several characteristics of workers within each household, such as their net yearly wages, average weekly hours of work and number of months of employment per year, educational attainment (the highest completed school degree), job experience, gender, marital status, sector of employment, household composition, parents background, regions of residence, and town size.

We consider a sub-sample of men and women between 15-64 years old, full time and part time employees, working either in the public or in the private sector and such that information about wages are available. In the analysis, we exclude self-employed because of the low reliability of their declared earnings. As discussed by Brandolini and Cannari (1994), SHIW seems to underestimate the self-employed earnings of about 50 percentage points.

1.3.1 Variables Used in the Analysis

As shown by Equation (1.11), wages, schooling attainment, and working experience of each individual are the key variables in the estimate of Mincer equation.

Mincer equation refers to the (log of) hourly price of labor as the correct measure of worker's wages (LOGY_H), and, indeed, this is the measure used by most empirical studies[3] (Brunello and Miniaci, 1999; Blundell, Dearden and Sianesi, 2005; Ciccone, Cingano and Cipollone, 2006). SHIW contains yearly net wages of taxes and social security contributions. Additional information on the average number of hours worked per week and on the number of months worked per year, can be used to estimate the hourly net wage, which is calculated by yearly net wages divided by months worked multiplied by hours worked each month.

Schooling attainment (SCHOOL) is generally measured by the number of years spent at school. SHIW does not contain information about this number of years, but only on the highest degree attained by individuals. Following a common approach in literature (Vieira, 1999; Brunello and Miniaci, 1999) we calculate the educational attainment of the individual by imputing the number of years required to complete her/his reported maximum level of educational attainment[4]. More precisely, we consider that the (statutory) numbers of years required to obtain a primary and a junior school certificate is 5 and 8 years respectively; instead, for the upper secondary school the number of years ranges from 11 (vocational or technical school) to 13 (classical or scientific studies); finally, for tertiary education, we consider 16, 18 and 21 years for the university diploma, the college degree, and the postgraduate degree (e.g. Ph.D.) respectively. In the analysis of Section 1.4.2, we will also treat education as a categorical variable divided into 4

[3] Hourly wages can be affected by measurement errors because we calculate them as total wages divided by hours of work.

[4] Standard and not actual years of formal schooling are recorded. Since students who fail to reach a standard have to repeat the year, the actual number of years is likely to be underestimated.

categories: no education or primary school or junior high school (COMP_SCHOOL), 3-year vocational school (VOCATIONAL), upper secondary school (UPPER SECONDARY), tertiary education (TERTIARY; including university diploma, college and post-graduate education). It is important to remark that in Italy the statutory number of years can be significantly different from the actual number of years spent to obtain a degree, especially at college because of the high percentage of irregular student.

Many empirical studies use age as a proxy for the (working) experience of individuals. But this choice can be severely biased, especially for young cohorts. Other authors use potential experience, defined as the difference between the current age and the age at the labor market entry, but they ignore the possibility of unemployment or underemployment, again a crucial feature for young cohorts. In this work, we use as proxy for experience (EXPERIENCE), the number of years for which a worker has been paid social security contribution; they should reflect the effective years of training on the job and learning-by-doing activities.

We introduce several control variables in the analysis to account for individual characteristics and for differences in the labor market. A gender dummy (DUMMY_MALE) controls for different wage levels between men and women. Marital status also enter into the analysis as a dummy variable (DUMMY_MARRIED) taking the value 1 if the person is formally married, 0 otherwise. Part-time work is captured through a separate dummy variable (DUMMY_PART_TIME), since the assumption that each working hour makes the same contribution to weekly wages (constancy of the hourly wage) cannot hold across workers with different time status (part time versus full time).

In addition, controls are introduced for family composition, as a proxy for the influence of housework, particularly important in the female labor supply (Heckman and Killingsworth, 1986). We control for the number of components of the family (NCOMP) and for the fact that the individual is the head of his/her household (DUMMY_HOUSEHOLD).

Table 1.1 - Means and standard deviations of the variables used in the empirical analysis for the entire sample (1995, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012)

Variable	Mean	S. d.	Description
LOGY_H	2,265	0,438	Logarithm of the hourly wages less tax
SCHOOL	11,373	3,800	Schooling attainment, that is the number of years spent at school
COMP_SCHOOL	0,383	0,486	Compulsory school: no schooling, primary school and junior high school
VOCATIONAL	0,090	0,288	3-years Vocational degree
UPPER_SECONDARY	0,379	0,485	Upper secondary degree
TERTIARY	0,146	0,354	Tertiary degree
EXPERIENCE	17,683	10,673	Number of years for which it has been paid social security contributions, as a proxy for years of training on the job
DUMMY_MALE	0,578	0,494	Gender dummy
DUMMY_MARRIED	0,647	0,478	Dummy variable for marital status
NCOMP	3,329	1,185	Number of components of the family
DUMMY_HOUSEHOLD	0,475	0,499	Household dummy, that is equal to 1 if the individual is the household of the family
DUMMY_PART_TIME	0,094	0,292	Dummy variable for part time work
DUMMY_AGRICULTURAL	0,034	0,180	Dummy variable for agricultural sector
DUMMY_INDUSTRIAL	0,312	0,463	Dummy variable for industrial sector
DUMMY_PUBLIC	0,320	0,466	Dummy variable for public administration sector
DUMMY_OTHER_SECTOR	0,335	0,472	Dummy variable for other sector
DUMMY_TOWN	0,083	0,275	Dummy variable for the town of residence that has more than 500.000 inhabitants
DUMMY_NORTH	0,501	0,500	Dummy variable for North regions
DUMMY_CENTER	0,214	0,410	Dummy variable for Center regions
DUMMY_SOUTH	0,286	0,452	Dummy variable for South regions
DUMMY_SETT_GEN	0,374	0,484	Dummy variable equal to 1 if the individual works in the same sector of the father and/or of the mother
SCHOOL_F	6,094	4,094	Schooling attainment of the father's worker
SCHOOL_M	5,346	3,711	Schooling attainment of the mother's worker

Controls for sector (DUMMY_AGRICULTURAL for the agricultural sector, DUMMY_INDUSTRIAL for the industrial sector, DUMMY_PUBLIC for the public sector and DUMMY_OTHER_SECTOR for other sector different from the previous ones) should capture potential factor from the demand side of labor market (e.g. imperfectly competitive labor markets). In the same light, we add some controls for the geographical area of residence: one dummy for the town of residence that has more than 500.000 inhabitants (DUMMY_TOWN), and three different dummies for the Italian macro-regions: North, Center and South (DUMMY_NORTH, DUMMY_CENTER and DUMMY_SOUTH)[5].

Table 1.1 reports some descriptive statistics of the main variables used in the empirical analysis for all the waves (wages are expressed in euro 2012).

1.4 Estimates

In a first model, we consider schooling as measured by the years of schooling. In a second step of analysis, we consider separately different level of educational attainment.

1.4.1 Mincer Wage Model with Years of Education

For each available wave, we estimate the Mincer wage model reported in Equation (1.11). However, as discussed in a very large literature reviewed by Card (1995), OLS estimation of the returns to education via Mincer wage Equation are not consistent either because of i) the measurement errors in the schooling variable, and ii) the endogeneity bias of schooling.

In particular, the measurement of years of schooling in our data is exposed to error because it is possible to observe only the last completed degree. However, individuals with the same completed degree could have spent a significantly

[5] Card and Krueger (1992) showed how students who grew up in states with better quality schools acquire more education. Moreover, the place of residence is linked to the possibility to find a job and be well-paid.

different number of years in education. Moreover, the endogeneity bias arise either from unobserved differences in the individual ability or from a general unobserved heterogeneity. Indeed, if individual with higher education have greater ability than others, the estimated return to education is biased upwards since part of the productivity differential is due to their ability or to other skills acquired outside the school (ability bias). Thus, the ability bias interacts with heterogeneous subjective discount rates that result in under-estimating the true effect of schooling on wages when workers with lower education are the more able ones (heterogeneity bias). The total effect of the bias in the OLS estimates is ambiguous.

One way to deal with measurement errors and the endogeneity of schooling is to estimate the Equation (1.11) by using instrumental variables (IV). The identification of a valid instrument is not an easy task and it has been reviewed among others by Card (1999) and Ashenfelter, Harmon and Oosterbeek (1999). The requirements for an instruments to be valid are that it should be correlated with educational choice but not correlated (with the log of) wages conditional on schooling (Wooldridge, 2012).

There is a long tradition in using family background variables, typically the level of parent’s schooling, as a valid instruments (Cannari and D’Alessio, 1995; Colussi, 1997; Card, 1999). The idea is based on the observation of persistence across generation about the level of schooling and it is theoretically justified by involuntary transmission of human capital. Some previous articles on returns to education in Italy derived instrumental variables in the SHIW data, exploiting information provided by the school reforms of the 1960s (Brunello and Miniaci, 1999). However, this type of instrumental variables becomes much less convincing when the focus of the analysis is the time dynamics of return to education. Since the effects of school reforms change according to the population sub-group involved in the reforms, the group of people affected by the instruments changes over time, affecting in turn dynamic comparison of the estimates.

Our instruments will be a set of variables that measure family background, including the highest completed educational level by the father and the mother of the respondents. More educated parents are likely to value education more and to fill better jobs. Furthermore, early educational investment decisions are usually taken not by the individual him/herself, but rather by other agents such as the parents. The assumption is that not only the level and also the kind of education owned by the parents affects the children's one, both through direct decisions, when children are young, and indirect decisions, by encouraging a certain career over another. Checchi, Ichino, Rustichini (1999) show that students choose the level and kind of education not only in relation to their previous curricula but also according to the level and type of education of their parents.

In our estimation strategy, the instruments validity are tested by computing Sargan test, which is an over-identification test with an asymptotic χ^2 distribution and degrees of freedom equal to the number of over-identifying restrictions. The test verifies whether the instruments play a direct role, through predicting educational attainment (Wooldridge, 2012). An important requirement is also that selected instrument should be correlated with the endogenous variable and to test for this, as suggested by Bound et al. (1995)[6], in the first-stage regression of the endogenous variable we compute the F-statistic on the excluded instruments. The F-test on excluded variables shows that our set of instruments is valid, meaning that instruments play a significant role in the reduced form for education and it explains a substantial share of variation in education. Hence, the condition for a valid instrument is satisfied.

[6] If the instruments are weakly correlated with the endogenous variable, this is likely to produce estimates with large standard errors. In particular, if the correlation between the instrument and the endogenous explanatory variable is weak, then even a small correlation between the instrument and the error can produce a larger inconsistency in the IV estimate of the coefficients than in the OLS estimates.

Table 1.2 - IV estimates[7]. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
SCHOOL	0.0643***	0.0619***	0.0687***	0.0712***	0.0668***	0.0786***	0.0587***	0.0685***	0.0542***
	(0.00368)	(0.00764)	(0.00475)	(0.00760)	(0.00678)	(0.00621)	(0.00613)	(0.00784)	(0.00813)
EXPERIENCE	0.0189***	0.0188***	0.0209***	0.0246***	0.0144***	0.0250***	0.0226***	0.0151***	0.0169***
	(0.00331)	(0.00671)	(0.00354)	(0.00530)	(0.00450)	(0.00375)	(0.00447)	(0.00439)	(0.00435)
EXPERIENCE^2	-0.000142*	-0.000142	-0.000199**	-0.000278**	-0.000149	-0.000326***	-0.000270**	-4.20e-05	-7.98e-05
	(7.92e-05)	(0.000155)	(8.13e-05)	(0.000128)	(0.000115)	(9.04e-05)	(0.000112)	(0.000100)	(0.000100)
DUMMY_MALE	0.132***	0.114***	0.0967***	0.0984***	0.0812***	0.109***	0.156***	0.154***	0.101***
	(0.0250)	(0.0354)	(0.0186)	(0.0253)	(0.0261)	(0.0210)	(0.0215)	(0.0266)	(0.0232)
DUMMY_MARRIED	0.00438	0.0501	0.0562**	0.00849	0.0369	-0.00942	-0.0498*	0.0292	0.00841
	(0.0249)	(0.0448)	(0.0251)	(0.0382)	(0.0317)	(0.0235)	(0.0279)	(0.0283)	(0.0326)
NCOMP	0.0177**	0.0150	-0.00134	0.00120	-0.00221	0.0315***	0.0277***	-0.00232	0.0210*
	(0.00728)	(0.0146)	(0.00777)	(0.0101)	(0.00898)	(0.00896)	(0.00893)	(0.0110)	(0.0117)
DUMMY_HOUSEHOLD	-0.00637	-0.00124	0.00590	0.0225	0.0188	0.0306	0.00893		
	(0.0254)	(0.0369)	(0.0184)	(0.0247)	(0.0245)	(0.0200)	(0.0237)		
DUMMY_TOWN	0.00582	0.0310	0.0126	-0.0814**	-0.0184	0.0423*	0.0164	-0.0339	-0.00380
	(0.0210)	(0.0405)	(0.0215)	(0.0360)	(0.0447)	(0.0242)	(0.0305)	(0.0395)	(0.0407)
DUMMY_NORTH	0.0378**	0.0671**	0.0459***	0.0455*	0.0667**	-0.00831	-0.00197	0.0514*	0.0404
	(0.0167)	(0.0286)	(0.0171)	(0.0238)	(0.0297)	(0.0193)	(0.0230)	(0.0288)	(0.0273)
DUMMY_SOUTH	-0.0239	0.0635**	-0.00599	0.00619	0.0224	-0.0493**	-0.0344	0.0201	-0.00710
	(0.0185)	(0.0319)	(0.0224)	(0.0280)	(0.0353)	(0.0230)	(0.0254)	(0.0316)	(0.0340)
DUMMY_AGRICULTURAL	-0.0394	0.0209	-0.116*	-0.0404	-0.0661	-0.127*	-0.0606	0.0278	-0.102*
	(0.0703)	(0.104)	(0.0629)	(0.0578)	(0.0435)	(0.0742)	(0.0481)	(0.0692)	(0.0574)
DUMMY_PUBLIC	0.109***	0.0435	0.0199	0.00801	0.0525*	0.0100	0.0947***	0.0677*	0.0689*
	(0.0218)	(0.0343)	(0.0216)	(0.0314)	(0.0311)	(0.0290)	(0.0288)	(0.0356)	(0.0400)
DUMMY_OTHER_SECTOR	0.0156	-0.00728	-0.00811	-0.0140	-0.0144	-0.0299	-0.00180	0.00405	-0.0520**
	(0.0179)	(0.0397)	(0.0196)	(0.0232)	(0.0263)	(0.0204)	(0.0235)	(0.0265)	(0.0252)
DUMMY_SECT_PARENTS	-0.00735	0.0423*	-0.0131	-0.00227	-0.0113	-0.0150	0.0307*	0.00793	-0.0157
	(0.0182)	(0.0250)	(0.0151)	(0.0186)	(0.0188)	(0.0172)	(0.0186)	(0.0224)	(0.0236)
DUMMY_PART_TIME	0.0387	0.0781	0.0826**	-0.0605	-0.0123	0.0191	0.0131	0.0444	-0.0191
	(0.0360)	(0.0651)	(0.0322)	(0.0444)	(0.0385)	(0.0415)	(0.0366)	(0.0333)	(0.0348)
Constant	1.130***	1.085***	1.112***	1.088***	1.240***	0.980***	1.172***	1.082***	1.209***
	(0.0596)	(0.132)	(0.0692)	(0.0939)	(0.0854)	(0.0879)	(0.0877)	(0.112)	(0.110)
Observations	4,352	1,468	3,783	3,321	3,405	3,437	2,836	2,145	2,112
R-squared	0.403	0.308	0.294	0.261	0.206	0.267	0.331	0.250	0.302

Robust standard errors in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Table 1.2 presents the IV estimates for the period 1995-2012[8]. The Sargan test never rejects the null hypothesis of no miss specification (see the first stage estimation and all the tests in the appendix A), so we cannot reject the validity of over-identifying restrictions. In addition, the Bound test always rejects the null hypothesis of no correlation between education and additional instruments.

We confirm for this sample the finding that the estimated returns to education are significantly larger with IV than with OLS, as stressed by large part of the

[7] In the SHIW waves, information about family background is available only for the households and for his/her spouse or cohabitant. For year 2008 for the households and for his/her spouse or cohabitant if the households is borne in an odd year, while for year 2010 and year 2012 only for the households.

[8] We also estimate return to education by applying OLS (the results are showed in the Appendix A). Consistent with the existing literature, we find large positive returns to education after instrumenting for education; the two-stage least squares estimates are much larger than their OLS counterparts. OLS approach, failing to address endogeneity and measurement errors problems consistently underestimates the returns to education. IV estimates are generally 20–40% above their OLS counterpart (Trostel et al., 2002).

international literature. The downward OLS bias implied by IV estimates could arise from the attenuation effect of a measurement error in the schooling variables, but also a distortion from omission of the variable “ability” could lead to a similar result. This means that the more “able” (in terms of capacity to earn higher wages) individuals have lower preference for schooling, and those preferences could be justified by the higher opportunity costs faced by the “able” individuals.

1.4.1.1 The Return on Schooling

The main features of empirical research on returns to education in Italy are shown in Table 1.3. The estimated rate of return to an additional year of schooling vary across studies, also for the method used in the estimate. Antonelli (1985), who consider regional data, estimates that an additional year of schooling increases annual net wages by 4.6 percent. Cannari et al. (1989) use a larger sample from the 1986 wave of the Bank of Italy, finding a similar result of a return around 4 per cent. While Lucifora and Reilly (1990) estimate the mincerian wages function using the ENI special survey on earning and they find that the marginal return to schooling is slightly higher for women than for men but again around 4 percent.

Table 1.3 – A summary of the estimated rates of return to schooling of an additional year of schooling in Italy

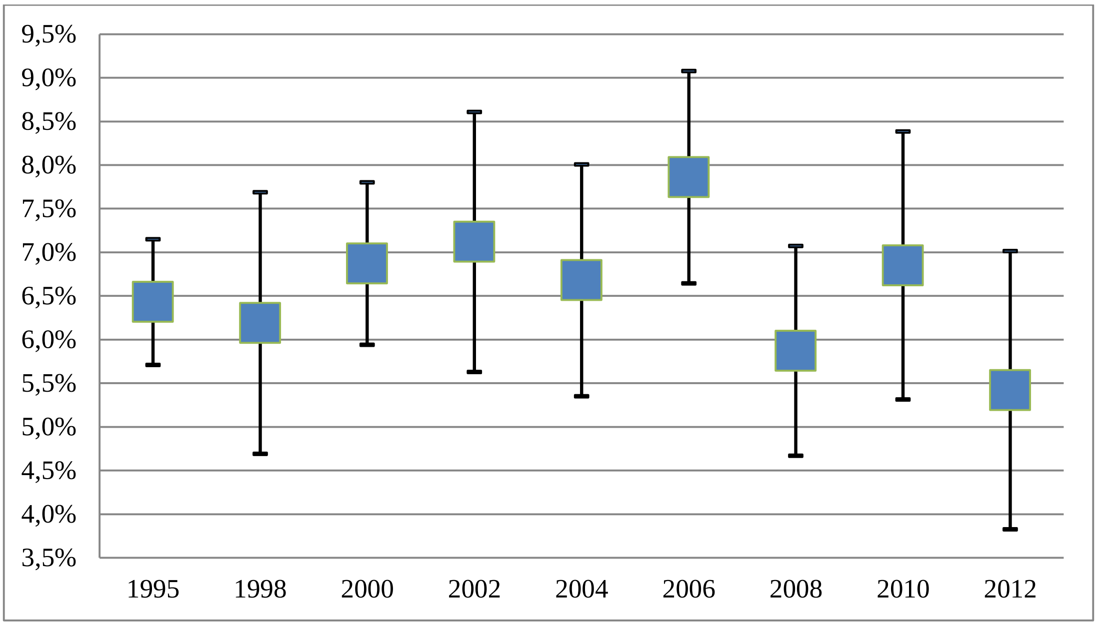
Author	Method of Estimation	Years	Estimated Rates of Return to Schooling (%)
Antonelli (1985)	OLS	1977	4.6
Cannari, Pellegrini, and Sestito (1989)	OLS	1986	4.0
Lucifora and Reilly (1990)	OLS	1985	3.6 (men) 4.0 (women)
Cannari and D'Alessio (1995)	IV	1993	7.0
Colussi (1997)	IV	1993	7.6
Flabbi (1997)	IV	1991	6.2 (men) 5.6 (women)
Brunello and Miniaci (1999)	IV	1993 and 1995	5.7
Brunello, Comi, and Lucifera (2000)	OLS	1995	6.2 (men) 7.7 (women)
Ciccone (2004)	OLS	1987-2000	6.1
Ciccone, Cingano, and Cipollone (2006)	OLS	1987-2000	6.9
Mendolicchio (2006)	PV	2002	5.3 (men) 6.5 (women)
Cingano and Cipollone (2009)	OLS	1987-2000	6.0

For the 1993 wave of Bank of Italy Cannari and D'Alessio (1995), using family background variables as instruments of educational outcomes, find that the marginal return to education is around 7 percent, much higher than previous results. Also Colussi (1997) obtain a similar result, using the same wave and a similar set of instruments. For 1991 wave, Flabbi (1997) calculates the returns to education separately for men and women with an instrumental variable approach based upon the identification of exogenous changes in the schooling system; he finds that the marginal effect of education is 6.2 percent for men and 5.6 percent for women, confirming the gender gap in wages. For the 1993 and 1995 waves, Brunello and Miniaci (1999) estimate a return to education equal to 5.7 percent (taking into account the endogeneity of schooling). The estimated coefficient on the mincerian rate of return to schooling is around 6 percent in Ciccone (2004) and Cingano and Cipollone (2009).

Brunello, Comi and Lucifora (2000) find evidence of a greater return to schooling for women, that is also confirmed in the work of Mendolicchio (2006), in which proxy variables approach is applied to deal with the endogeneity of the schooling variable.

In our results from the estimations of the Mincerian wage equation, the evidence is that returns have changed over the period considered. The estimations of the returns to schooling are between 5.4 percent and 7.9 percent, recording the highest level in 2006 and the lowest in 2012, and on average the rate of return to schooling is equal to 6.6 percent. Looking at the previous estimates made for Italy, as shown in Figure 1.1, we can notice that our estimate are in line with the literature. Moreover, from 1995 to 2012, it is not present a clear patterns of the return to schooling, either increasing or decreasing.

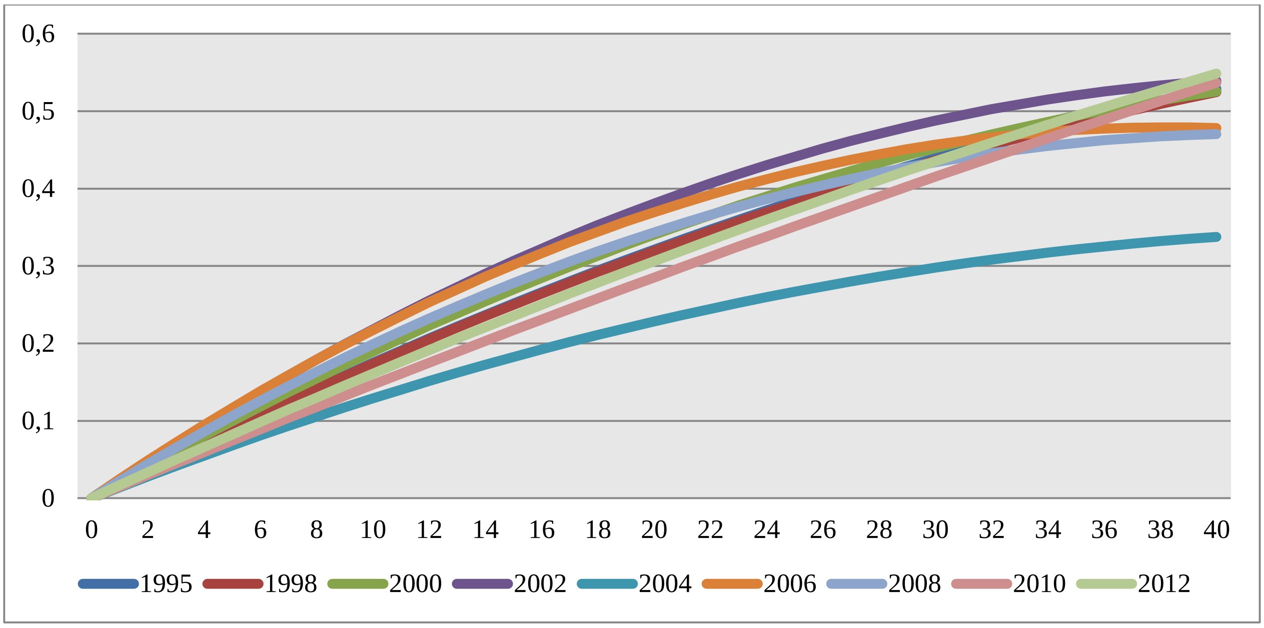
Figure 1.1 - Estimates of the Return to Education, 1995-2012 (with confidence intervals at 95%)



1.4.1.2 The Return on Experience

The dynamics of experience is drawn in Figure 1.2. We observe different pattern for each year of the sample: from 1995 to 2008 the experience profile is a concave function, more or less steeper, while in 2010 it is approximately a linear function. Therefore, we can affirm that the experience profile is not linear function (except for 2010 and 2012) and that the estimates are quite stable over the time period considered.

Figure 1.2 - Estimates of the Experience Profile, 1995-2012



1.4.1.3 The Impact of Other Variables

If we consider the DUMMY_MALE variable, we observe a strong evidence of a gender pay gap, in favor of men for all the period considered, with an increasing trend, from 13.2 percent in 1995 to 15.4 percent in 2010 and to 10.1 in 2012.

Considering the geographical residence of the workers and the sector of employment, differences in estimates mainly reflect territorial and sectorial performance of Italy.

The DUMMY_NORTH is positive while the DUMMY_SOUTH in negative. This means that it is more convenient to work in the north regions in comparison to the central regions, instead if an individual works in the south region he will earns less than in the center regions. Therefore, working in the same sector of the father or the mother (DUMMY_SECT_PARENTS) seems to not bring particular benefits, except for year 1998 and 2008 where this dummy is significant and positive.

Finally, considering different sector of employment, working in the agricultural sector is less convenient than working in the industrial sector. On the contrary, working in the public sector is more convenient than working in the industrial sector.

1.4.2 Mincer Wage Model with Different Types of School

The current Italian education system is composed by primary, secondary, upper secondary and tertiary education. Primary school is compulsory for children aged between 6 and 11 years. Lower secondary education is also compulsory, free of charge and lasts three years. Post compulsory education is divided into the following categories: classical, scientific and pre-school teacher training, artistic education, technical school and vocational education. Upper secondary education lasts from three to five years, depending on the type of school. Since 1969, the selection of the type school does not preclude access to tertiary education. Graduation from upper secondary schools requires a leaving school certificate examination and access to tertiary education is only conditional on passing this exam.

In comparison with other OECD countries in 2012, average education attainments of the upper secondary education in Italy is substantially low as shown in Table 1.4. On average across OECD countries, the percentage of 25-34 year-olds with at least upper secondary education is 18 percent higher than that among 55-64 year-olds (about 82 percent against 64 percent). This difference for cohort can be explained by the observed general decline in demand for manual labor and for basic cognitive skills (easily replicated by computers), in favor of a sharp increase in the demand for complex communication and advanced analytical skills, which require a more educated labor force.

Table 1.4 - Percentage of adults who have attained at least upper secondary education, by age group (2012)

	25-34 years old	55-64 years old
OECD average	82	64
Italy	72	42

Source: OECD (2014)

In Italy, just 72 percent of the age-group 25-34 (versus an OECD average of 82 percent) has attained at least upper secondary education; however, such a percentage is much higher than the 42 percent of the 55-64 age-group.

For what concerns tertiary education in OECD countries we observe the same upward trend of education attainment for younger cohorts of population as reported in Table 1.5 (from 24 percent to 39 percent): younger adults have higher tertiary education than older adults by an average of 15 percentage points.

Table 1.5 - Percentage of adults who have attained tertiary education, by age group (2012)

	25-34 years old	55-64 years old
OECD average	39	24
Italy	22	11

Source: OECD (2014)

In Italy in 2012 the percentage of population in the 25-34 years-olds cohort with a university degree is equal to 22 percent, much lower than the OECD average of 39 percent. Although Italy shows a very significant increase over time of the percentage of the population attaining tertiary education (22 percent of the 25-34 age group must be compared with 11 percent of the 55-64 age group), we notice that such difference is well below that observed for OECD countries (from 24 percent to 39 percent).

Considering gender in OECD and Italy, evident disparities in educational attainments between women and men are present in the older generations, but with a significant inversion in the more recent cohorts (see Tables 1.6 and 1.7). In particular, in OECD countries while for older generation (e.g. 55-64 age group) the percentage of people attaining upper secondary and tertiary education is significantly larger for men, for the 25-34 age group the educational level is higher for women.

Table 1.6 – Percentage of adults who have attained at least upper secondary education, by age group and gender (2012)

Women, by age group

	25-64	25-34	35-44	45-54	55-64
OECD average	75	84	79	72	61
Italy	59	76	65	55	40

Men, by age group

	25-64	25-34	35-44	45-54	55-64
OECD average	76	81	78	74	68
Italy	56	68	59	51	45

Source: OECD (2014)

The gender gap in education in favor of women is recorded also in Italy: 8 percent higher for the same group for upper secondary education, and 10 percent higher for women aged 25-34 for tertiary education.

Table 1.7 - Percentage of adults who have attained tertiary education, by age group and gender (2012)

Women, by age group

	25-64	25-34	35-44	45-54	55-64
OECD average	34	44	38	30	23
Italy	17	27	19	13	11

Men, by age group

	25-64	25-34	35-44	45-54	55-64
OECD average	30	34	33	28	25
Italy	14	17	15	11	11

Source: OECD (2014)

In all OECD countries, adults with tertiary education earn more than adults with upper secondary or post-secondary non-tertiary education, who, in turn, earn more than adults without upper secondary education. Across OECD countries, compared with adults with upper secondary education who have income from employment, those without this qualification earn about 20 percent less, those with post-secondary non-tertiary education about 10 percent more, those with tertiary-vocationally oriented education about 30 percent more, and those with tertiary-academically oriented education or advanced research earn about 70 percent more.

Higher educational attainment is associated with higher wages during a person's working life. On average across OECD countries, wages increase with the level of educational attainment, but this increase is particularly large for older workers. People with higher levels of education are more likely to be employed, and remain employed, and have more opportunities to gain experience on the job. On average, the wages of tertiary-educated 55-64 year-olds is larger than that for 25-64 year-olds: by 36 percent for OECD countries, by 43 percent for Italy.

Regardless of the level of education, the gender gap in wages persists. Across OECD countries, a tertiary-educated woman earns about 73 percent of what a tertiary educated man earns (in Italy women who have obtained a tertiary degree earn 69 percent or less of tertiary-educated men).

Finally, in all OECD countries, individuals with a tertiary-level degree have a greater chance of being employed than those without such a degree. In general, higher education improves job prospects and the likelihood of remaining employed in times. In 2012, in Italy 79 percent of the population with a tertiary education is employed against 71 percent with an upper secondary education (84 percent against 74 percent in OECD countries).

1.4.2.1 The Return of Different Level of Schooling

The empirical specification in Equation (1.11) is based on the assumption that the return to education is constant and independent of the level of attained education. In this section, we allow the marginal return to schooling to vary with the level of completed education by replacing years of schooling with three educational dummies, one for each level of completed schooling above compulsory school, that is vocational school, secondary and tertiary education. This is the multiple factor model, an alternative way to estimate returns to schooling, where different educational levels have separate effects on wages.

As suggested by the "credentialism" hypothesis, in the presence of heterogeneity what really matters is the type of school rather than the overall number of years spent in formal education. We investigate these issues by considering the highest degree attained by individual using educational dummies rather than years of schooling in our wages regressions. In particular, we look at education achievements by broad levels: compulsory school (no schooling, primary school and junior high school), vocational, upper secondary and tertiary education.

Also in the case of the estimate the returns of education from different type of school, we deal with the problem of endogeneity by using instrumental variables. We apply the two step methodology proposed by Vella and Gregory (1996). The empirical strategy consists of estimating the two following equations:

$$\ln w(s_i, x_i) = \alpha_{0i} + \sum_{h=1,3} \varphi_h E_{ih} + \beta_0 x_i + \beta_1 x_i^2 + \varepsilon_i \tag{1.12}$$

$$s_i^* = z_i \gamma + v_i \tag{1.13}$$

where w_i is the real hourly wage, E_{ih} are educational dummies that correspond to the highest degree achieved by the individual, x_i and z_i are observed attributes, ε_i and v_i are normally distributed error terms with zero means and finite variances, s_i^* is the latent level of education. We define s_i as the observed level of education, that takes the following discrete values:

$$s_i = \begin{cases} 1 & \text{if } s_i^* < \mu_0 \\ 2 & \text{if } \mu_0 \leq s_i^* \leq \mu_1 \\ 3 & \text{if } s_i^* \geq \mu_1 \end{cases} \tag{1.14}$$

and associate s to the educational dummies by setting $E_{ih} = 1$ if $s_i = h$ and $E_{ih} = 0$ otherwise.

We use a two steps procedure to estimate the coefficients. In the first step we estimate an ordered Probit model for educational attainment as a function of the instrument used in the previous IV estimation. In the second step, we include the score[9] associated to the ordered Probit in the wages equation and we then apply ordinary least squares. Our specification of the ordered Probit includes the same covariates of the instrumental equation used before.

The interpretation of the estimated coefficients is in terms of additional return that the educational level provides to the individual with respect to the reference group that is compulsory school. Our results are reported in Table 1.8. For instance, in 2012, an employee with a high school degree earns, on average, 25.6 percent more than an employee with the same covariate belonging to the

[9] See Idson and Feaster (1990) for details on the computation of the score.

reference group. This differential increase to 56.5 percent for graduated individuals.

The estimated coefficients of the score have always a negative sign, implying that the covariance between unobservable variables that affect wages and educational choice is negative. This means that an individual attains a lower educational level than predicted, because individuals with higher ability have a higher marginal cost of schooling in terms of foregone wages, due to more attractive wage offer. Hence, these individuals tend to acquire less education that predicted education and earn higher wages.

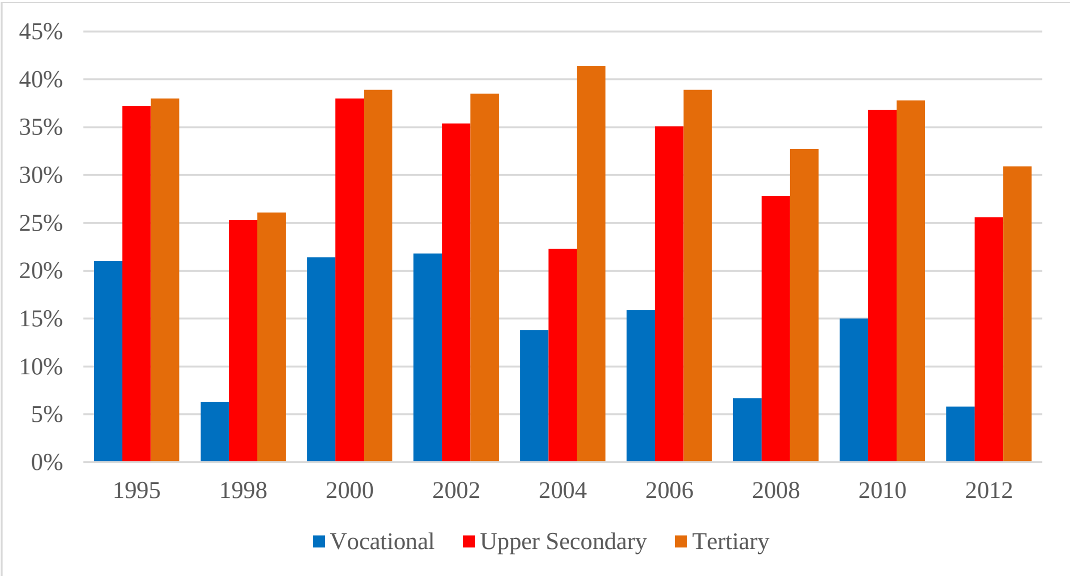
Table 1.8 – Second stage OLS estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
VOCATIONAL	0.210***	0.0632	0.214***	0.218***	0.138***	0.159***	0.0669*	0.150***	0.0581
	(0.0349)	(0.0558)	(0.0332)	(0.0368)	(0.0404)	(0.0313)	(0.0391)	(0.0570)	(0.0417)
UPPER_SECONDARY	0.372***	0.253***	0.380***	0.354***	0.223***	0.351***	0.278***	0.368***	0.256***
	(0.0304)	(0.0504)	(0.0362)	(0.0440)	(0.0556)	(0.0404)	(0.0398)	(0.0578)	(0.0532)
TERTIARY	0.752***	0.514***	0.769***	0.739***	0.637***	0.740***	0.605***	0.746***	0.565***
	(0.0508)	(0.0838)	(0.0567)	(0.0828)	(0.0880)	(0.0682)	(0.0695)	(0.105)	(0.0924)
EXPERIENCE	0.0213***	0.0191***	0.0239***	0.0281***	0.0155***	0.0264***	0.0242***	0.0160***	0.0183***
	(0.00328)	(0.00669)	(0.00348)	(0.00509)	(0.00433)	(0.00355)	(0.00436)	(0.00437)	(0.00442)
EXPERIENCE^2	-0.000239***	-0.000185	-0.000303***	-0.000394***	-0.000198*	-0.000386***	-0.000311***	-6.27e-05	-0.000107
	(7.85e-05)	(0.000155)	(7.95e-05)	(0.000122)	(0.000111)	(8.39e-05)	(0.000111)	(0.000101)	(0.000103)
DUMMY_MALE	0.135***	0.124***	0.102***	0.0987***	0.0684***	0.109***	0.155***	0.160***	0.105***
	(0.0256)	(0.0344)	(0.0181)	(0.0245)	(0.0265)	(0.0200)	(0.0215)	(0.0268)	(0.0226)
DUMMY_MARRIED	0.00754	0.0535	0.0531**	0.00621	0.0434	-0.00680	-0.0444	0.0209	0.00593
	(0.0254)	(0.0447)	(0.0244)	(0.0371)	(0.0310)	(0.0222)	(0.0276)	(0.0276)	(0.0325)
NCOMP	0.0165**	0.0112	0.00124	0.00223	-0.00471	0.0322***	0.0250***	-0.000861	0.0202*
	(0.00734)	(0.0144)	(0.00742)	(0.0101)	(0.00910)	(0.00840)	(0.00860)	(0.0108)	(0.0113)
DUMMY_HOUSEHOLD	-0.00255	-0.00839	0.00905	0.0269	0.0268	0.0376**	0.0125		
	(0.0257)	(0.0352)	(0.0179)	(0.0240)	(0.0239)	(0.0187)	(0.0237)		
DUMMY_TOWN	0.00673	0.0374	0.00600	-0.0792**	-0.00873	0.0444*	0.0189	-0.0296	-0.0102
	(0.0211)	(0.0411)	(0.0209)	(0.0349)	(0.0422)	(0.0239)	(0.0304)	(0.0373)	(0.0394)
DUMMY_NORTH	0.0366**	0.0775***	0.0463***	0.0494**	0.0599**	-0.00521	0.00325	0.0606**	0.0492*
	(0.0168)	(0.0282)	(0.0162)	(0.0232)	(0.0282)	(0.0182)	(0.0227)	(0.0283)	(0.0275)
DUMMY_SOUTH	-0.0345*	0.0428	-0.0284	-0.0114	-0.00934	-0.0761***	-0.0489**	0.0167	-0.0119
	(0.0187)	(0.0314)	(0.0214)	(0.0273)	(0.0330)	(0.0219)	(0.0249)	(0.0289)	(0.0332)
DUMMY_AGRICULTURAL	-0.118	-0.143	-0.190***	-0.0991*	-0.129***	-0.151**	-0.0997**	-0.00494	-0.115**
	(0.0717)	(0.0921)	(0.0578)	(0.0577)	(0.0406)	(0.0746)	(0.0502)	(0.0694)	(0.0564)
DUMMY_PUBLIC	0.114***	0.0991***	0.0230	0.0216	0.0935***	0.0464*	0.105***	0.0545	0.0664
	(0.0227)	(0.0330)	(0.0230)	(0.0313)	(0.0346)	(0.0282)	(0.0273)	(0.0390)	(0.0422)
DUMMY_OTHER_SECTOR	0.0103	0.0165	-0.000773	-0.00667	0.00656	-0.0134	0.00719	0.00186	-0.0487*
	(0.0179)	(0.0388)	(0.0197)	(0.0232)	(0.0260)	(0.0194)	(0.0235)	(0.0263)	(0.0254)
DUMMY_SECT_PARENTS	0.00569	0.0543**	-0.0108	-0.000419	0.00189	-0.0120	0.0293	0.00830	-0.0150
	(0.0180)	(0.0255)	(0.0146)	(0.0185)	(0.0188)	(0.0160)	(0.0189)	(0.0215)	(0.0240)
DUMMY_PART_TIME	0.0355	0.0622	0.0729**	-0.0739*	-0.0360	0.00195	0.0121	0.0393	-0.0212
	(0.0346)	(0.0637)	(0.0315)	(0.0433)	(0.0391)	(0.0398)	(0.0353)	(0.0368)	(0.0332)
SCORE	-0.0543***	0.00650	-0.0958***	-0.0770***	-0.0362	-0.0857***	-0.0437*	-0.0960***	-0.0556*
	(0.0187)	(0.0317)	(0.0214)	(0.0275)	(0.0319)	(0.0238)	(0.0255)	(0.0369)	(0.0317)
Constant	1.576***	1.596***	1.596***	1.613***	1.803***	1.608***	1.634***	1.601***	1.636***
	(0.0462)	(0.0984)	(0.0456)	(0.0612)	(0.0568)	(0.0596)	(0.0545)	(0.0705)	(0.0655)
Observations	4,352	1,468	3,783	3,321	3,405	3,437	2,836	2,145	2,112
R-squared	0.412	0.317	0.339	0.286	0.245	0.327	0.345	0.295	0.327

Robust standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

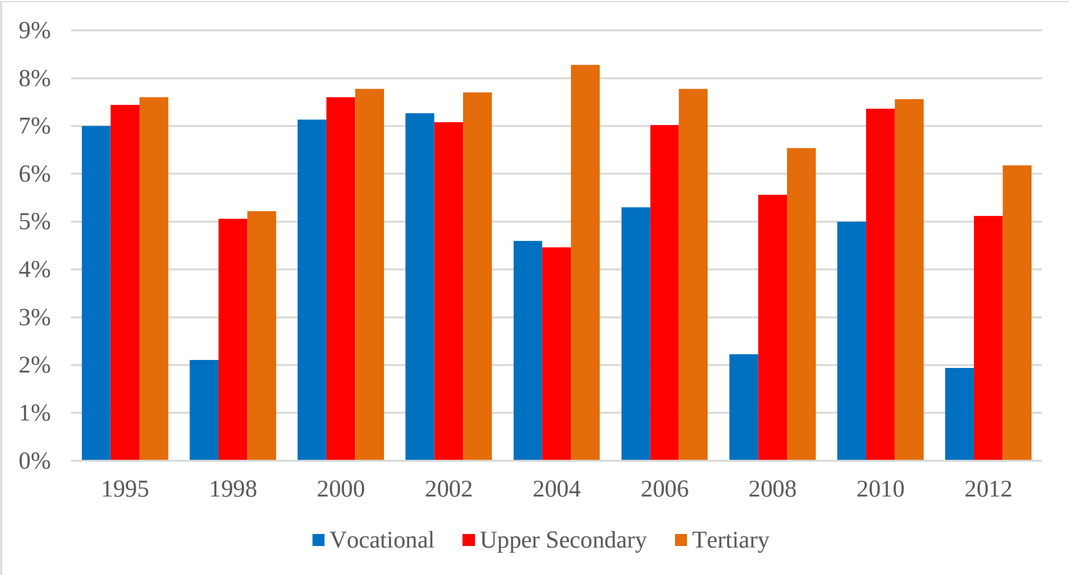
Considering different educational attainment, vocational school seems to have a not clear pattern, from 21 percent in 1995 to 15 percent in 2010. The rate of return of secondary school is not constant over the period considered, but it shows a slightly decreasing trend from 1995 to 2010. The same trend is observed for the rate of return of tertiary education (university).

Figure 1.3 – Rate of Return of Different Types of School 1995-2012



However, even if the college premium does not have a particular trend, attending college let to have between 30 percent and 40 percent of higher wages.

Figure 1.4 – Annual Rate of Return of Different Types of School 1995-2012 (reference category: compulsory school)



Moreover, we assume that these returns can be spread evenly among the years of school required to complete a degree (see Figure 1.4). It turns out that the increase in wages due to an additional year of vocational school, upper secondary school and college is respectively 5 percent, 7.4 percent and 7.6 percent in 2010. Hence, there is evidence that returns to education are not constant but increase with the level of attained education.

Finally, considering experience and the other control variables that are included in the estimation, we do not observe significant changes from the IV estimates.

5 Concluding Remarks

We have studied the wage function in Italy, focusing on the role of return to education. Using cross-sectional data from the 1995 to 2012 waves of the Bank of Italy survey on the income and wealth of Italian household, we have applied instrumental variables estimation to solve the problem of endogeneity. The evidence is that returns to schooling have changed over the period considered, 1995-2012, and are between 5.4 percent and 7.9 percent, recording the highest level for 2006 and the lowest in 2012. Considering different sector of employment, a relative convenience to work in the public sector emerges. In addition, there is an evidence of a gender pay gap, in favor of men for all the period considered.

When the type of school attended is taken into consideration, we also find that the returns to education increase with higher levels of educational attainment. In this case, to solve the problem of endogeneity, an ordered Probit is applied to the choice of educational attainment and then we add the score of the Probit estimation, to the original equation and apply OLS. In particular, for 2010, the estimated coefficient of the educational dummy is respectively 15 percent for vocational school, 36.8 percent for upper secondary, and 74.6 percent for college education. More able subjects, who received better wage offers, have lower education than predicted, because of the relative incentive to anticipate labor market entry (as signaled by the negative coefficient of the score).

In this analysis we take into consideration only employees excluding self-employed because of low reliability of their declared earnings. Restricting the analysis only to employees probably leads to an underestimation of the returns to education in Italy. However, the possible presence of outliers in earnings of certain categories of self-employed (typically professionals and managers) could lead to an upward bias and the solution to this problem and is left to future research.

Chapter 2

A Semiparametric Estimate of the Mincer Wage Function in Italy

Abstract. Past studies on return of education assume that the marginal rate of return is constant over all levels of education. The main objective of this chapter is to relax this assumption and the parametric structure of the related econometric model. In this matter, it is possible to test the non-linearity of the returns of schooling. Moreover, the findings allow exploring the nature of the shape of the returns function. In order to pursuit the aim of this work, a semiparametric additive model is applied using data from the Bank of Italy's Survey of Household Income and Wealth (SHIW). To deal with the endogeneity of omitted variables in the wage equation a control function method is performed. Results show that the wages−schooling relationship is non-linear, allowing for return to education that depend on the level of schooling.

2.1 Introduction

The most of contributes estimating the Mincer wage function inspired by the pioneering work by Mincer (1974) assume a constant rate of return on schooling and a quadratic specification for the impact of experience on wage. In this chapter, we estimate the wage function without imposing any restriction on the shape of marginal impact of education and experience on wage.

The policy implications of our analysis range from the possibility to target more precisely the effort of Government on education, to provide a differentiated support to on-the-job training conditioned to the year of experience.

The rate of return on education has been estimated in literally hundreds of studies (for a review see, e.g., Psacharopoulos 1994, Ashenfelter et al. 1999, Harmon et al. 2000). The vast majority of these works assumes that the marginal rate of return is constant over all levels of education, even though some studies have found significant nonlinearities. In particular, Mincer (1974), Psacharopoulos (1985, 1994) and Hamon and Walker (1999) document significant diminishing return on education, while Card and Kruger (1992) provide evidence in favour of increasing returns at low level of education. Finally, Heckman and Polacheck (1974) and Card (1995, 1999) argue that the return on education appears approximately constant. Heckman et al. (2008) provide evidence against the assumptions that schooling has a constant marginal impact on wage, and that the impact on wage of schooling and experience can be separately estimated[10]. The literature is therefore not conclusive on this point.

The literature has instead paid less attention to the shape of the impact of experience on wage. The use a quadratic specification makes the estimate subject to an important misspecification bias. Murphy and Welch (1990) using a quadratic specification represent an attempt to limit this bias; in the same respect Zheng

[10] They formally reject the hypothesis of linearity in returns to education in the Mincer regression using US national level census data for all census years from 1940 to 1990.

(2000) proposes the use of higher ordered polynomials. Also for this feature of the model a conclusion is not still reached.

Semiparametric techniques appear particularly well suited to deal with the problem of specification of Mincer wage function. So far their use has been limited by endogeneity, generally due to the presence of omitted variables (typically the unobserved individual ability). But the use of control function approach appears to effectively circumvent this drawback[11]. In particular, we estimate a semiparametric Mincer wage function, in which both return on education and experience enter as nonparametric terms. Our methodology for the estimate of semiparametric model is based on Wood (2011), integrated with the control function method as discussed by Blundell and Powell (2003).

To our knowledge, no analysis on a semiparametric specification of Mincer wage function is available for Italy. More importantly, we provide evidence that the return on education for all the available waves in the period 1995-2012 are increasing in the level of education, starting from 4 percent for five years of schooling to 8 percent for fifteen year of schooling on average. However, decreasing returns are observed for workers with 18 years of schooling (tertiary education), from 2008 to 2012. The impact of experience, on the contrary, appears well approximated by the quadratic form generally used in the literature, even though in some waves the estimated relationship is approximately linear. Endogeneity of education appears pervasive in all the estimates, justifying the use of the control function approach.

The chapter is organized as follows. Section 2 presents a more general version of Mincer wage function, which includes the possibility of varying return on education and experience. Section 3 describes the sample used in the analysis, while Section 4 contains the estimates. Finally, Section 5 concludes. Technical stuff are gathered in appendix B.

[11] Garen (1984) represents an early application of the control function approach to estimate Mincer wages function, but he does not consider a semiparametric specification.

2.2 Semiparametric Specification of the Mincer Wage Function

In Section 2 of Chapter 1, we derived the Mincer wage function as:

$$\ln w(s_i, x_i) = \alpha_{0i} + \rho_s s_i + \beta_0 x_i + \beta_1 x_i^2 + \varepsilon_i, \tag{2.1}$$

where ε_i is a mean zero residual with $E(\varepsilon_i | s_i, x_i) = 0$, and $\alpha_0 = \ln E_0 - \kappa(1 + \rho_{ps})$, $\beta_0 = \kappa \left[\rho_{ps}\left(1 + \frac{1}{2T}\right) + \frac{1}{T}\right] - \delta$, $\beta_1 = -\frac{\rho_{ps}\kappa}{2T}$.

Equation (2.1) is based on two key assumptions: the return on education is independent of the level of education and the investment in human capital of employed workers is hyperbolic declining with experience as suggested by Ben-Porath (1967) model. Relaxing these two assumptions leads to the more general Mincer wage function as follows:

$$\ln w(s_i, x_i) = \alpha_{0i} + f_1(s_i) + f_2(x_i) + \varepsilon_i, \tag{2.2}$$

Equation (2.2) represent a semiparametric specification of Mincer wage function, which allows both to alleviate the potential misspecification bias of original formulation and at the same time limits the computational burden of a full nonparametric specification.

The first derivative of f_1 and $f_2$respectively represent the marginal return on education and experience.

2.2.1 Endogeneity of Schooling

In the literature, there is a wide consensus that the presence of omitted variables in the estimate of Mincer wage function poses a key problem of endogeneity. In particular, the unobserved individual ability could significantly affect the choice of education; in this regard parents' schooling is generally used as instrumental variable as we discuss in Chapter 1.

Blundell and Powell (2003) discuss three approaches to deal with endogeneity in semiparametric models: standard instrumental variable approach, fitted value approach and control function approach, arguing that the last provides the best choice. In particular, the control function approach treats the endogeneity problem as a problem of omitted variables, where omitted variable is estimated in the first stage by regressing the endogenous variable on the instrumental variable along with other independent variables. The estimated residual is then included as an independent variable in the second stage to control for endogeneity.

The implementation of the control function approach in the semiparametric model proceeds with the calculation of residuals from the first-stage semiparametric regression:

$$s_i = \pi_1(z_i) + \pi_2(x_i) + v_i, \tag{2.3}$$

where z_i is an instrumental variable, x_i is the experience variable, v_i is the unobserved error term, $\pi_1(.)$and $\pi_2(.)$ are the unspecified functions on the instrumental variable and on experience. Then, the estimated residuals from Equation (2.3) are inserted in the second stage semiparametric regression as an independent variable, i.e.:

$$\ln w(s_i, x_i) = \alpha_{0i} + f_1(s_i) + f_2(x_i) + \gamma_1 \widehat{v_i} + \varepsilon_i, \tag{2.4}$$

where $\gamma_1$is an unspecified parameter on estimated residuals, $\widehat{v_i}$, from the first stage.

Test on the presence of endogeneity are made on the estimate of parameter γ_1; in particular, the rejection of null hypothesis of γ_1=0 results in the not possibility to reject the presence of endogeneity in the estimate.

Finally, Blundell and Powell (2003) discusses how the following restriction:

$$E(\varepsilon|S, X, Z) = E(\varepsilon|S, X, v) = E(\varepsilon|v) \tag{2.5}$$

must hold for a correct application of the control function approach. We will discuss all these diagnostics of estimation in the section of results.

2.3 Data and Sources

The analysis is based on the same sample used in Chapter 1. In particular, data come from the Bank of Italy's Survey of Household Income and Wealth (SHIW), which reports several socio-economic characteristics of Italian households. The SHIW is a biannual survey on Italian families with a sample of approximately 8,000 household per year. From 1995 to 2012 observations from nine subsequent surveys are available. SHIW contains information both on households (family composition) and on individuals; moreover, it provides detailed information on several characteristics of workers within each household, such as their net yearly wages, average weekly hours of work and number of months of employment per year, educational attainment (the highest completed school degree), job experience, gender, marital status, sector of employment, household composition, parents background, regions of residence, and town size.

We consider a sub-sample of men and women between 15-64 years old, full time and part time employees, working either in the public or in the private sector and such that information about wages are available. In the analysis, we exclude self-employed because of the low reliability of their declared earnings. As discussed by Brandolini and Cannari (1994), SHIW seems to underestimate the self-employed earnings of about 50 percentage points.

2.3.1 Variables Used in the Analysis

Mincer wage function in Eq. (2.2) refers to the (log of) hourly price of labor as correct measure of worker's wages (LOGY_H), and, this is indeed the measure used by most empirical studies[12] (Brunello and Miniaci, 1999; Blundell, Dearden and Sianesi, 2005; Ciccone, Cingano and Cipollone, 2006). SHIW contains yearly net wages of taxes and social security contributions. Additional information on the average number of hours worked per week and on the number of months worked per year, can be used to estimate the hourly net wage, which is calculated by yearly net wages divided by months worked multiplied by hours worked each month.

Schooling attainment (SCHOOL) is generally measured by the number of years spent at school. SHIW does not contain information about this number of years, but only on the highest degree attained by individuals. Following a common approach in literature (Vieira, 1999; Brunello and Miniaci, 1999) we calculate the educational attainment of the individual by imputing the number of years required to complete her/his reported maximum level of educational attainment[13]. More precisely, we consider that the (statutory) numbers of years required to obtain a primary and a junior school certificate is 5 and 8 years respectively; instead, for the upper secondary school the number of years ranges from 11 (vocational or technical school) to 13 (classical or scientific studies); finally, for tertiary education, we consider 16, 18 and 21 years for the university diploma, the college degree, and the postgraduate degree (e.g. Ph.D.) respectively. It is important to remark that in Italy the statutory number of years can be significantly different from the actual number of years spent to obtain a degree, especially at college because of the high percentage of irregular student.

Many empirical studies use age as a proxy for the (working) experience of individuals. But this choice can be severely biased, especially for young cohorts. Other authors use potential experience, defined as the difference between the current age and the age at the labor market entry, but they ignore the possibility

[12] Hourly wages can be affected by measurement errors because we calculate them as total wages divided by hours of work.

[13] Standard, not actual, years of formal schooling are recorded. Since students who fail to reach a standard have to repeat the year, the actual number of years is likely to be underestimated.

of unemployment or underemployment, again a crucial feature for young cohorts. In this work we use, as proxy for experience (EXPERIENCE), the number of years for which a worker has been paid social security contribution; they should reflect the effective years of training on the job and learning-by-doing activities.

We introduce several control variables in the analysis to account for individual characteristics and for differences in the labor market. A gender dummy (DUMMY_MALE) controls for different wage levels between men and women. Marital status also enter into the analysis as a dummy variable (DUMMY_MARRIED) taking the value 1 if the person is formally married, 0 otherwise. Part-time work is captured through a separate dummy variable (DUMMY_PART_TIME), since the assumption that each working hour makes the same contribution to weekly wages (constancy of the hourly wage) may not hold across workers with different time status (part time versus full time).

In addition, controls are introduced for family composition, as a proxy for the influence of housework, particularly important in the female labor supply (Heckman and Killingsworth, 1986). We control for the number of components of the family (NCOMP) and for the fact that the individual is the head of his/her household (DUMMY_HOUSEHOLD).

Controls for sector (DUMMY_AGRICULTURAL for the agricultural sector, DUMMY_INDUSTRIAL for the industrial sector, DUMMY_PUBLIC for the public sector and DUMMY_OTHER_SECTOR for other sector different from the previous ones) should capture potential factors from the demand side of labor market (e.g. imperfectly competitive labor markets). In the same light, we add some controls for the geographical area of residence: one dummy for the town of residence that has more than 500.000 inhabitants (DUMMY_TOWN), and three different dummies for the Italian macro-regions: North, Center and South (DUMMY_NORTH, DUMMY_CENTER and DUMMY_SOUTH).

Our instruments for schooling will be a set of variables that measure family background, including the highest completed educational level by the father and the mother of the interviewed individual (SCHOOL_F and SCHOOL_M). There is a long tradition in using family background variables, typically the level's of parent's schooling, as a valid instruments (Cannari and D'Alessio, 1995; Card, 1999; Trostel et al., 2002). The idea is based on the observation of persistence across generation about the level of schooling and it is theoretically justified by involuntary transmission of human capital.

Table 2.1 reports some descriptive statistics of the main variables used in the empirical analysis for all the waves (wages are expressed in euro 2012).

Table 2.1 - Means and standard deviations of the variables used in the empirical analysis for the entire sample (1995, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012)

Variable	Mean	S. d.	Description
LOGY_H	2,265	0,438	Logarithm of the hourly wages less tax
SCHOOL	11,373	3,800	Schooling attainment, that is the number of years spent at school
EXPERIENCE	17,683	10,673	Number of years for which it has been paid social security contributions, as a proxy for years of training on the job
DUMMY_MALE	0,578	0,494	Gender dummy
DUMMY_MARRIED	0,647	0,478	Dummy variable for marital status
NCOMP	3,329	1,185	Number of components of the family
DUMMY_HOUSEHOLD	0,475	0,499	Household dummy, that is equal to 1 if the individual is the household of the family
DUMMY_PART_TIME	0,094	0,292	Dummy variable for part time work
DUMMY_AGRICULTURAL	0,034	0,180	Dummy variable for agricultural sector
DUMMY_INDUSTRIAL	0,312	0,463	Dummy variable for industrial sector
DUMMY_PUBLIC	0,320	0,466	Dummy variable for public administration sector
DUMMY_OTHER_SECTOR	0,335	0,472	Dummy variable for other sector
DUMMY_TOWN	0,083	0,275	Dummy variable for the town of residence that has more than 500.000 inhabitants
DUMMY_NORTH	0,501	0,500	Dummy variable for North regions
DUMMY_CENTER	0,214	0,410	Dummy variable for Center regions
DUMMY_SOUTH	0,286	0,452	Dummy variable for South regions
SCHOOL_F	6,094	4,094	Schooling attainment of the father's worker
SCHOOL_M	5,346	3,711	Schooling attainment of the mother's worker

2.4 Results

Table 2.2 presents the estimates for the first stage regression (Equation 2.3). The first stage model suggests that the individual's education level is correlated with father's and mother's education level. The high statistical significance of the instrumental variables suggest that the residuals of Equation (2.3) are independent of these two variables. This ensures that the restriction for applying the control function approach in Equation (2.5) is satisfied.

The coefficients of the correction terms provide a direct test for the presence of "selection bias" (endogeneity) and of "return bias" induced by sorting gains (self-selection). The coefficient of the control function significant, indicating that schooling is endogenous, and negative. One possible explanation is that the correction term picks up the correlation between 'ability' and education. In this case, we would interpret the result as a signal of 'negative selection' into education: individuals with higher absolute unobservable wages (say, ability) would be less likely to get high education levels. However, the negative correlation between education and unobservable earnings may simply reflect a downward bias (in both the linear and the semiparametric model without correction term) induced by large measurement errors in education, as it is likely to be the case, given that they are imputed.

2.4.1 The Return on Schooling

Figure 2.1 shows that the marginal effects of schooling from 1995 to 2012 with 95 percent confidence limits, shown as dashed lines, which are estimated from the linear model (LM) and from the generalized additive model (GAM) with the control function. Significant non-linearity emerges from the semiparametric estimation of schooling that draw the shape of the entire wages-education profile

Based on Figure 2.1, increasing returns are evident for workers until 8 years of schooling (junior high school), from 1995 to 2004; however, in the following years they show a flat pattern. If we consider worker with almost 13 years of schooling (secondary school), the marginal effects across year continue to increase. On the other hand, decreasing returns are observed for workers with 18 years of schooling (tertiary education), from 2008 to 2012.

Overall, we find that the rate of return on schooling ranges from 4 percent for low level of schooling to 8 percent for medium-high level of schooling. This finding provides a strong support to any policy of incentive/support to tertiary education.

2.4.2 The Experience Profile

Figure 2.2 shows the estimated function for the experience. Also in this case, the semiparametric approach (GAM with CF) demonstrates a stronger power in comparison to the quadratic specification of the standard Mincer wage function (LM with CF).

In particular, we observe different pattern for each year of the sample: from 1995 to 2008 the experience profile is approximately a concave function, more or less steeper, while in 2010 and 2012 it is approximately a linear function.

2.4.3 Control Variables

All the control variables enter linearly in the semiparametric model. The estimates confirm what we have found in the first chapter as regards the sign of the impact.

In particular, for the DUMMY_MALE variable, we observe a strong evidence of a gender pay gap, in favor of men for all the period considered.

Considering the geographical variables (DUMMY_NORTH and DUMMY_SOUTH), we observe that the DUMMY_NORTH is positive while the other one in negative. This means that it is more convenient to work in the north regions in comparison to the central regions, instead if an individual works in the south region he will earns less than in the center regions.

Finally, considering different sector of employment, working in the agricultural sector is less convenient than working in the industrial sector. On the contrary, working in the public sector is more profitable than working in the industrial sector.

Figure 2.1 – Estimated schooling

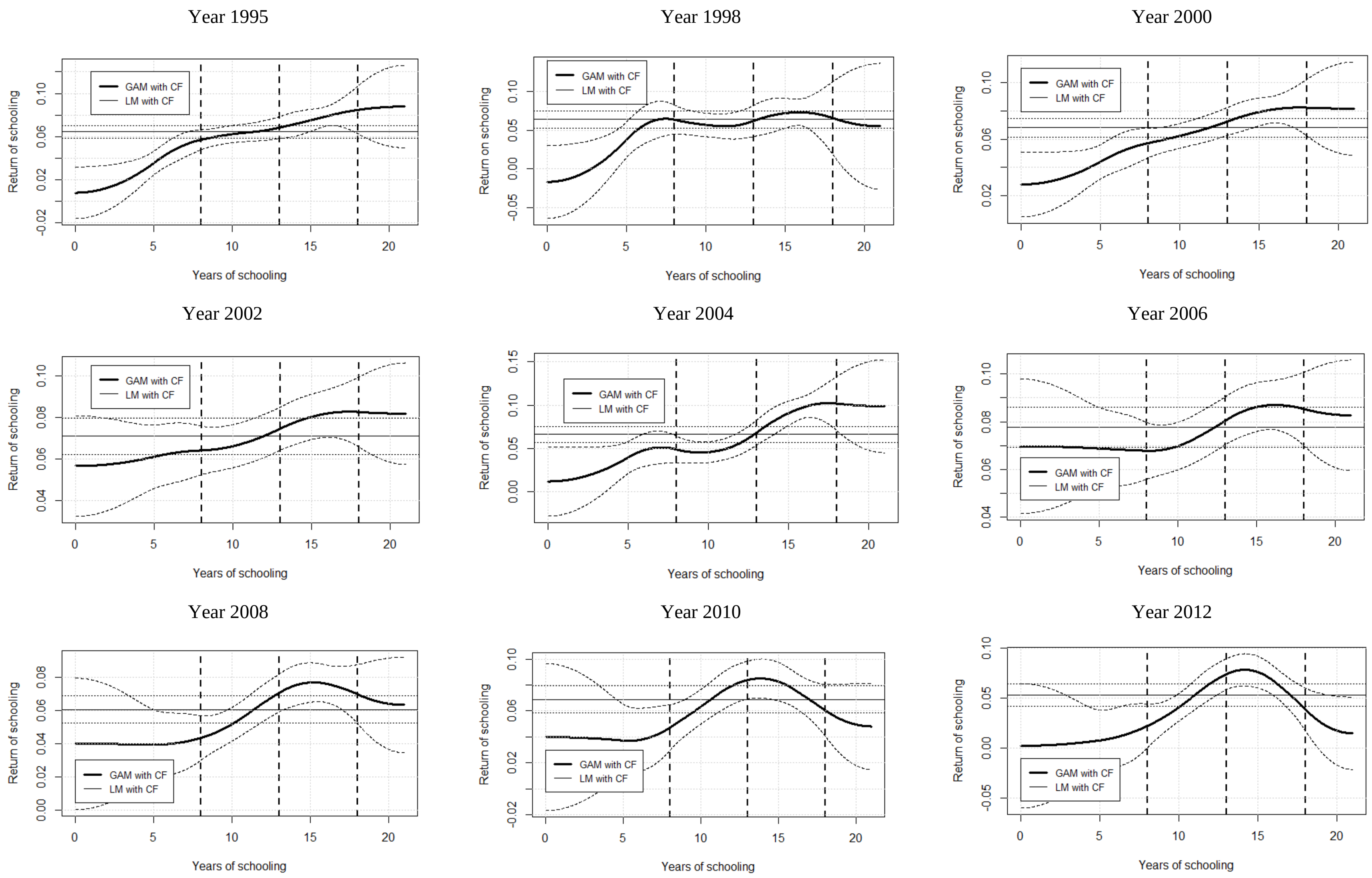


Figure 2.2 – Estimated experience

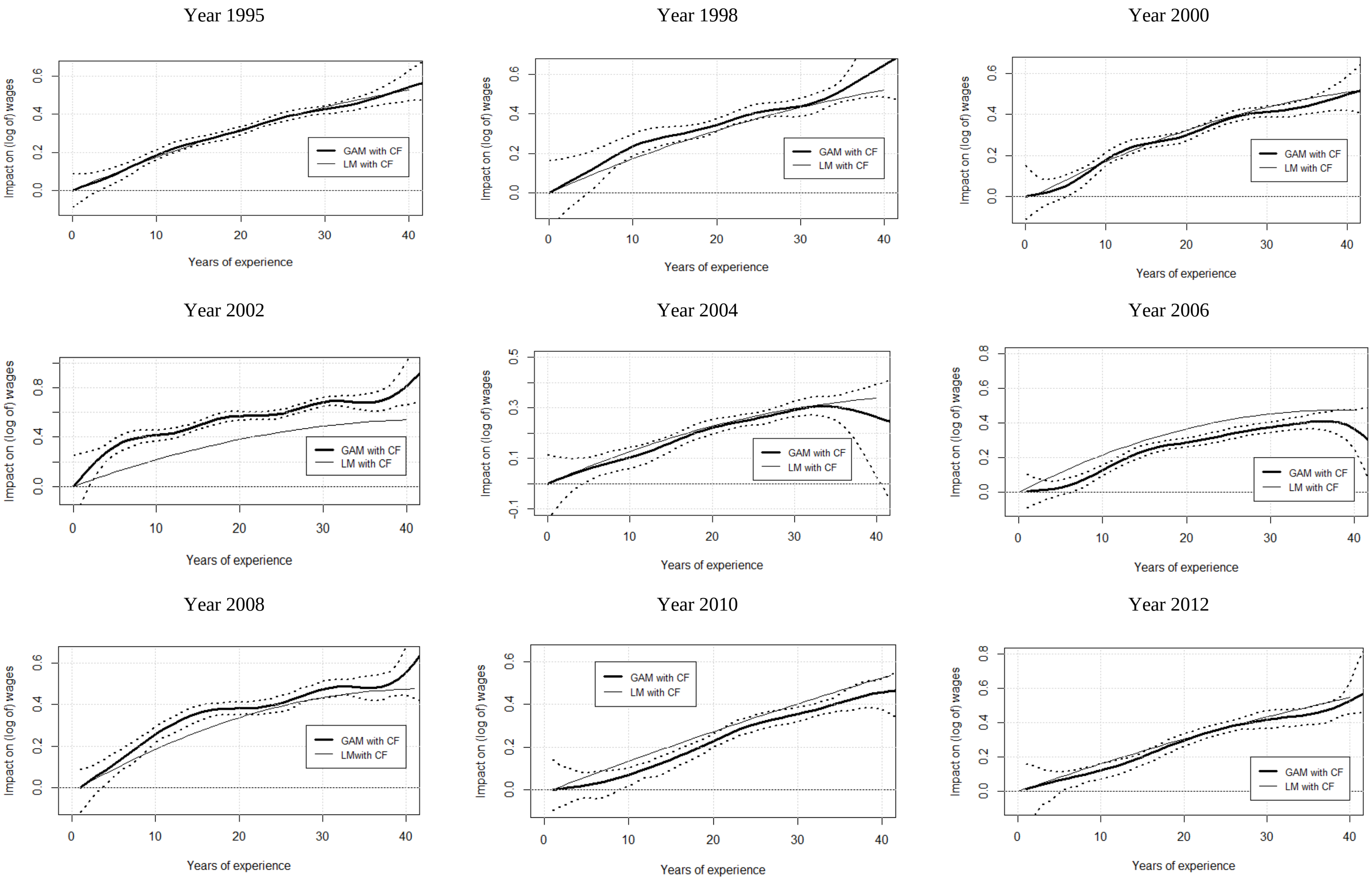


Table 2.2 - First Stage GAM Estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are Center (DUMMY_CENTER) and Industrial sector (DUMMY_INDUSTRIAL), 1995-2012.

	1995	1998	2000	2002	2004	2006	2008	2010	2012
Parametric coefficients:	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*
Constant	9,741 ***	11,245 ***	9,792 ***	9,507 ***	9,894 ***	10,480 ***	10,477 ***	11,056 ***	11,351 ***
DUMMY_MALE	0,265 *	0,368	-0,302 **	-0,415 ***	-0,419 ***	-0,342 ***	-0,341 **	-0,588 ***	-0,399 ***
DUMMY_MARRIED	0,521 **	-0,356	-0,017	0,290	0,644 ***	0,472 ***	0,849 ***	-0,076	0,085
NCOMP	-0,154 **	-0,192 **	-0,002	0,086	-0,021	-0,013	-0,166 **	0,215 ***	0,043
DUMMY_HOUSEHOLD	-0,162	-0,612 **	0,174	-0,027	0,112	0,181	0,076		
DUMMY_TOWN	0,254 *	-0,068	0,900 ***	0,490 ***	0,091	0,423 **	0,384 **	0,414 **	0,662 **
DUMMY_NORTH	-0,169	-0,013	0,232 *	0,377 ***	0,055	-0,258 *	0,184	-0,250	-0,445 **
DUMMY_SOUTH	-0,199	0,024	0,040	0,255	-0,350 **	-0,655 ***	-0,245	-0,307	-0,351 *
DUMMY_AGRICULTURAL	-1,222 ***	-2,611 ***	-1,411 ***	-1,522 ***	-1,381 ***	-0,742 ***	-1,020 ***	-1,150 ***	-0,899 **
DUMMY_PUBLIC	2,361 ***	2,241 ***	2,507 ***	2,491 ***	2,465 ***	2,431 ***	2,384 ***	2,236 ***	2,502 ***
DUMMY_OTHER_SECTOR	-0,002	0,486 **	0,792 ***	0,489 ***	0,865 ***	0,742 ***	0,645 ***	0,637 ***	0,544 ***
DUMMY_PART_TIME	-0,662 ***	-0,605 *	-0,614 ***	-0,526 ***	-0,797 ***	-0,884 ***	-0,566 ***	-0,811 ***	-0,608 ***
Non Parametric coefficients:	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*
SCHOOL_F	2,907 ***	2,712 ***	2,838 ***	1,012 ***	2,607 ***	2,763 ***	2,813 ***	2,625 ***	2,767 ***
SCHOOL_M	2,371 ***	2,151 ***	1,001 ***	2,845 ***	2,331 ***	1,003 ***	1,979 ***	1,003 ***	2,575 ***
EXPERIENCE	4,812 ***	1,003 *	1,849 ***	1,782 ***	2,845 **	1,009 ***	1,002 ***	1,003 ***	4,118 ***
R-sq.(adj)	0,413	0,406	0,389	0,373	0,366	0,362	0,365	0,320	0,343
Dev. Exp.	41,5%	41,3%	39,2%	37,6%	37,0%	36,5%	36,8%	32,4%	34,9%
REML score	11.84	3.981	10.324	8.977	9.075	9.094	7.580	5.943	5.734
Scale est.	9,531	8,360	9,739	10,224	9,435	8,398	9,907	11,505	10,908
Obs.	4.352	1.468	3.783	3.321	3.405	3.437	2.836	2.145	2.112

(***p<0.01, **p<0.05, *p<0.1)

Table 2.3 - Second Stage GAM Estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are Center (DUMMY_CENTER) and Industrial sector (DUMMY_INDUSTRIAL), 1995-2012.

	1995	1998	2000	2002	2004	2006	2008	2010	2012
Parametric coefficients:	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*
Constant	2,120 ***	2,115 ***	2,195 ***	2,238 ***	2,218 ***	2,220 ***	2,194 ***	2,214 ***	2,184 ***
DUMMY_MALE	0,131 ***	0,116 ***	0,095 ***	0,097 ***	0,075 ***	0,109 ***	0,150 ***	0,155 ***	0,103 ***
DUMMY_MARRIED	0,007	0,054	0,055 ***	0,004	0,038 *	-0,010	-0,049 **	0,025	0,005
NCOMP	0,017 ***	0,017 *	0,000	0,003	-0,004	0,031 ***	0,029 ***	-0,001	0,021 ***
DUMMY_HOUSEHOLD	-0,006	-0,002	0,009	0,024	0,022	0,034 **	0,009		
DUMMY_TOWN	0,004	0,032	0,007	-0,082 ***	-0,016	0,041 **	0,003	-0,035	-0,011
DUMMY_NORTH	0,038 ***	0,067 **	0,046 ***	0,046 **	0,060 ***	-0,008	0,000	0,055 ***	0,046 **
DUMMY_SOUTH	-0,027 *	0,056 **	-0,014	0,005	0,006	-0,053 **	-0,034 *	0,018	-0,008
DUMMY_AGRICULTURAL	-0,080 ***	-0,040	-0,148 ***	-0,056 *	-0,099 ***	-0,135 ***	-0,067 *	0,012	-0,118 **
DUMMY_PUBLIC	0,110 ***	0,047	0,028 *	0,009	0,064 ***	0,018	0,081 ***	0,058 **	0,070 **
DUMMY_OTHER_SECTOR	0,015	0,006	-0,003	-0,015	-0,005	-0,026 *	0,002	0,002	-0,049 **
DUMMY_PART_TIME	0,034	0,080 **	0,081 ***	-0,054 **	-0,017	0,016	0,012	0,041	-0,021
FIRST STAGE_RES	-0,013 ***	-0,014 **	-0,025 ***	-0,027 ***	-0,022 ***	-0,035 ***	-0,017 ***	-0,028 ***	-0,016 ***
Non Parametric coefficients:	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*	*E.D.F.*
SCHOOL	3,895 ***	3,630 ***	3,401 ***	2,584 ***	3,776 ***	2,652 ***	3,212 ***	3,548 ***	4,000 ***
EXPERIENCE	2,874 ***	2,284 ***	4,539 ***	6,407 ***	2,747 ***	5,510 ***	6,319 ***	1,345 ***	2,614 ***
R-sq.(adj)	0,420	0,317	0,345	0,299	0,250	0,338	0,352	0,295	0,323
Dev. Exp.	42,3%	32,7%	34,9%	30,4%	25,4%	34,2%	35,8%	30,1%	32,9%
REML score	2.005	834	1.792	2.042	2.168	1.664	1.388	1.268	1.199
Scale est.	0,102	0,110	0,105	0,153	0,160	0,109	0,122	0,142	0,145
Obs.	4.352	1.468	3.783	3.321	3.405	3.437	2.836	2.145	2.112

(***p<0.01, **p<0.05, *p<0.1)

2.4.4 The choice between Education and Experience

In this section, we simulate a counterfactual scenario: one-year education expansion for all individuals of the sample, holding other variables constant except experience (one year less), in order to assess the impact on the wages. Therefore, we calculate:

$$\ln w^{CF}(s_i, x_i) = \alpha_{0i} + f_1(s_i + 1) + f_2(x_i - 1) + \varepsilon_i, \tag{2.6}$$

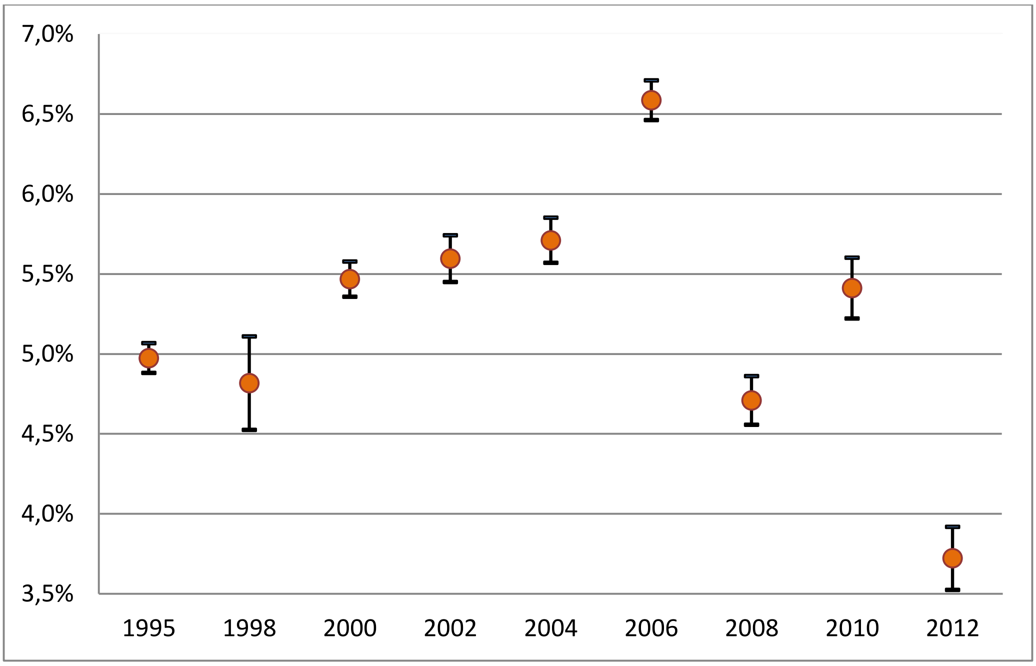
where w^{CF} is the counterfactual wage of an individual, where we simulate that the individual has attained one more year of schooling, and as a consequence one year less of experience, because of the postponement of entrance in the labor market. In particular, we calculate the difference between the expected counterfactual and the predicted wage (Equation 2.2) by the model:

$$\ln w^{\,CF}(s_i, x_i) - \ln w(s_i, x_i) = \ln \frac{w^{CF}(s_i, x_i)}{w(s_i, x_i)} \approx \frac{w^{CF} - w}{w} =$$
$$= (f_1(s_i + 1) - f_1(s_i)) + (f_2(x_i - 1) - f_2(x_i)) \tag{2.7}$$

Applying the semi-parametric estimates of the coefficients of the Mincer wage equation, we estimate the variation on the wage given in Equation (2.7). Our results show that the impact is positive (Figure 2.3) for all the years considered, but the magnitude is different. In particular, the impact is around 5 percent in 1995 and 1998, then increases until 2006 reaching the highest level at 6.6 percent and then decrease to 3.7 in 2012.

By a policy maker perspective, having one year education expansion in Italy, is powerful to have a positive impact in the wages, but the advantage is decreasing in the last period of the analysis.

Figure 2.3 – Estimated impact of one more year of education (and one less of experience) on the log of hourly wages with 95% confidence intervals



2.5 Concluding Remarks

The empirical returns to schooling is an important information for policy maker. In particular, understanding how the returns estimations vary with the level of schooling attainment could be important to a better tuning of the educational policies. Policy interest focuses not only on the average returns to education but also on the dispersion of returns across education levels. The shape of the wages function is a key factor for understanding how policies of education expansion will affect incomes.

Most empirical studies assume that (log of) wages are a linear function of the years of education and a quadratic function of the years of experience.

We have provided evidence that such parsimonious parametric model misses important features of the true relationships. In our results, linearity can typically be firmly rejected, thus marginal returns will differ from the average. Increasing returns are discovered for workers until 8 years of schooling (junior high school), from 1995 to 2004; however, in the following year they show a flat pattern. If we consider worker with almost 13 years of schooling (secondary school), the marginal effects across year continue to increase. On the other hand, decreasing returns are observed for workers with 18 years of schooling (tertiary education), from 2008 to 2012.

Considering experience, from 1995 to 2008 the experience profile is approximately a concave function, more or less steeper, while in 2010 and 2012 it is approximately a linear function. Therefore, semiparametric approach is superior to parametric ones in terms of flexibility and of predictability, even if partial coincident for some linear variables.

Even if the return to education is increasing with years of schooling, the absolute level is below that of other countries, i.e. USA and UK (Trostel et al., 2002; Harmon et al., 2003).

A future line of research should remove the assumption that post-school investment are identical across persons and do not depend on the schooling level (separability). The independence between education and the return to experience, typically illustrated by the fact that age earnings profiles are approximately parallel across broad education groups, is also being questioned (Heckman, Lochner and Todd, Lemieux, 2003). This suggests that log wages regression may not be separable in education and experience and, in particular, that the return to experience may be affected by schooling. Various economic models[14] may be able to explain this. These include models of endogenous post-schooling human capital investments as well as various lifecycle incentive models where wages are upward sloping.

[14] It is possible to use Mincer's approach to derive alternative wages equations that do not require that the investment ratio is independent of schooling.

As stated in this chapter, the semiparametric approach is more flexible to the parametric ones, so it would be interesting to estimate the effect of the interaction between education and experience applying a semiparametric estimation with a control function to deal with the endogeneity of schooling.

Chapter 3

An Estimate Across Quantiles of the Mincer Wage Function in Italy

Abstract.

Several studies focused on the estimation of the average impact of schooling, experience and other variables on wages without investigate if they affect individuals differently over the wage distribution. Understanding the heterogeneity of education is relevant because allows to test if education can reduce or increase inequality. In order to take into account simultaneously endogeneity and heterogeneity of education, an instrumental variables quantile regression is applied. Our results show that, while returns to schooling are positive everywhere, there exists a large degree of heterogeneity in returns to education across the wage distribution. In particular, gains are higher for individuals in the upper tail of the wages distribution than for those in the lower tail. This means that education have an inequality-increasing effect over time, because individuals with high ability, those at the upper quantile of the wage distribution, seem to benefit more from formal education.

3.1 Introduction

Many studies have investigated the average impact of several determinants, such as schooling, experience, gender, etc. on wages (Dickson and Harmon, 2011, Cingano and Cipollone, 2009). This chapter aims to understand if individuals in different quantile of the wage distribution are differently affected by these determinants.

Applying a quantile regression, we test the hypothesis of heterogeneous effects of schooling, experience, gender, etc. on wages, i.e. if the effects of the variables are increasing, decreasing, or u-shaped across the quantiles. In addition, we assess the relationship between education and wage inequality and the changes of trends during the period under analysis. Indeed, if returns to schooling are heterogeneous along the wage distribution, schooling can have an impact upon wage inequality.

In a simple human capital model, wage inequality can increase because returns to education and experience increase, or because residual or within-group inequality increases (Lemieux, 2008). In the case that returns are increasing from the lower to the higher end of the wage distribution, it can be interpreted as an indication that ability and education (or skills) are complement between them, and more able workers can benefit from additional investment in education. Consequently, a negative relationship between ability and returns to education (decreasing returns with quantile) can be interpreted as evidence of substitutability between education and ability. Finally, if there is no distinct pattern, then average returns (in the absence of biases in their estimation) capture the overall profitability of education (Patrinos et al., 2006, Chernozhukov, Hasen and Janson, 2007).

However, two potential issues complicate the estimation of the effects of education on the whole wage distribution: the endogeneity of education attainment and heterogeneity in the returns to education. The first issue concerns the causal effects of education on wages. Although there is little doubts that education plays an important role in determining individuals' wages, the estimation of the causal effects of education on wages is not straightforward due to potential endogeneity and measurement error problems. Indeed, more able workers may get more education as well as earn more in the labor market. In this case, the positive observed relationship between education and wages may be driven by a third variable, namely ability. Moreover, information on schooling gathered during surveys may also be misreported. Therefore, not controlling for observable and unobservable determinants can preclude estimation of the causal inference of the underlying effects of education on wages.

Previous empirical studies have typically relied on regression analysis and linear specification, focusing mainly on average effects. Although of great interest, the average effects can hide important information of the wage distribution. Moreover, average estimates do not capture information about inequality effects of education. For example, if the education positively affects more the upper tail of the wages distribution than education increases inequality rather than decreases it. Therefore, in order to foster equality through education, schooling should increase wages more for individuals in the lower tail of the wages distribution than for those in the upper tail. When the average effects are the only available information, it is not clear whether expanding educational opportunities will increase or decrease inequality.

Several econometrics models deal with endogeneity and heterogeneity issues simultaneously (Card, 1999; Arias et al., 2001). However, empirical studies often deal with only one issue at a time. To overcome the endogeneity, the most applied approach is instrumental variable estimation (Angrist and Krueger, 1991; Harmon and Walker, 2000; Trostel et al., 2002; Dickson, 2013). Instead, to deal with the heterogeneity issue, researchers rely on different methods to account it to study the returns to education, i.e.: random coefficient model (Harmon et al., 2003), nonparametric estimation (Henderson et al., 2011), Bayesian hierarchical models

(Koop and Tobias, 2004) and quantile regression (Arias et al., 2001; Martins and Pereira, 2004; Fasih et al., 2012). Most of these studies indicate that the impact of education on wages is far from homogeneous. Then, the population does not seem to be reasonably described by a single parameter for the relation between wages and education.

Harmon et al. (2003) utilize random-coefficient models to estimate the variance of returns to education, where returns to education were the random coefficient in the Mincer wage equation. They find that the dispersion of returns to education was quite high in the UK, and the dispersion of individual returns remained stable during the 1990s.

Henderson et al. (2011) employ generalized non-parametric kernel estimation to estimate heterogeneous rates of return across different demographic groups in the USA. They find that the non-parametric median rate of schooling return for US workers increased significantly in the long run, from 8.2 percent in 1950 to 14.3 percent in 2005.

Using data from the US, Koop and Tobias (2004) estimate Bayesian hierarchical models to investigate the nature of heterogeneity in returns to schooling. They not only found strong evidence of heterogeneity in schooling returns, but also noted that it followed a continuous distribution, rather than a discrete one.

For the purpose of our research, the more appropriate method is the quantile regression. This econometric model allows to estimate the returns to education over the wage distribution considering the heterogeneity through quantile-specific intercepts and quantile-specific slopes. However, to overcome endogeneity of schooling, recent studies by Chernozhukov and Hansen (2006; 2008; 2013) have proposed an instrumental variable quantile regression (IVQR) approach to estimate rates of return within a distributional framework that addresses both heterogeneity and endogeneity issues at the same time.

In this chapter, that approach is applied to estimate the effect of schooling, experience and other variables on the entire distribution of wages in Italy for the period 1995-2012. Finally, the evolution over time of the quantile returns to education and what impact the returns have on the structure of wages are investigated.

The reminder of the chapter is organized as follows. Section 3.2 introduce individual heterogeneity in the Mincer wage equation and present a summary of the empirical literature review on that topic. Section 3.3 describes the dataset used in the empirical estimation and the characteristics of the sample. Section 3.4 describes the econometric methods used to estimates returns to education. Section 3.5 presents the results, and Section 3.6 concludes.

3.2 Individual Heterogeneity in the Mincer Wage Function

In the first Chapter (in particular, in Section 1.2), the Mincerian wage equation has been derived as follow:

$$\ln w(s_i, x_i) = \alpha_{0i} + \rho_s s_i + \beta_0 x_i + \beta_1 x_i^2 + \varepsilon_i, \tag{3.1}$$

where ε_i is a mean zero residual with $E(\varepsilon_i | s_i, x_i) = 0$, and $\alpha_0 = \ln E_0 - \kappa(1 + \rho_{ps})$, $\beta_0 = \kappa \left[\rho_{ps}\left(1 + \frac{1}{2T}\right) + \frac{1}{T}\right] - \delta$, $\beta_1 = -\frac{\rho_{ps}\kappa}{2T}$.

Individual heterogeneity (talents) potentially affects both the intercept of the wage equation (through α_{0i}) and the slope of the wage-education relation (through ρ_s) in Equation (3.1). Therefore, three hypothesis can be tested. First, evidence of 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 pointed out by Ashenfelter and Rouse (1998), more able workers acquire more schooling 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 estimate the true impact of education on wage without solving the bias introduced by the endogeneity of schooling attainment, since cross-sectional estimates are (marginally) biased by an omitted ability variable (Heckman et al., 2006). Third, we want to study how education affects individuals differently taking into account both heterogeneity and endogeneity issues simultaneously. To incorporate these features, an instrumental variable quantile regression is applied, which estimates the causal effect of education on conditional quantiles of the wage distribution, allowing for quantile-specific intercepts and quantile-specific slopes.

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

$$Q_{\ln(w)}[\tau|s,x] = \alpha_0(\tau) + \rho_s(\tau)s + \beta_0(\tau)x + \beta_1(\tau)x^2, \tag{3.2}$$

where ρ_s is the return to schooling at the τ^{th} quantile, and $\tau \in (0,1) \rightarrow \alpha_0(\tau) + \rho_s(\tau)s + \beta_0(\tau)x + \beta_1(\tau)x^2$ is strictly increasing in τ. In Equation (3.2) the returns to education, experience and the intercept are function of τ, allowing for heterogeneous effects of these variables on wages.

The existing literature uses conventional quantile regression to investigate the heterogeneous effects on wage. Koenker and Basset (1978) first introduced the quantile regression model. Since then, several authors have used this framework to explore the wage effects of schooling over the entire wage distribution.

Buchinsky (1994) for US, shows that education is more profitable at the top of the distribution. This finding helps explain the rapidly increasing earnings inequality associated with rising rewards for educational qualifications at a time of schooling expansion.

Martins and Pereira (2004), utilizing quantile regression techniques, show that for nearly all EU countries (but Denmark, Germany and Italy could be considered borderline cases), returns to schooling are significantly higher at the top of the wage distribution. Individuals who are the top of the conditional wage distribution are there because of their unobserved characteristics and the results suggest that this group receives higher education increment. Their results imply that schooling aggravates within group inequality. They explain that factors such as over-education, ability– schooling interactions and school quality or different fields of study may be driving this result.

Giustinelli (2004) applies quantile regressions to investigate the dynamic of educational wage premia over the period 1993–2000 for Italy. The main result is that the schooling premium shows a U-shaped pattern across the wage distribution in each sample year. Naticchioni et al. (2009), by means of quantile regressions, show that educational wage premia in the private sector decline across the entire wage distribution, for the period 1993-2004 in Italy. Patrinos et al. (2009) and Fasih et al. (2012) find the same evidence i.e. that returns increase across quantiles for Latin American countries, while returns decrease across quantile for most East Asian countries (the exception is Singapore, a high-income country). Hartog et al. (2001) examine the evolution of the returns to education in Portugal over the 1980s and early 1990s. They apply a quantile regression analysis and they find that returns are higher for those at higher quantiles in the conditional wage distribution, and education has an important role in the expansion of wage inequality in Portugal.

Carrasco et al. (2014) find that the compression of the wage distribution in Spain between 1995 and 2006 is largely explained by a decrease in the returns to education, due to an increase in the supply of high-skilled workers and the increasing weight of low-skilled occupations. In contrast, the widening of the wage distribution after 2006 is largely explained by an increase in the relative demand for high-skilled workers generating an increase in the school premium.

From all 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 tend to be increasing in the quantiles of wage distribution for developed countries. This is interpreted as a positive impact of education within-groups inequality. However, all this kind of studies ignore the issue of endogeneity.

3.3 Data and Sources

The analysis is based again on data drawn from the Bank of Italy's Survey of Household Income and Wealth (SHIW), which reports several socio-economic characteristics of Italian households.

The SHIW is a biannual survey on Italian families with a sample of approximately 8,000 household per year. From 1995 to 2012 observations from nine subsequent surveys are available. In particular, the SHIW contains information both on households (family composition) and on individuals. Moreover, it provides detailed information on several characteristics of workers within each household, such as their net yearly wages, average weekly hours of work and number of months of employment per year, educational attainment (the highest completed school degree), job experience, gender, marital status, sector of employment, household composition, parents background, regions of residence, and town size.

We consider a sub-sample of men and women between 15-64 years old, full time and part time employees, working either in the public or in the private sector and such that information about wages are available. In the analysis, we exclude self-employed because of the low reliability of their declared earnings. As discussed by Brandolini and Cannari (1994), SHIW seems to underestimate the self-employed earnings of about 50 percentage points.

3.3.1 Variables Used in the Analysis

The variables used in the empirical model are the same introduced previously. For the sake of clarity, discussion about them is repeated.

As shown by Equation (3.1), wages, schooling attainment, and working experience of each individual are the key variables in the estimation of Mincer equation.

Mincer equation refers to the (log of) hourly price of labor as correct measure of worker's wages (LOGY_H), and, indeed, this is the measure used by most empirical studies[15] (Brunello and Miniaci, 1999; Blundell, Dearden and Sianesi, 2005; Ciccone, Cingano and Cipollone, 2006). SHIW contains yearly net wages of taxes and social security contributions. Additional information on the average number of hours worked per week and on the number of months worked per year, can be used to estimate the hourly net wage, which is calculated by yearly net wages divided by months worked multiplied by hours worked each month.

Schooling attainment (SCHOOL) is generally measured by the number of years spent at school. SHIW does not contain information about this number of years, but only on the highest degree attained by individuals. Following a common approach in literature (Vieira, 1999; Brunello and Miniaci, 1999) we calculate the educational attainment of the individual by imputing the number of years required to complete her/his reported maximum level of educational attainment[16]. More precisely, we consider that the (statutory) numbers of years required to obtain a primary and a junior school certificate is 5 and 8 years respectively; instead, for

[15] Hourly wages can be affected by measurement errors because we calculate them as total wages divided by hours of work.

[16] Standard, not actual, years of formal schooling are recorded. Since students who fail to reach a standard have to repeat the year, the actual number of years is likely to be underestimated.

the upper secondary school the number of years ranges from 11 (vocational or technical school) to 13 (classical or scientific studies); finally, for tertiary education, we consider 16, 18 and 21 years for the university diploma, the college degree, and the postgraduate degree (e.g. Ph.D.) respectively. It is important to remark that in Italy the statutory number of years can be significantly different from the actual number of years spent to obtain a degree, especially at college because of the high percentage of irregular student.

Many empirical studies use age as a proxy for the (working) experience of individuals. But this choice can be severely biased, especially for young cohorts. Other authors use potential experience, defined as the difference between the current age and the age at the labor market entry, but they ignore the possibility of unemployment or underemployment, again a crucial feature for young cohorts. In this work we use, as proxy for experience (EXPERIENCE), the number of years for which a worker has been paid social security contribution; they should reflect the effective years of training on the job and learning-by-doing activities.

We introduce several control variables in the analysis to account for individual characteristics and for differences in the labor market.

A gender dummy (DUMMY_MALE) controls for different wage levels between men and women. Marital status also enter into the analysis as a dummy variable (DUMMY_MARRIED) taking the value 1 if the person is formally married, 0 otherwise. Part-time work is captured through a separate dummy variable (DUMMY_PART_TIME), since the assumption that each working hour makes the same contribution to weekly wages (constancy of the hourly wage) may not hold across workers with different time status (part time versus full time).

In addition, controls are introduced for family composition, as a proxy for the influence of housework, particularly important in the female labor supply (Heckman and Killingsworth, 1986). We control for the number of components of the family (NCOMP) and for the fact that the individual is the head of his/her household (DUMMY_HOUSEHOLD).

Controls for sector (DUMMY_AGRICULTURAL for the agricultural sector, DUMMY_INDUSTRIAL for the industrial sector, DUMMY_PUBLIC for the public sector and DUMMY_OTHER_SECTOR for other sector different from the previous ones) should capture potential factors from the demand side of labor market (e.g. imperfectly competitive labor markets). In the same light, we add some controls for the geographical area of residence: one dummy for the town of residence that has more than 500.000 inhabitants (DUMMY_TOWN), and three different dummies for the Italian macro-regions: North, Center and South (DUMMY_NORTH, DUMMY_CENTER and DUMMY_SOUTH).

Table 3.1 reports some descriptive statistics of the main variables used in the empirical analysis for all the waves (wages are expressed in euro 2012).

Table 3.1 - Means and standard deviations of the variables used in the empirical analysis for the entire sample (1995, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012)

Variable	Mean	S. d.	Description
LOGY_H	2,265	0,438	Logarithm of the hourly wages less tax
SCHOOL	11,373	3,800	Schooling attainment, that is the number of years spent at school
EXPERIENCE	17,683	10,673	Number of years for which it has been paid social security contributions, as a proxy for years of training on the job
DUMMY_MALE	0,578	0,494	Gender dummy
DUMMY_MARRIED	0,647	0,478	Dummy variable for marital status
NCOMP	3,329	1,185	Number of components of the family
DUMMY_HOUSEHOLD	0,475	0,499	Household dummy, that is equal to 1 if the individual is the household of the family
DUMMY_PART_TIME	0,094	0,292	Dummy variable for part time work
DUMMY_AGRICULTURAL	0,034	0,180	Dummy variable for agricultural sector
DUMMY_INDUSTRIAL	0,312	0,463	Dummy variable for industrial sector
DUMMY_PUBLIC	0,320	0,466	Dummy variable for public administration sector
DUMMY_OTHER_SECTOR	0,335	0,472	Dummy variable for other sector
DUMMY_TOWN	0,083	0,275	Dummy variable for the town of residence that has more than 500.000 inhabitants
DUMMY_NORTH	0,501	0,500	Dummy variable for North regions
DUMMY_CENTER	0,214	0,410	Dummy variable for Center regions
DUMMY_SOUTH	0,286	0,452	Dummy variable for South regions
DUMMY_SETT_GEN	0,374	0,484	Dummy variable equal to 1 if the individual works in the same sector of the father and/or of the mother
SCHOOL_F	6,094	4,094	Schooling attainment of the father's worker
SCHOOL_M	5,346	3,711	Schooling attainment of the mother's worker

3.4 Methodology for the Estimates

Many previous applied econometrics studies have typically relied on regression analysis and linear specification, thereby focusing mainly on average effects. While of interest, the average effects may hide important information in the rest of the wage distribution. Many variables, such as wages, have continuous distributions, and these distributions can change in response to treatments in ways not fully revealed by averages. In the following, we briefly describe the IVQR method that we use to estimate the causal heterogeneous returns to education across the wages distribution.

In the ordinary quantile regression method, assume that the error term in the wage function, Equation (3.2), is 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 $\rho_s(\tau)$ and $\alpha_0(\tau)$, $\beta_0(\tau)$, $\beta_1(\tau)$ in Equation (3.2) by solving the following minimization problem:

$$Q_{\ln(w)}[\tau|s,X] = arg \min_{\rho_s(\tau),\beta(\tau)} E[\varphi_\tau\, (\ln(w) - \alpha_0(\tau) - \rho_s(\tau)s - \beta_0(\tau)x - \beta_1(\tau)x^2], \quad (3.3)$$

where $\varphi_\tau(\varepsilon_i)$ is the "check function" defined as $\varphi_\tau(\varepsilon_i) = [\tau - 1(\varepsilon_i \leq 0)]\varepsilon_i$.

Assuming independence between education variable and the error term may be too stringent because of potential unobserved wage determinants. In this respect, to account for potential dependence between s and u in a distributional framework, as aforementioned, we apply the instrumental variable quantile regression approach developed by Chernozhukov and Hansen (2006, 2008, 2013).

As in the case of two-stage least squares, the identification of this approach relies on the existence of a vector of instrumental variables z that is statistically related to s but independent of the error term. In addition, 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"). 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 (2006) show that assuming rank similarity implies the following moment condition:

$$P\big[(\ln(w) \leq Q_{\ln(w)}(\tau|x,z)\big|x,z)\big] = \tau; \tag{3.4}$$

$$P[(\ln(w) - \alpha_0(\tau) - \rho_s(\tau)s - \beta_0(\tau)x - \beta_1(\tau)x^2 \leq 0|x,z)] = \tau. \tag{3.5}$$

The moment condition given in Equation (3.5) provides a statistical restriction that can be used to estimate the parameters $\rho_s(\tau)$, $\beta_0(\tau)$, $\beta_1(\tau)$ and $\alpha_0(\tau)$. Equation (3.5) is equivalent to the statement that zero is the τ^{th} quantile of the random variable $\ln(w) - Q_{\ln(w)}(\tau|x,s)$ conditional on (x,z). Chernozhukov and Hansen (2008) formulate the problem as finding $[\rho_s(\tau)\,, \beta_0(\tau), \beta_1(\tau)\,, \alpha_0(\tau)]$ such that zero is the solution to the standard quantile regression of $[\ln(w) - \alpha_0(\tau) - \rho_s(\tau)s - \beta_0(\tau)x - \beta_1(\tau)x^2]$ on (x,z):

$$\begin{aligned} 0 = arg\min_{f\in F} E[\varphi_\tau\,(\ln(w) - \alpha_0(\tau) - \rho_s(\tau)s - \beta_0(\tau)x - \\ \beta_1(\tau)x^2 - f(x,z)], \end{aligned} \tag{3.6}$$

where *F* is the class of measurable functions of (x,z). In empirical application, *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 $\rho_s(\tau)$, we look for a value ρ_s that makes the estimated coefficient on the instrumental variable $\hat{\gamma}(\tau,\rho_s)$ in Equation (3.6) as close to zero as in ordinary quantile regression.

The IVQR estimator consists of a two-step procedure:

i) for a given value of $\rho_s^j(\tau)$, run the ordinary quantile regression of $\ln(w_i) - \rho_s^j(\tau)s_i$ on x_i and z_i to obtain estimates $\hat{\alpha}_0(\rho_s^j(\tau),\tau), \hat{\beta}_0(\rho_s^j(\tau),\tau),\ \hat{\beta}_1(\rho_s^j(\tau),\tau), \hat{\gamma}(\rho_s^j(\tau),\tau)$;

ii) then test $\hat{\gamma}\left(\rho_s^j(\tau), \tau\right) = 0$ and save the corresponding Wald Statistics W^j.

Then the estimation procedure has to be repeated these two steps for all the values in a pre-specified support for $\rho_s^j(\tau)$ and the values that minimizes the F-statistic is the IVQR estimator of $\widehat{\rho_s}(\tau)^{IVQR}$ and the corresponding $\widehat{\alpha}_0\left(\rho_s^j(\tau), \tau\right)$, $\hat{\beta}_0\left(\rho_s^j(\tau), \tau\right)$ and $\hat{\beta}_1\left(\rho_s^j(\tau), \tau\right)$ are the IVQR estimate of $\alpha_0(\tau)$, $\beta_0(\tau)$ and $\beta_1(\tau)$.

The IVQR approach[17] allows interpretation of the $\widehat{\rho_s}(\tau)^{IVQR}$ 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 individual 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 methods put no restriction of the form of the endogenous variables and instruments.

For these estimations, a set of background variables of family (Cannari and D'Alessio, 1995; Card, 1999; Trostel et al., 2002) will be used as instruments for the implementation of the IVQR methods. The idea is based on the observation of persistence across generation about the level of schooling and it is theoretically justified by involuntary transmission of human capital. In particular, our instruments will be a set of variables that measure family background: the highest completed educational level by the father and the mother of the interviewed individual. Therefore, more educated parents are likely to value education more and to fill better jobs.

[17] We use the Stata command ivqreg, that performs an instrumental variable quantile regression using robust standard error formula in Chernozhukov and Hansen (2006, 2008) to evaluate heterogeneous marginal effect of endogenous variable.

3.5 Estimated Results

3.5.1 Returns to Education across the Wage Distribution

The results[18] of estimation of Equation (3.2) by applying IVQR methods are shown in Table 3.2. Finding show a large degree of heterogeneity in returns to education. The estimated coefficients are larger in the upper tail of the wages distribution than in the lower tail. This means that the highest wage earners enjoy larger gains from having an additional year of education.

Table 3.2 – Returns to education, IVQR estimates by quantile

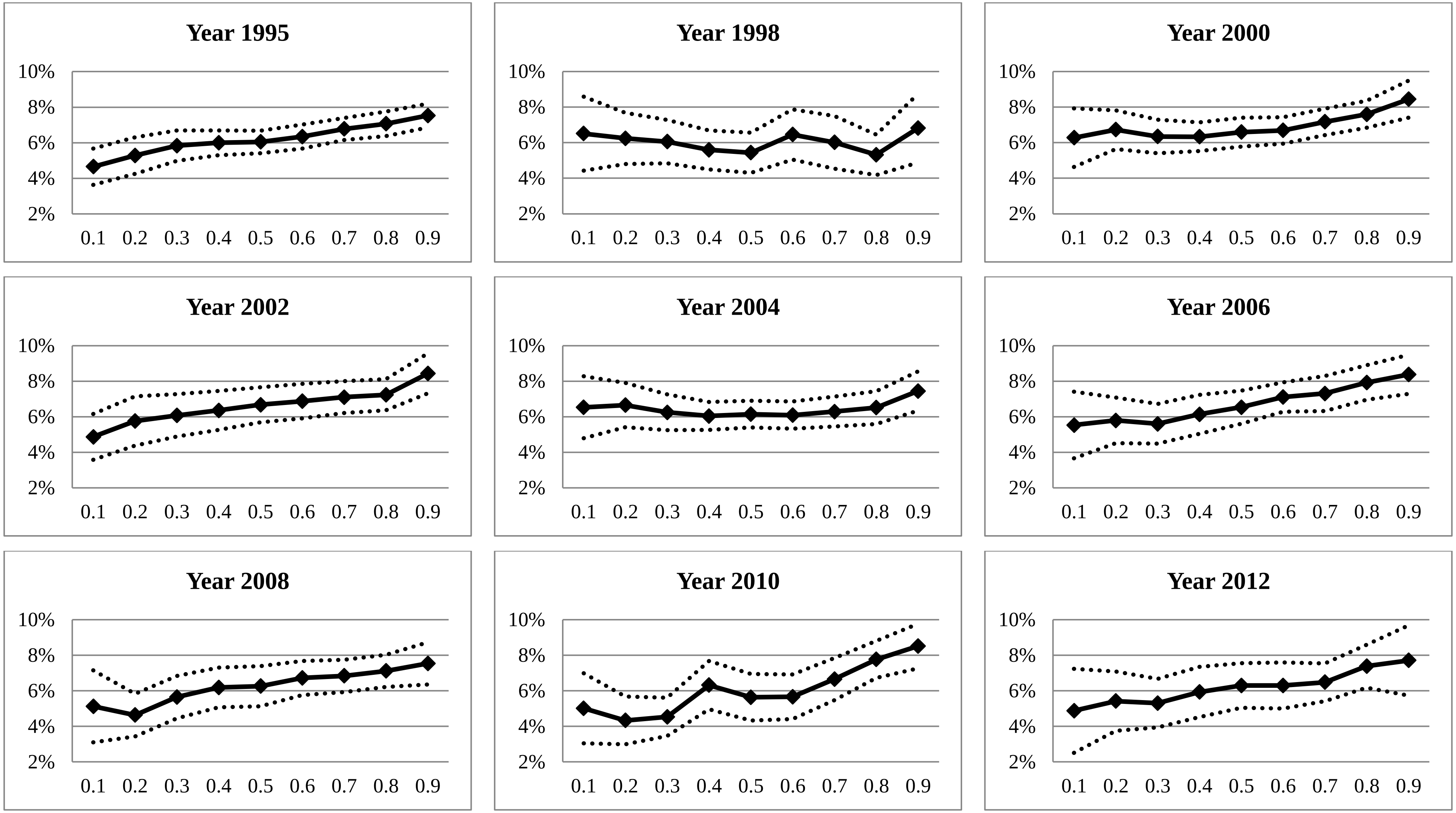
Year	τ=0.1	τ=0.2	τ=0.3	τ=0.4	τ=0.5	τ=0.6	τ=0.7	τ=0.8	τ=0.9
1995	0.047***	0.053***	0.058***	0.060***	0.060***	0.063***	0.068***	0.071***	0.075***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
1998	0.065***	0.062***	0.061***	0.056***	0.054***	0.065***	0.060***	0.053***	0.068***
	(0.011)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.008)	(0.006)	(0.010)
2000	0.063***	0.067***	0.063***	0.063***	0.066***	0.067***	0.072***	0.076***	0.084***
	(0.008)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
2002	0.049***	0.058***	0.061***	0.064***	0.067***	0.069***	0.071***	0.072***	0.084***
	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)
2004	0.065***	0.067***	0.062***	0.060***	0.061***	0.061***	0.063***	0.065***	0.074***
	(0.009)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
2006	0.055***	0.058***	0.056***	0.061***	0.065***	0.071***	0.073***	0.079***	0.084***
	(0.010)	(0.007)	(0.006)	(0.006)	(0.005)	(0.004)	(0.005)	(0.005)	(0.006)
2008	0.051***	0.046***	0.056***	0.062***	0.063***	0.067***	0.068***	0.071***	0.075***
	(0.010)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)
2010	0.050***	0.043***	0.045***	0.063***	0.056***	0.057***	0.067***	0.078***	0.085***
	(0.010)	(0.007)	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.006)
2012	0.049***	0.054***	0.053***	0.059***	0.063***	0.063***	0.065***	0.074***	0.077***
	(0.012)	(0.009)	(0.007)	(0.007)	(0.006)	(0.007)	(0.005)	(0.006)	(0.010)

*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in brackets

Moreover, for year 1995 and year 2002 the rates of return appear to monotonically increase as the quantile increases. In particular, for 1995, the estimated return to education at the 10^{th} percentile is equal to 4.7 percent and increase to 7.5 percent at the 90^{th} percentile. For 2002, the increase is even more pronounced, from 4.9 percent to 8.4 percent for the same percentiles.

[18] The results are based on the log of net hourly wages. Progressive taxation is likely to have a stronger impact in eroding the returns to education at the top of the distribution than at its bottom (Martins and Pereira, 2004).

Figure 3.1 – Estimated return to schooling over the wage distribution, from 1995 to 2012, by quantile, with 95% confidence intervals

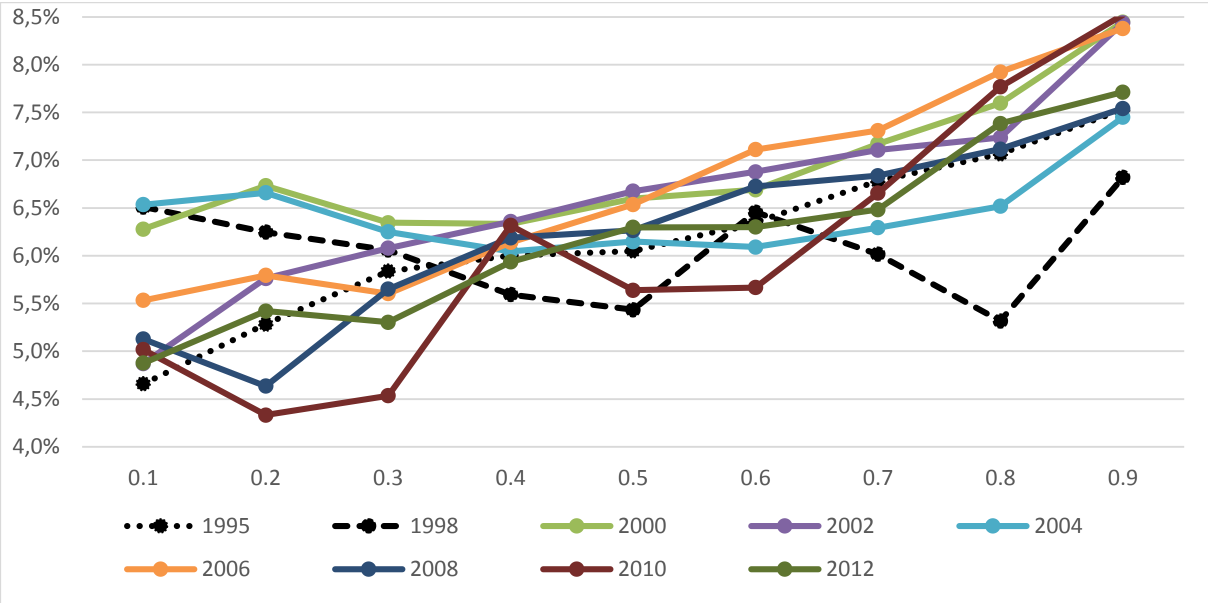


However, for 1998 IVQR results do not show a clear pattern and for 2004 returns to education first decrease until 40th percentile and then increase. Overall, for all the other year of the sample the results still show an increasing pattern. These results underline that average effects are not fully informative about the distribution.

To further underline the heterogeneous effects of education on wage, Figure 3.1 provides a graphical illustration of these results reporting the quantile-specific returns to education from $\tau = 0.1$ to $\tau=0.9$ and for each year of the sample period (from 1995 to 2012) with the relative 95 percent pointwise confidence interval.

Then, we examine how the impact of education on wage levels and wage dispersion has evolved from 1995 to 2012. To describe changes in the conditional wage distribution, Figure 3.2 plots all the quantile-return profiles at different years of the sample period.

Figure 3.2 – Estimated return to schooling over the wage distribution at different years, by quantile



Comparing the IVQR results for return to schooling for each year, we do not find any clear pattern, either increasing or decreasing for all the parts of the wage distribution, moving from 1995 to 2012. In particular, in the lower part of the wage distribution the return to schooling show a decrease from 1995 to 2010 and then a recovery in 2012. In the middle part of the distribution of wages, the returns to education are around 6 percent (except for 1998) and then in the upper part of the distribution we observe an increase until 2010 and then a slightly decrease in 2012.

3.5.2 Experience across the Wage Distribution

We find heterogeneity also in the estimate of the return of experience variable. The estimated coefficients are larger in the lower tail of the wages distribution than in the upper tail. This means that the least wage earners enjoy larger gains from having an additional year of experience than do the highest wage earners.

The rates of return appear to monotonically decrease as the quantile increases. For all the year of the sample, the results show a decreasing pattern. In Figure 3.3, we calculated the estimate return to experience for a worker with 20 years-experience. In particular, for 1995, the estimated return to education at the 10^{th} percentile is equal to 35.3 percent and decrease to 21.5 percent at the 90^{th} percentile. For 2012, the increase is even more pronounced, from 48 percent to 22.3 percent for the same percentiles.

Figure 3.3 – Estimated return to experience(=20) over the wage distribution, from 1995 to 2012, by quantile

(not reported standard errors show that differences between 0.1 and 0.9 quantiles are significant)

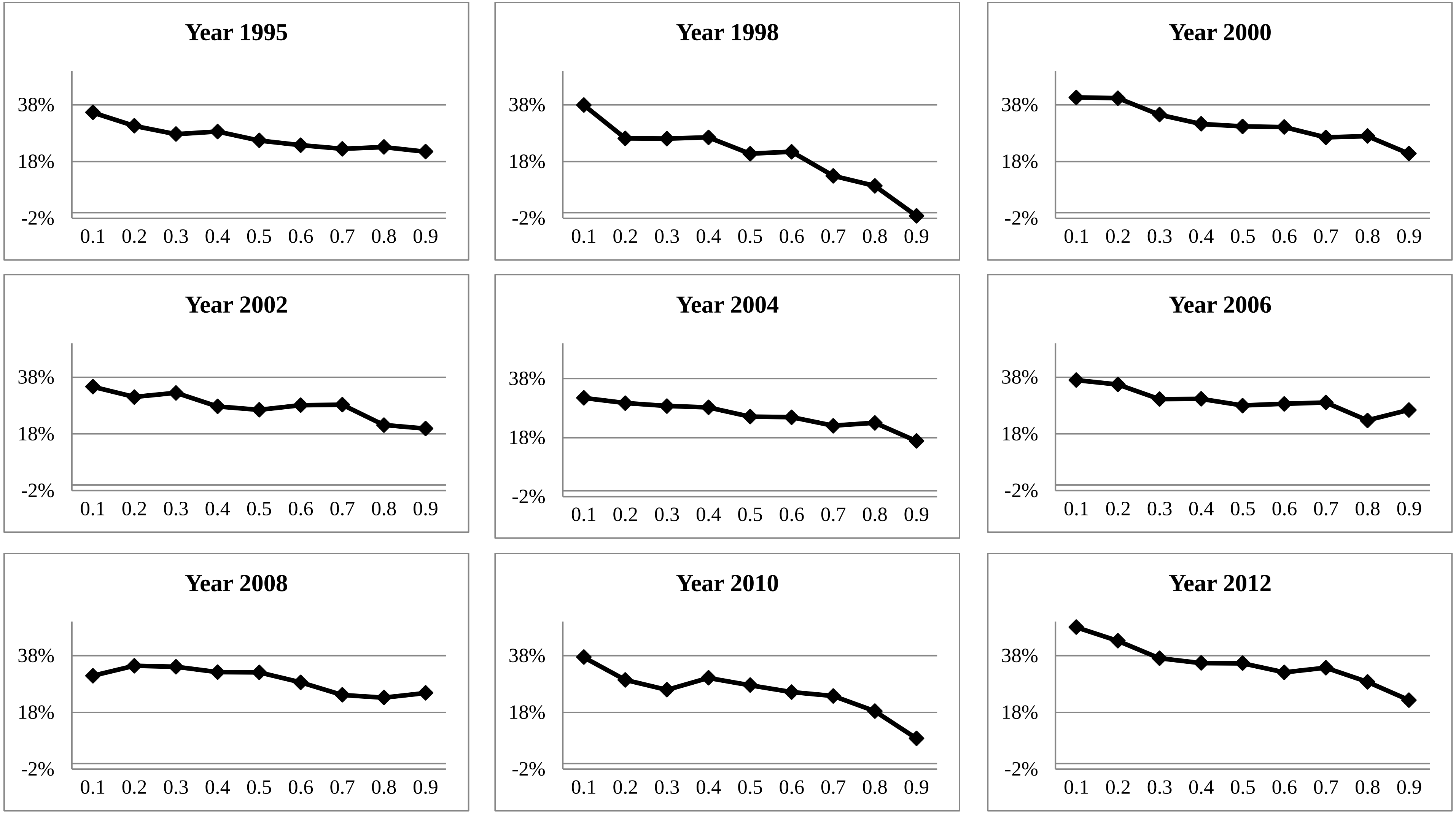
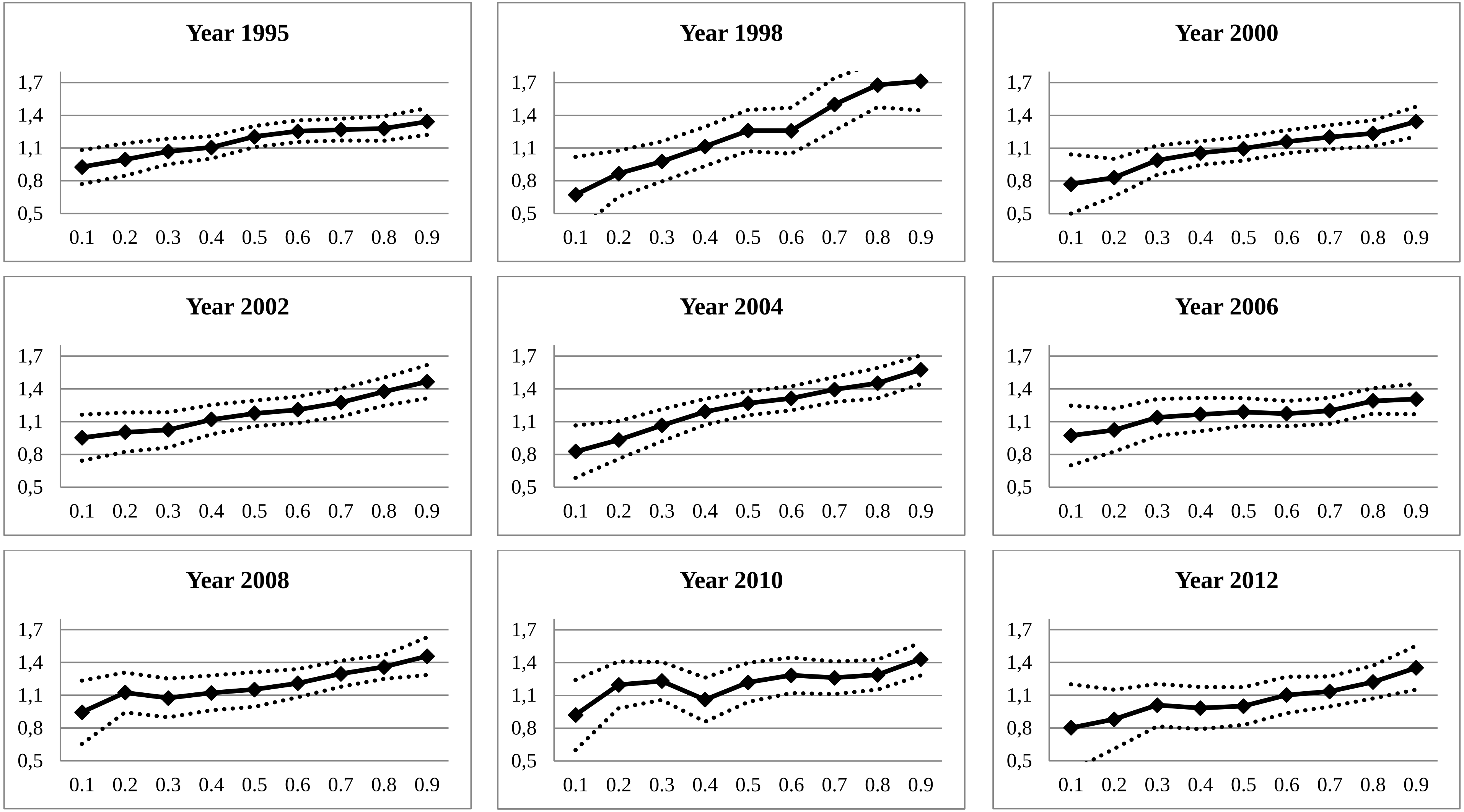


Figure 3.4 – Estimated specific intercept over the wage distribution, from 1995 to 2012, by quantile, with 95% confidence intervals



3.5.3 Quantile Specific Intercept across the Wage Distribution

The econometric model we apply in the empirical analysis allows considering the heterogeneity also through quantile-specific intercepts. In Figure 3.4, the estimated specific intercepts over the wage distribution are plotted for all the year selected in the analysis. The intercept shows an increasing pattern in quantile for almost all the year.

Our results suggest that the impact of education on the distribution of wages depends on the initial distribution of ability across population and, consequently, formal education does not compensate for differences in innate abilities and early life conditions.

3.5.4 Control Variables across the Wage Distribution

Results about control variables are in line with previous literature on these topics. Difference reflects mainly geographical and sectorial performance in Italy. In general, the results show a large degree of heterogeneity also across gender. In particular, gains are uniformly larger for men than for women across the whole distribution (around 9-17 percent) and for all the year that we consider in the analysis.

Considering the geographical residence of the workers and the sector of employment, differences in estimates mainly reflect territorial and sectorial performance of Italy. It is more convenient to work in the north regions in comparison to the center regions, and the returns decrease with quantile. Instead, if an individual works in the south region he will earns less than in the center regions.

Considering different sector of employment, working in the agricultural sector is less convenient than working in the industrial sector for almost all the quantile across the wage quantile distribution (but for year 1995 and year 2000, we find significant positive effect for working in agricultural sector at high quantile).

Finally, we find a wage gap in favor of public sector employees.

3.6 Concluding Remarks

In this chapter, we present evidence of heterogeneous returns to education over the wage distribution. We estimate causal link between education and wages at different quantiles of the conditional distribution of wages. The results provide evidence that there is not unique causal effect of schooling and that for each individual the effect depends on his position in the wage distribution and his unobservable wage determinants, such as ability.

The IVQR estimates show that returns to schooling vary substantially over the wage distribution. This means that returns to education are heterogeneous and the shape of the estimated returns over the quantiles are different and also different for each year of the sample period. In particular, taking into account the endogeneity of schooling, we observe that returns to education show an increasing pattern in the quantile index. If we interpret the quantile index as a measure of unobserved individual ability (Chernozhukov, Hasen and Janson, 2007), our results suggest that more able individuals profit more from one additional year of education. This means that more able individuals acquire more schooling because of higher marginal benefits.

According to our results, education should have an inequality-increasing effect over time, because individuals with high ability, those at the upper quantile of the wage distribution, seem to profit more from formal education. Therefore, considering also the endogeneity of the schooling variable, we confirm for Italy a previous result in the literature for other developed countries. Our results are in contrast with the ordinary quantile estimates of return to education in Martins and

Pereira (2014) for Italy that do not show any increasing or decreasing pattern across the quantile.

Finally, this analysis shows that estimates of the average return to education do not provide a complete characterization of the impact of education on labor market outcomes. Instrumental quantile regression methods is an appropriate tool to describe the impact of education on wages distribution, controlling for unobserved heterogeneity. Our results suggest that the impact of education on the distribution of wages depends on the initial distribution of ability across population and, as a consequence, formal education does not compensate for differences in innate abilities and early life conditions. Our findings contributes to a growing literature estimating heterogeneous effects of education. By illustrating an application of the IVQR approach in Italy our results highlights the importance in looking beyond the average causal effects of the variables of main interest in empirical analysis of the estimation of the return to education.

Appendix A to Chapter 1

A.1 Some Descriptive Statistics of the Sample

Figure A.1 – Mean of the Log of hourly wages less tax (1995 -2012)

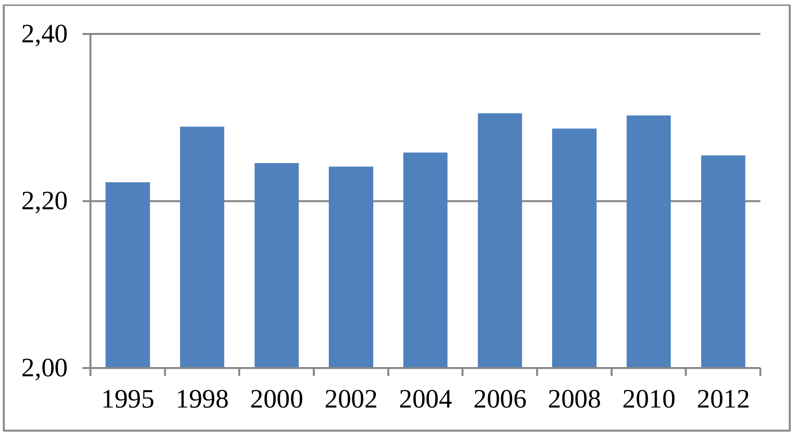


Figure A.2 – Mean of the number of year of Schooling (1995 -2012)

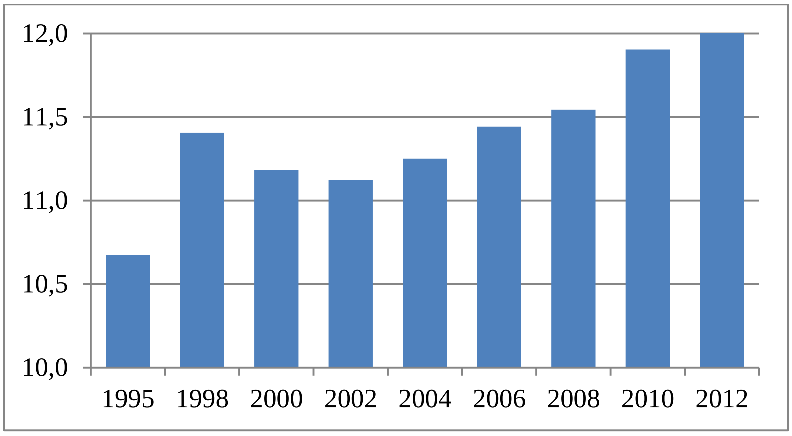
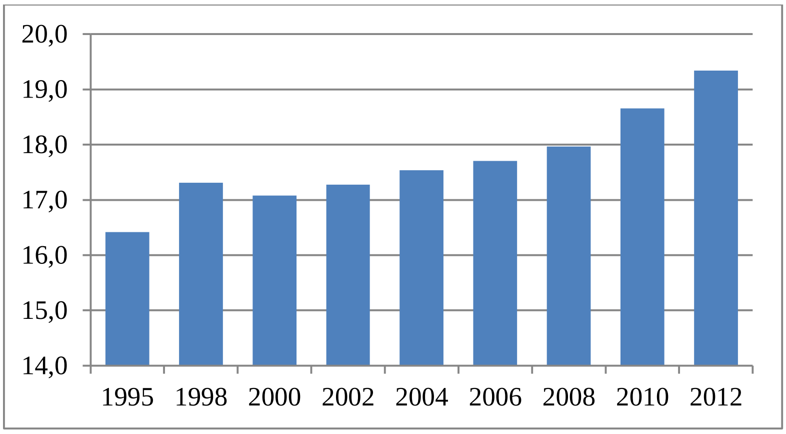


Figure A.3 – Mean of the number of year of Experience (1995 -2012)



A.2 OLS Estimates

Table A.1 shows OLS estimates, obtained by including in the original specification controls for the composition of her/his family, the geographical area of residence and the sector in which the individual is currently working.

Table A.1 - OLS estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
SCHOOL	0.0514***	0.0447***	0.0425***	0.0454***	0.0409***	0.0451***	0.0441***	0.0416***	0.0377***
	(0.00173)	(0.00315)	(0.00188)	(0.00229)	(0.00224)	(0.00196)	(0.00229)	(0.00214)	(0.00214)
EXPERIENCE	0.0272***	0.0275***	0.0255***	0.0271***	0.0210***	0.0250***	0.0274***	0.0194***	0.0221***
	(0.00277)	(0.00458)	(0.00246)	(0.00285)	(0.00300)	(0.00268)	(0.00275)	(0.00249)	(0.00285)
EXPERIENCE^2	-0.000352***	-0.000351***	-0.000365***	-0.000362***	-0.000308***	-0.000386***	-0.000405***	-0.000207***	-0.000226***
	(6.82e-05)	(0.000114)	(6.05e-05)	(7.46e-05)	(8.14e-05)	(6.75e-05)	(7.15e-05)	(6.20e-05)	(6.73e-05)
DUMMY_MALE	0.0855***	0.0422	0.0785***	0.106***	0.0790***	0.0905***	0.112***	0.117***	0.0654***
	(0.0155)	(0.0282)	(0.0137)	(0.0163)	(0.0174)	(0.0159)	(0.0143)	(0.0143)	(0.0165)
DUMMY_MARRIED	0.0739***	0.0666**	0.105***	0.0617***	0.0702***	0.0536***	0.0312**	0.0684***	0.0476***
	(0.0165)	(0.0331)	(0.0148)	(0.0191)	(0.0180)	(0.0161)	(0.0157)	(0.0153)	(0.0175)
NCOMP	-0.00338	0.00475	-0.0104*	-0.00937	-0.0122*	0.0140**	0.00783	-0.00193	0.0179**
	(0.00535)	(0.0105)	(0.00550)	(0.00653)	(0.00665)	(0.00654)	(0.00573)	(0.00609)	(0.00705)
DUMMY_HOUSEHOLD	0.0436***	0.0463	0.0325**	0.0381**	0.0385**	0.0651***	0.0503***	0.0168	0.0268*
	(0.0167)	(0.0285)	(0.0134)	(0.0159)	(0.0165)	(0.0147)	(0.0135)	(0.0128)	(0.0144)
DUMMY_TOWN	0.0333*	0.0209	0.0457**	-0.0369	0.0233	0.0587***	0.0280	-0.0120	0.0152
	(0.0175)	(0.0340)	(0.0178)	(0.0278)	(0.0313)	(0.0215)	(0.0243)	(0.0240)	(0.0296)
DUMMY_NORTH	0.0404***	0.0778***	0.0479***	0.0398**	0.0473**	-0.00852	-0.0296*	0.0441***	0.0249
	(0.0140)	(0.0241)	(0.0136)	(0.0169)	(0.0198)	(0.0160)	(0.0164)	(0.0170)	(0.0179)
DUMMY_SOUTH	-0.0379**	0.0570**	-0.0293	0.00369	-0.0233	-0.0821***	-0.0783***	-3.96e-05	-0.0216
	(0.0172)	(0.0288)	(0.0189)	(0.0232)	(0.0241)	(0.0188)	(0.0186)	(0.0189)	(0.0210)
DUMMY_AGRICULTURAL	-0.117*	-0.0967	-0.131***	-0.0424	-0.0935***	-0.168***	-0.00792	-0.0596	-0.0701
	(0.0679)	(0.0705)	(0.0437)	(0.0568)	(0.0329)	(0.0480)	(0.0419)	(0.0388)	(0.0518)
DUMMY_PUBLIC	0.174***	0.109***	0.108***	0.110***	0.141***	0.126***	0.143***	0.148***	0.128***
	(0.0168)	(0.0268)	(0.0150)	(0.0182)	(0.0197)	(0.0188)	(0.0190)	(0.0176)	(0.0212)
DUMMY_OTHER_SECTOR	0.0109	-0.00144	0.0288*	0.0103	0.00798	0.000915	-0.00276	0.0159	-0.0460***
	(0.0149)	(0.0301)	(0.0152)	(0.0177)	(0.0189)	(0.0161)	(0.0160)	(0.0146)	(0.0169)
DUMMY_SECT_PARENTS	0.00296	0.0797***	0.0176	0.0139	0.00659	-0.00164	0.0340***	0.0113	0.00327
	(0.0151)	(0.0228)	(0.0121)	(0.0157)	(0.0146)	(0.0135)	(0.0131)	(0.0133)	(0.0153)
DUMMY_PART_TIME	0.0734**	0.0348	0.0475*	-0.0834**	-0.0480	-0.00720	0.0260	-0.00241	-0.0312
	(0.0324)	(0.0527)	(0.0267)	(0.0367)	(0.0310)	(0.0301)	(0.0268)	(0.0215)	(0.0230)
Constant	1.173***	1.192***	1.309***	1.280***	1.437***	1.342***	1.307***	1.360***	1.340***
	(0.0408)	(0.0708)	(0.0354)	(0.0420)	(0.0460)	(0.0438)	(0.0437)	(0.0459)	(0.0462)
Observations	6,066	2,016	5,724	5,461	5,425	5,378	5,409	5,161	4,975
R-squared	0.450	0.366	0.353	0.306	0.261	0.326	0.353	0.327	0.314

Robust standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Table A.2 shows OLS estimates of the empirical specification, including interaction of the variable schooling with experience and with gender.

Table A.2 - OLS estimates with interactions. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
SCHOOL	0.0471***	0.0445***	0.0291***	0.0388***	0.0420***	0.0267***	0.0242***	0.0330***	0.0251***
	(0.00663)	(0.0108)	(0.00625)	(0.00831)	(0.00657)	(0.00555)	(0.00553)	(0.00539)	(0.00620)
EXPERIENCE	0.0216***	0.0455***	0.00857	0.0158*	0.0199***	0.00515	-0.000317	0.00709	0.0203**
	(0.00810)	(0.0138)	(0.00760)	(0.00923)	(0.00772)	(0.00846)	(0.00880)	(0.00714)	(0.00837)
EXPERIENCE^2	-0.000294	-0.000970***	-6.86e-05	-0.000172	-0.000413**	-0.000174	0.000133	-0.000119	-0.000474**
	(0.000196)	(0.000344)	(0.000183)	(0.000229)	(0.000199)	(0.000219)	(0.000228)	(0.000177)	(0.000193)
SCHOOL*EXPER	0.000492	-0.00170	0.00138**	0.000930	-5.80e-07	0.00155**	0.00230***	0.000961	3.08e-05
	(0.000704)	(0.00112)	(0.000640)	(0.000824)	(0.000702)	(0.000678)	(0.000700)	(0.000590)	(0.000696)
SCHOOL*EXPER^2	-4.17e-06	5.86e-05*	-2.29e-05	-1.48e-05	1.22e-05	-1.33e-05	-4.39e-05**	-5.10e-06	2.45e-05
	(1.78e-05)	(3.04e-05)	(1.60e-05)	(2.13e-05)	(1.85e-05)	(1.81e-05)	(1.82e-05)	(1.51e-05)	(1.64e-05)
DUMMY_MALE	0.129***	-0.0536	0.102**	0.173***	0.200***	0.146***	0.136***	0.236***	0.0299
	(0.0432)	(0.0821)	(0.0435)	(0.0568)	(0.0505)	(0.0439)	(0.0483)	(0.0474)	(0.0534)
SCHOOL*MALE	-0.00403	0.00848	-0.00222	-0.00589	-0.0106**	-0.00499	-0.00203	-0.00996**	0.00302
	(0.00360)	(0.00645)	(0.00365)	(0.00482)	(0.00426)	(0.00361)	(0.00418)	(0.00395)	(0.00416)
DUMMY_MARRIED	0.0732***	0.0700**	0.105***	0.0611***	0.0703***	0.0572***	0.0286*	0.0658***	0.0476***
	(0.0166)	(0.0323)	(0.0148)	(0.0190)	(0.0182)	(0.0161)	(0.0156)	(0.0153)	(0.0175)
NCOMP	-0.00398	0.00307	-0.0114**	-0.00998	-0.0130*	0.0133**	0.00836	-0.00200	0.0164**
	(0.00537)	(0.0105)	(0.00546)	(0.00646)	(0.00664)	(0.00653)	(0.00571)	(0.00606)	(0.00701)
DUMMY_HOUSEHOLD	0.0441***	0.0438	0.0332**	0.0392**	0.0391**	0.0624***	0.0491***	0.0172	0.0245*
	(0.0168)	(0.0281)	(0.0134)	(0.0160)	(0.0165)	(0.0146)	(0.0135)	(0.0127)	(0.0144)
DUMMY_TOWN	0.0317*	0.0142	0.0464***	-0.0388	0.0216	0.0585***	0.0265	-0.0177	0.0119
	(0.0175)	(0.0342)	(0.0178)	(0.0278)	(0.0312)	(0.0214)	(0.0241)	(0.0239)	(0.0296)
DUMMY_NORTH	0.0399***	0.0745***	0.0467***	0.0402**	0.0481**	-0.0101	-0.0320*	0.0436***	0.0245
	(0.0140)	(0.0238)	(0.0136)	(0.0169)	(0.0198)	(0.0159)	(0.0164)	(0.0168)	(0.0179)
DUMMY_SOUTH	-0.0416**	0.0563*	-0.0324*	0.00143	-0.0274	-0.0864***	-0.0825***	-0.00393	-0.0225
	(0.0172)	(0.0287)	(0.0189)	(0.0232)	(0.0237)	(0.0188)	(0.0185)	(0.0186)	(0.0210)
DUMMY_AGRICULTURAL	-0.116*	-0.108	-0.128***	-0.0434	-0.0942***	-0.173***	-0.00883	-0.0568	-0.0747
	(0.0686)	(0.0679)	(0.0440)	(0.0567)	(0.0331)	(0.0482)	(0.0421)	(0.0384)	(0.0503)
DUMMY_PUBLIC	0.173***	0.111***	0.104***	0.106***	0.138***	0.120***	0.140***	0.141***	0.121***
	(0.0169)	(0.0268)	(0.0150)	(0.0181)	(0.0197)	(0.0187)	(0.0191)	(0.0174)	(0.0213)
DUMMY_OTHER_SECTOR	0.00895	-0.00335	0.0275*	0.0107	0.0108	-3.01e-05	-0.00259	0.0144	-0.0480***
	(0.0151)	(0.0301)	(0.0150)	(0.0179)	(0.0190)	(0.0161)	(0.0158)	(0.0145)	(0.0170)
DUMMY_SECT_PARENTS	0.00368	0.0778***	0.0179	0.0146	0.00648	0.00102	0.0354***	0.0122	0.00324
	(0.0151)	(0.0225)	(0.0121)	(0.0157)	(0.0146)	(0.0135)	(0.0132)	(0.0131)	(0.0152)
DUMMY_PART_TIME	0.0763**	0.0280	0.0505*	-0.0794**	-0.0440	-0.00747	0.0251	-0.000550	-0.0349
	(0.0325)	(0.0519)	(0.0264)	(0.0369)	(0.0312)	(0.0300)	(0.0268)	(0.0215)	(0.0228)
Constant	1.225***	1.209***	1.478***	1.362***	1.432***	1.575***	1.553***	1.472***	1.513***
	(0.0821)	(0.144)	(0.0764)	(0.0980)	(0.0819)	(0.0812)	(0.0766)	(0.0773)	(0.0793)
Observations	6,066	2,016	5,724	5,461	5,425	5,378	5,409	5,161	4,975
R-squared	0.451	0.371	0.356	0.308	0.264	0.334	0.358	0.333	0.324

Robust standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

A.3 First Stage Regression of IV Estimates

Table A.3 shows the estimates of the first stage regression of the instrumental variables estimation.

Table A.3 – First stage of IV estimates. Dependent Variable: schooling. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
SCHOOL_F	0.296***	0.292***	0.298***	0.267***	0.237***	0.261***	0.260***	0.281***	0.248***
	(0.0213)	(0.0393)	(0.0248)	(0.0253)	(0.0240)	(0.0247)	(0.0289)	(0.0431)	(0.0348)
SCHOOL_M	0.216***	0.207***	0.150***	0.172***	0.189***	0.135***	0.183***	0.145***	0.159***
	(0.0254)	(0.0513)	(0.0284)	(0.0290)	(0.0261)	(0.0285)	(0.0332)	(0.0433)	(0.0371)
EXPERIENCE	0.0363	-0.0374	-0.0130	-0.00816	-0.0495	-0.0453	-0.0308	-0.0588	-0.00311
	(0.0268)	(0.0492)	(0.0305)	(0.0350)	(0.0302)	(0.0297)	(0.0311)	(0.0449)	(0.0419)
EXPERIENCE^2	-0.00226***	0.000500	-0.000833	-0.000699	0.000762	-0.000104	-0.000199	0.000484	-0.000570
	(0.000662)	(0.00118)	(0.000740)	(0.000844)	(0.000727)	(0.000696)	(0.000718)	(0.00101)	(0.000910)
DUMMY_MALE	0.269	0.409	-0.317*	-0.411**	-0.395**	-0.310	-0.333	-0.592***	-0.412**
	(0.195)	(0.340)	(0.172)	(0.190)	(0.174)	(0.193)	(0.202)	(0.219)	(0.209)
DUMMY_MARRIED	0.510**	-0.428	-0.0373	0.294	0.618***	0.418**	0.792***	-0.0912	0.0893
	(0.246)	(0.419)	(0.243)	(0.251)	(0.221)	(0.209)	(0.251)	(0.250)	(0.250)
NCOMP	-0.169***	-0.196*	-0.0204	0.0844	-0.0318	-0.0220	-0.170**	0.226**	0.0448
	(0.0643)	(0.114)	(0.0755)	(0.0777)	(0.0754)	(0.0732)	(0.0784)	(0.0980)	(0.0935)
DUMMY_HOUSEHOLD	-0.168	-0.629*	0.188	-0.0163	0.109	0.177	0.0701		
	(0.199)	(0.344)	(0.165)	(0.181)	(0.171)	(0.180)	(0.218)		
DUMMY_TOWN	0.223	-0.179	0.908***	0.531**	0.110	0.398*	0.348	0.435	0.717**
	(0.189)	(0.354)	(0.209)	(0.247)	(0.216)	(0.225)	(0.273)	(0.292)	(0.323)
DUMMY_NORTH	-0.158	0.0635	0.238	0.378**	0.0239	-0.235	0.173	-0.252	-0.453*
	(0.159)	(0.268)	(0.168)	(0.181)	(0.172)	(0.179)	(0.214)	(0.247)	(0.248)
DUMMY_SOUTH	-0.197	0.0832	0.0470	0.310	-0.315	-0.622***	-0.240	-0.281	-0.340
	(0.177)	(0.293)	(0.190)	(0.225)	(0.222)	(0.212)	(0.240)	(0.304)	(0.291)
DUMMY_AGRICULTURAL	-1.372***	-2.769***	-1.473***	-1.443***	-1.327***	-0.756**	-1.060***	-1.098*	-0.958**
	(0.466)	(0.562)	(0.373)	(0.332)	(0.335)	(0.312)	(0.397)	(0.636)	(0.382)
DUMMY_PUBLIC	2.465***	2.268***	2.551***	2.454***	2.520***	2.491***	2.423***	2.235***	2.529***
	(0.177)	(0.285)	(0.175)	(0.205)	(0.192)	(0.175)	(0.201)	(0.294)	(0.274)
DUMMY_OTHER_SECTOR	0.00794	0.473	0.800***	0.501***	0.866***	0.715***	0.645***	0.643***	0.521**
	(0.159)	(0.300)	(0.172)	(0.176)	(0.180)	(0.176)	(0.192)	(0.229)	(0.229)
DUMMY_SECT_PARENTS	0.154	0.273	0.291**	-0.160	0.183	0.393***	0.193	-0.0290	0.255
	(0.151)	(0.218)	(0.135)	(0.151)	(0.139)	(0.144)	(0.163)	(0.204)	(0.187)
DUMMY_PART_TIME	-0.629***	-0.656*	-0.594**	-0.528**	-0.825***	-0.857***	-0.599**	-0.822**	-0.651**
	(0.228)	(0.378)	(0.232)	(0.262)	(0.248)	(0.237)	(0.295)	(0.411)	(0.265)
Constant	7.434***	8.901***	7.903***	7.570***	7.969***	8.906***	8.516***	9.259***	8.996***
	(0.385)	(0.732)	(0.399)	(0.467)	(0.412)	(0.423)	(0.461)	(0.599)	(0.548)
Observations	4,352	1,468	3,783	3,321	3,405	3,437	2,836	2,145	2,112
R-squared	0.408	0.408	0.390	0.373	0.365	0.365	0.364	0.322	0.338
Sargan test $\chi 2(1)$	1.691	1.891	0.239	0.05	0.515	0.197	0.026	0.457	0.868
p-Value	0.1935	0.1691	0.6248	0.8239	0.473	0.657	0.8716	0.5038	0.3516
F-test on excl. instrum.	461.28	147.101	288.83	201.88	225.05	195.76	188.55	107,67	106.27
p-Value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0000

Robust standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

A.4 First Step in the Ordered Probit

Table A.4 reports the results of the ordered probit model for educational attainment as a function of the instrument used in the IV estimation. This is the first step necessary to estimate the score associated to the ordered probit that we add in the wages equation in order to apply ordinary least squares as second step.

Table A.4 – Ordered probit estimates. Dependent Variable: education. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL).

VARIABLES	1995	1998	2000	2002	2004	2006	2008	2010	2012
PRIMARY_F	-0.836***	-0.950***	-0.716***	-0.586***	-0.628***	-0.677***	-0.725***	-0.564***	-0.523***
	(0.0653)	(0.109)	(0.0762)	(0.0774)	(0.0700)	(0.0735)	(0.0807)	(0.0912)	(0.0819)
PRIMARY_M	-0.495***	-0.375***	-0.421***	-0.553***	-0.436***	-0.293***	-0.453***	-0.280***	-0.468***
	(0.0729)	(0.120)	(0.0834)	(0.0838)	(0.0784)	(0.0775)	(0.0848)	(0.0989)	(0.0847)
EXPERIENCE	0.00682	-0.0205	-0.0171	-0.0141	-0.0386***	-0.0194	-0.0130	-0.0178	0.00268
	(0.0102)	(0.0189)	(0.0111)	(0.0133)	(0.0121)	(0.0125)	(0.0127)	(0.0157)	(0.0165)
EXPERIENCE^2	-0.000571**	0.000449	0.000122	8.38e-05	0.000782***	3.84e-05	3.19e-05	0.000118	-0.000270
	(0.000254)	(0.000449)	(0.000272)	(0.000319)	(0.000283)	(0.000301)	(0.000294)	(0.000364)	(0.000358)
DUMMY_MALE	0.0436	0.113	-0.177***	-0.147**	-0.175**	-0.115	-0.137*	-0.240***	-0.179**
	(0.0738)	(0.140)	(0.0648)	(0.0718)	(0.0687)	(0.0810)	(0.0766)	(0.0805)	(0.0806)
DUMMY_MARRIED	0.194**	-0.188	0.0600	0.112	0.292***	0.130	0.343***	-0.0254	0.0954
	(0.0983)	(0.159)	(0.0881)	(0.0985)	(0.0896)	(0.0869)	(0.0979)	(0.0946)	(0.0977)
NCOMP	-0.0637**	-0.117***	-0.00877	0.0137	-0.0325	-0.0213	-0.0470	0.0634*	0.00537
	(0.0247)	(0.0443)	(0.0263)	(0.0304)	(0.0302)	(0.0308)	(0.0312)	(0.0364)	(0.0354)
DUMMY_HOUSEHOLD	-0.0893	-0.320**	0.0640	-0.0679	0.0225	0.0290	0.0406		
	(0.0749)	(0.140)	(0.0624)	(0.0698)	(0.0681)	(0.0792)	(0.0856)		
DUMMY_TOWN	0.113	0.0415	0.297***	0.194**	0.0734	0.220**	0.149	0.190*	0.263**
	(0.0703)	(0.126)	(0.0784)	(0.0957)	(0.0901)	(0.0885)	(0.102)	(0.106)	(0.116)
DUMMY_NORTH	-0.0315	-0.0375	0.142**	0.148**	0.0360	-0.130*	0.0132	-0.103	-0.183**
	(0.0614)	(0.100)	(0.0643)	(0.0729)	(0.0715)	(0.0724)	(0.0830)	(0.0889)	(0.0888)
DUMMY_SOUTH	-0.0787	0.0318	0.0541	0.0981	-0.135	-0.292***	-0.196**	-0.192*	-0.170
	(0.0675)	(0.106)	(0.0705)	(0.0876)	(0.0868)	(0.0877)	(0.0941)	(0.106)	(0.105)
DUMMY_AGRICULTURAL	-0.426**	-1.420***	-0.407***	-0.564***	-0.770***	-0.528***	-0.463**	-0.334	-0.591**
	(0.191)	(0.282)	(0.151)	(0.199)	(0.183)	(0.167)	(0.186)	(0.251)	(0.234)
DUMMY_PUBLIC	0.929***	0.749***	0.950***	0.977***	0.987***	0.991***	0.922***	0.892***	0.967***
	(0.0682)	(0.113)	(0.0646)	(0.0760)	(0.0760)	(0.0736)	(0.0790)	(0.104)	(0.106)
DUMMY_OTHER_SECT	0.0287	0.0758	0.297***	0.255***	0.342***	0.309***	0.285***	0.298***	0.236***
	(0.0651)	(0.116)	(0.0670)	(0.0720)	(0.0735)	(0.0709)	(0.0775)	(0.0869)	(0.0902)
DUMMY_SECT_PARENTS	0.0724	0.127	0.136***	-0.00329	0.117**	0.222***	0.110*	0.0442	0.110
	(0.0616)	(0.0843)	(0.0505)	(0.0592)	(0.0560)	(0.0585)	(0.0629)	(0.0734)	(0.0713)
DUMMY_PART_TIME	-0.264**	-0.333**	-0.169*	-0.151	-0.235**	-0.292***	-0.194*	-0.327**	-0.229**
	(0.104)	(0.157)	(0.0918)	(0.108)	(0.0968)	(0.0960)	(0.111)	(0.141)	(0.101)
Constant cut1	-1.009***	-1.807***	-0.830***	-0.796***	-0.917***	-1.076***	-0.985***	-1.079***	-1.020***
	(0.153)	(0.260)	(0.152)	(0.174)	(0.163)	(0.172)	(0.175)	(0.201)	(0.219)
Constant cut2	-0.771***	-1.549***	-0.519***	-0.520***	-0.644***	-0.757***	-0.664***	-0.705***	-0.640***
	(0.153)	(0.258)	(0.152)	(0.173)	(0.162)	(0.174)	(0.177)	(0.201)	(0.219)
Constant cut3	0.649***	-0.163	0.818***	0.858***	0.816***	0.702***	0.748***	0.575***	0.677***
	(0.153)	(0.255)	(0.152)	(0.175)	(0.165)	(0.178)	(0.182)	(0.199)	(0.215)
Observations	4,352	1,468	3,783	3,321	3,405	3,437	2,836	2,145	2,112

Robust standard errors in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Appendix B to Chapter 2

B.1 GAM Estimation

A Generalized Additive Model (Hastie and Tibshirani, 1986, 1990) is a generalized linear model with a linear predictor involving a sum of smooth functions of covariates. In general, the model has a structure something like:

$$g(\mu_i) = X_i^*\theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \cdots, \tag{B.1}$$

Where:

$$\mu_i \equiv E(Y_i) \text{ and } Y_i \sim \text{some exponential family distribution}$$

Y_i is a response variable, X_i^* is a row of the model matrix for any strictly parametric model components, θ is the corresponding parameter vector, and the f_j are smooth functions of the covariate, x_k.

The model allows for rather flexible specification of the dependence of the response on the covariates, but specifying the model only in terms of "smooth functions", rather than detailed parametric relationships, it is possible to avoid cumbersome and unwieldy models. This flexibility and convenience comes to the cost of two new theoretical problems. It is necessary both to represent the smooth functions is some way and to choose how smooth they should be.

We estimate the model (2.2) following the method described in Wood (2011), and implemented by the package *mgcv* in R. The estimation is obtained by penalized likelihood maximization. The model is fitted by minimizing:

$$\|\mathbf{y} - \mathbf{X}\beta\|^2 + \sum_{k=1}^{K} \gamma_k \int_{-\infty}^{+\infty} [\mu_k''(x)]^2 \, dx \tag{B.2}$$

where **y** is the vector of observations, **X** is the matrix of explanatory variables, β is a vector of parameters to be estimated, $\gamma_{k,} k = 1, \dots, K$ are smoothing parameters, and the penalty, which controls the smoothness of the estimate, is represented by the integrated square of the second derivatives of the smooth terms. The vector of parameters originates from expressing every smooth term in model (2.2), $\mu_k(\cdot)$, as:

$$\mu_k(x) = \sum_{l=1}^{q} b_l(x)\beta_l \tag{B.3}$$

where $b_l(x)$ are *basis functions* and q is their number.

Parameters β are chosen to minimize the function in Equation (B.2) for given values of the smoothing parameters γ_k. Smoothing parameters are, in turn, chosen by the minimization of the restricted maximum likelihood (REML) score. Estimation proceeds by penalized iteratively re-weighted least squares (P-IRLS), until convergence in the estimates is reached.

B.2 Bootstrap Procedure to Compute Confidence Intervals

Since the second-stage regression contains generated regressors (i.e. the first-stage residuals), to obtain the appropriate standard errors we use the following bootstrap procedure. Given a sample of observations (*y, X, Z*), where *y* is the vector of dimension *N* of dependent variable, *X* is the $N \times K$ matrix of explanatory variables (including the endogenous variable), and *Z* is the $N \times K$ matrix of instruments:

1. select a bootstrap sample (y_b^*, X_b^*, Z_b^*) drawn with replacement from (*y, X, Z*);
2. run a semiparametric regression of the endogenous variable on the exogenous variables and the instruments;
3. insert the first-stage residuals in the original semiparametric regression;
4. repeat B = 1000 times points 1-3;

5. for each estimated parametric coefficients compute the corresponding equal-tail bootstrap *p-value* (see Davidson and MacKinnon (2007)):

$$P^*(\hat{\beta}) = 2 * \min\left(\frac{1}{B}\sum_{b=1}^{B} \#\{\hat{\beta}_b^* \leq 0\}, \frac{1}{B}\sum_{b=1}^{B} \#\{\hat{\beta}_b^* > 0\}\right) \qquad \text{(B.4)}$$

6. for each estimated non-parametric coefficients compute the average partial effect and the 95 percent confidence bands.

B.3 Linear Model with Control Function Estimates

Table B.1 - First Stage LM estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL), 1995-2012.

	1995	1998	2000	2002	2004	2006	2008	2010	2012
Parametric coefficients:	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*
Constant	7,5023 ***	8,9688 ***	7,9623 ***	7,5281 ***	7,9987 ***	8,9869 ***	8,5361 ***	9,2506 ***	9,0686 ***
SCHOOL_F	0,2960 ***	0,2932 ***	0,3008 ***	0,2654 ***	0,2394 ***	0,2663 ***	0,2628 ***	0,2813 ***	0,2501 ***
SCHOOL_M	0,2176 ***	0,2144 ***	0,1537 ***	0,1693 ***	0,1918 ***	0,1385 ***	0,1853 ***	0,1445 ***	0,1637 ***
EXPERIENCE	0,0361 *	-0,0391	-0,0091	-0,0085	-0,0480 **	-0,0429 **	-0,0285	-0,0588 **	-0,0025
EXPERIENCE^2	-0,0023 ***	0,0005	-0,0009 *	-0,0007	0,0007	-0,0001	-0,0003	0,0005	-0,0006
DUMMY_MALE	0,2674 *	0,4095	-0,3024 **	-0,4119 ***	-0,3978 ***	-0,3277 **	-0,3265 **	-0,5919 ***	-0,4209 ***
DUMMY_MARRIED	0,5250 ***	-0,3986	-0,0250	0,2724	0,6338 ***	0,4694 ***	0,7943 ***	-0,0911	0,0843
NCOMP	-0,1688 ***	-0,1909 **	-0,0163	0,0819	-0,0294	-0,0208	-0,1652 ***	0,2258 ***	0,0486
DUMMY_HOUSEHOLD	-0,1701	-0,6443 **	0,1762	-0,0105	0,1133	0,1804	0,0546		
DUMMY_TOWN	0,2222	-0,1761	0,9159 ***	0,5292 ***	0,1141	0,4326 ***	0,3596 **	0,4355 **	0,7292 ***
DUMMY_NORTH	-0,1524	0,0574	0,2331 *	0,3815 ***	0,0240	-0,2544 *	0,1839	-0,2510	-0,4748 ***
DUMMY_SOUTH	-0,1973	0,0841	0,0459	0,3059 *	-0,3154 *	-0,6324 ***	-0,2380	-0,2783	-0,3545 *
DUMMY_AGRICULTURAL	-1,3261 ***	-2,7019 ***	-1,4081 ***	-1,4715 ***	-1,2805 ***	-0,6751 **	-1,0124 ***	-1,1045 ***	-0,9089 **
DUMMY_PUBLIC	2,3773 ***	2,2571 ***	2,5076 ***	2,4837 ***	2,4847 ***	2,4360 ***	2,4034 ***	2,2401 ***	2,4971 ***
DUMMY_OTHER_SECTOR	0,0105	0,5209 **	0,8101 ***	0,5107 ***	0,8757 ***	0,7378 ***	0,6541 ***	0,6453 ***	0,5114 ***
DUMMY_PART_TIME	-0,6396 ***	-0,6406 **	-0,6110 ***	-0,5200 ***	-0,8130 ***	-0,8624 ***	-0,5888 ***	-0,8230 ***	-0,6506 ***

(***p<0.01, **p<0.05, *p<0.1)

Table B.2 - Second Stage LM estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL), 1995-2012.

	1995	1998	2000	2002	2004	2006	2008	2010	2012
Parametric coefficients:	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*	*Estimate*
Constant	1,1280 ***	1,0736 ***	1,1140 ***	1,0880 ***	1,2435 ***	0,9841 ***	1,1610 ***	1,0820 ***	1,2130 ***
SCHOOL	0,0642 ***	0,0644 ***	0,0681 ***	0,0711 ***	0,0661 ***	0,0778 ***	0,0604 ***	0,0688 ***	0,0532 ***
EXPERIENCE	0,0189 ***	0,0186 ***	0,0207 ***	0,0246 ***	0,0143 ***	0,0249 ***	0,0230 ***	0,0151 ***	0,0169 ***
EXPERIENCE^2	-0,0001 ***	-0,0001	-0,0002 ***	-0,0003 ***	-0,0001 **	-0,0003 ***	-0,0003 ***	0,0000	-0,0001
DUMMY_MALE	0,1323 ***	0,1138 ***	0,0959 ***	0,0984 ***	0,0811 ***	0,1096 ***	0,1573 ***	0,1539 ***	0,1007 ***
DUMMY_MARRIED	0,0037	0,0553	0,0556 ***	0,0082	0,0363 ·	-0,0110	-0,0508 **	0,0292	0,0088
NCOMP	0,0176 ***	0,0161	-0,0015	0,0012	-0,0024	0,0314 ***	0,0287 ***	-0,0023	0,0209 ***
DUMMY_HOUSEHOLD	-0,0063	-0,0025	0,0065	0,0226	0,0186	0,0306 **	0,0064		
DUMMY_TOWN	0,0059	0,0318	0,0128	-0,0814 ***	-0,0185	0,0414 **	0,0176	-0,0341	-0,0038
DUMMY_NORTH	0,0376 ***	0,0663 **	0,0462 ***	0,0456 ***	0,0667 ***	-0,0078	-0,0005	0,0512 **	0,0413 **
DUMMY_SOUTH	-0,0239	0,0637 **	-0,0059	0,0062	0,0222	-0,0494 ***	-0,0337 *	0,0194	-0,0065
DUMMY_AGRICULTURAL	-0,0417	0,0376	-0,1196 ***	-0,0410	-0,0698 ·	-0,1311 ***	-0,0513	0,0300	-0,1062 **
DUMMY_PUBLIC	0,1131 ***	0,0361	0,0235	0,0088	0,0563 ***	0,0141	0,0874 ***	0,0655 **	0,0734 **
DUMMY_OTHER_SECTOR	0,0155	-0,0015	-0,0080	-0,0138	-0,0144	-0,0302 ·	-0,0015	0,0033	-0,0509 **
DUMMY_PART_TIME	0,0391 *	0,0822 **	0,0830 ***	-0,0605 **	-0,0136	0,0186	0,0157	0,0448	-0,0198
Residual First Stage	-0,0149 ***	-0,0192 ***	-0,0273 ***	-0,0268 ***	-0,0265 ***	-0,0360 ***	-0,0168 ***	-0,0285 ***	-0,0168 ***

(***p<0.01, **p<0.05, *p<0.1)

Appendix C to Chapter 3

C.1 Instrumental Variable Quantile Regression Estimates

Table C.1 - IVQR estimates. Dependent Variable: log of hourly wages less tax. Omitted categories are: Center (DUMMY_CENTER); Industrial sector (DUMMY_INDUSTRIAL), 1995-2012, by quantile

(*** p<0.01, ** p<0.05, * p<0.1)

Year 1995

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0466 ***	0,0528 ***	0,0584 ***	0,0600 ***	0,0605 ***	0,0635 ***	0,0678 ***	0,0707 ***	0,0753 ***
EXPERIENCE	0,0236 ***	0,0185 ***	0,0150 ***	0,0161 ***	0,0130 ***	0,0114 ***	0,0103 ***	0,0114 ***	0,0085 **
EXPERIENCE^2	-0,0003 ***	-0,0002 **	-0,0001	-0,0001	0,0000	0,0000	0,0000	0,0000	0,0001
DUMMY_MALE	0,1232 ***	0,1118 ***	0,0972 ***	0,0938 ***	0,0945 ***	0,1028 ***	0,0898 ***	0,1020 ***	0,1009 ***
DUMMY_MARRIED	0,0662 **	0,0328	0,0117	0,0123	0,0020	-0,0108	0,0128	0,0173	0,0088
NCOMP	0,0045	0,0168 **	0,0190 ***	0,0195 ***	0,0228 ***	0,0245 ***	0,0308 ***	0,0295 ***	0,0413 ***
DUMMY_HOUSEHOLD	0,0072	0,0235	0,0233	0,0207	0,0017	-0,0042	0,0055	0,0264	0,0355
DUMMY_TOWN	-0,0401 *	-0,0394 *	0,0012	0,0042	0,0041	0,0007	0,0171	0,0352	0,0551 *
DUMMY_NORTH	0,0425 **	0,0420 ***	0,0359 **	0,0347 **	0,0308 **	0,0311 **	0,0263 *	0,0416 **	0,0365 *
DUMMY_SOUTH	-0,0675 ***	-0,0292	-0,0131	-0,0029	0,0102	0,0151	0,0174	0,0216	0,0281
DUMMY_AGRICULTURAL	-0,3360 ***	-0,2335 ***	-0,1581 ***	-0,1340 ***	-0,1415 ***	-0,0767	0,0445	0,1728 ***	0,2708 **
DUMMY_PUBLIC	0,1753 ***	0,1259 ***	0,1061 ***	0,1098 ***	0,1168 ***	0,1229 ***	0,1046 ***	0,1028 ***	0,0986 ***
DUMMY_OTHER_SECTOR	-0,0319	-0,0359 **	-0,0147	0,0064	0,0105	0,0252	0,0468 ***	0,0559 ***	0,0760 ***
DUMMY_SECT_PARENTS	0,0346	0,0198	0,0099	0,0145	0,0094	0,0133	-0,0071	-0,0070	-0,0240
DUMMY_PART_TIME	-0,0372	0,0131	0,0158	0,0203	0,0137	-0,0011	-0,0053	0,0529	0,1775
Constant	0,9267 ***	0,9949 ***	1,0695 ***	1,1056 ***	1,2049 ***	1,2543 ***	1,2684 ***	1,2790 ***	1,3421 ***

Year 1998

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0651 ***	0,0625 ***	0,0606 ***	0,0559 ***	0,0543 ***	0,0646 ***	0,0602 ***	0,0532 ***	0,0682 ***
EXPERIENCE	0,0279 ***	0,0129 ***	0,0129 ***	0,0149 ***	0,0101 *	0,0103 **	0,0024	-0,0001	-0,0096
EXPERIENCE^2	-0,0004 **	0,0000	0,0000	-0,0001	0,0000	0,0000	0,0002	0,0002	0,0005 **
DUMMY_MALE	0,1120 **	0,1295 ***	0,1319 ***	0,1064 ***	0,1217 ***	0,1209 ***	0,1232 ***	0,1266 ***	0,1647 ***
DUMMY_MARRIED	0,0610	0,0791	0,0607	0,0629 *	0,0670 *	0,0591	0,0078	0,0256	-0,0064
NCOMP	0,0291 **	0,0149	0,0241 **	0,0215 **	0,0141	0,0232 *	0,0208	0,0061	0,0223
DUMMY_HOUSEHOLD	-0,0204	0,0186	-0,0107	0,0028	-0,0181	-0,0323	-0,0326	-0,0339	-0,0444
DUMMY_TOWN	-0,0343	-0,0172	-0,0523	-0,0505	0,0085	-0,0142	0,0206	0,0151	0,0180
DUMMY_NORTH	0,0449	0,0568 **	0,0519 **	0,0477 *	0,0373	0,0065	-0,0015	-0,0116	-0,0037
DUMMY_SOUTH	0,0273	0,0575 *	0,0637 **	0,0670 **	0,0648 **	0,0492	0,0452	0,0292	0,0518
DUMMY_AGRICULTURAL	-0,3055 ***	-0,1015	-0,1173	-0,0962	-0,0321	0,0350	0,0829	0,0675	0,3333
DUMMY_PUBLIC	-0,0075	0,0210	0,0275	0,0320	0,0371	0,0058	0,0453	0,1195 ***	0,0609
DUMMY_OTHER_SECTOR	-0,0151	0,0078	-0,0082	-0,0084	0,0097	-0,0366	-0,0176	0,0185	-0,0061
DUMMY_SECT_PARENTS	-0,0032	0,0222	0,0220	0,0163	0,0222	0,0251	0,0150	0,0456 *	0,0646 *
DUMMY_PART_TIME	-0,0129	0,0003	-0,0086	-0,0319	-0,0311	-0,0365	0,0017	0,0634	0,1041
Constant	0,6725 ***	0,8658 ***	0,9768 ***	1,1153 ***	1,2594 ***	1,2589 ***	1,4993 ***	1,6775 ***	1,7115 ***

Year 2000

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0628 ***	0,0673 ***	0,0635 ***	0,0633 ***	0,0660 ***	0,0669 ***	0,0717 ***	0,0760 ***	0,0845 ***
EXPERIENCE	0,0281 ***	0,0265 ***	0,0219 ***	0,0190 ***	0,0179 ***	0,0181 ***	0,0148 ***	0,0154 ***	0,0089 *
EXPERIENCE^2	-0,0004 ***	-0,0003 ***	-0,0002 ***	-0,0002 **	-0,0001 **	-0,0002 **	-0,0001	-0,0001	0,0001
DUMMY_MALE	0,1162 ***	0,1067 ***	0,1029 ***	0,1050 ***	0,1098 ***	0,1327 ***	0,1460 ***	0,1436 ***	0,1436 ***
DUMMY_MARRIED	0,0767 ***	0,0601 ***	0,0471 **	0,0352 *	0,0247	0,0138	0,0235	0,0420 *	0,0442
NCOMP	-0,0093	-0,0046	0,0072	0,0179 ***	0,0180 ***	0,0155 **	0,0143 **	0,0151 *	0,0120
DUMMY_HOUSEHOLD	-0,0108	-0,0035	-0,0040	0,0091	0,0129	-0,0050	-0,0079	-0,0069	0,0057
DUMMY_TOWN	-0,0329	0,0164	0,0330	0,0271	0,0277	0,0092	0,0289	0,0478 *	0,0520
DUMMY_NORTH	0,0868 ***	0,0655 ***	0,0476 ***	0,0402 ***	0,0494 ***	0,0594 ***	0,0591 ***	0,0575 ***	0,0700 ***
DUMMY_SOUTH	-0,0408	-0,0271	-0,0195	-0,0190	0,0108	0,0134	0,0284	0,0618 ***	0,0940 ***
DUMMY_AGRICULTURAL	-0,2117 **	-0,0915 *	-0,1045 **	-0,0823 **	-0,0665 *	-0,0361	0,0535	0,0930 *	0,1909 **
DUMMY_PUBLIC	0,0263	0,0145	0,0276	0,0395 **	0,0361 **	0,0439 **	0,0377 **	0,0264	0,0110
DUMMY_OTHER_SECTOR	-0,0358	-0,0135	-0,0146	0,0026	0,0163	0,0263 *	0,0365 **	0,0305	0,0274
DUMMY_SECT_PARENTS	-0,0032	-0,0034	0,0005	0,0013	-0,0106	-0,0196 *	-0,0293 **	-0,0418 ***	-0,0536 ***
DUMMY_PART_TIME	-0,0102	0,0060	-0,0126	-0,0182	0,0012	0,0051	0,0304	0,1062 **	0,1714 ***
Constant	0,7717 ***	0,8305 ***	0,9893 ***	1,0556 ***	1,0959 ***	1,1596 ***	1,2017 ***	1,2352 ***	1,3426 ***

	Year 2002																	
	0.1		0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9	
SCHOOL	0,0487	***	0,0576	***	0,0608	***	0,0636	***	0,0667	***	0,0688	***	0,0711	***	0,0724	***	0,0844	***
EXPERIENCE	0,0232	***	0,0203	***	0,0216	***	0,0166	***	0,0153	***	0,0167	***	0,0160	***	0,0100	***	0,0089	
EXPERIENCE^2	-0,0003	***	-0,0002	***	-0,0003	***	-0,0001	*	-0,0001		-0,0001	*	-0,0001		0,0000		0,0001	
DUMMY_MALE	0,0824	***	0,1008	***	0,1064	***	0,1033	***	0,0997	***	0,1056	***	0,1122	***	0,1326	***	0,1709	***
DUMMY_MARRIED	0,0431	*	0,0272		0,0160		0,0057		-0,0019		-0,0110		-0,0170		-0,0267		-0,0704	**
NCOMP	-0,0002		0,0088		0,0141	**	0,0146	**	0,0147	**	0,0132		0,0100		0,0216	**	0,0265	**
DUMMY_HOUSEHOLD	0,0404	**	0,0227		0,0193		0,0114		0,0000		0,0076		0,0061		0,0158		-0,0135	
DUMMY_TOWN	-0,0789	*	-0,0206		-0,0201		-0,0396	*	-0,0548	**	-0,0526	*	-0,0096		-0,0305		-0,0226	
DUMMY_NORTH	0,0793	***	0,0489	**	0,0345	**	0,0474	***	0,0457	***	0,0453	**	0,0219		0,0161		0,0177	
DUMMY_SOUTH	-0,0079		-0,0150		-0,0177		-0,0064		-0,0001		0,0137		0,0298		0,0144		0,0244	
DUMMY_AGRICULTURAL	-0,2478	***	-0,1222	***	-0,1403	***	-0,0943	**	-0,0743	**	-0,0643		-0,0362		-0,0424		0,0121	
DUMMY_PUBLIC	0,0366		0,0249		0,0202		0,0272		0,0325		0,0269		0,0462	*	0,0439	*	-0,0067	
DUMMY_OTHER_SECTOR	-0,0408	*	-0,0489	**	-0,0147		0,0021		0,0093		0,0016		0,0193		0,0299		0,0536	*
DUMMY_SECT_PARENTS	-0,0055		-0,0035		0,0209		0,0127		0,0048		0,0144		0,0211		0,0172		-0,0387	*
DUMMY_PART_TIME	-0,0716		-0,0267		-0,0331		-0,0174		-0,0098		0,0087		0,0110		0,0143		0,0392	
Constant	0,9525	***	1,0030	***	1,0249	***	1,1191	***	1,1754	***	1,2089	***	1,2745	***	1,3746	***	1,4660	***

	Year 2004																	
	0.1		0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9	
SCHOOL	0,0653	***	0,0666	***	0,0625	***	0,0605	***	0,0615	***	0,0609	***	0,0629	***	0,0652	***	0,0745	***
EXPERIENCE	0,0203	***	0,0197	***	0,0192	***	0,0186	***	0,0156	***	0,0154	***	0,0128	***	0,0135	***	0,0080	*
EXPERIENCE^2	-0,0002	***	-0,0002	***	-0,0002	***	-0,0002	***	-0,0002	**	-0,0001	**	-0,0001		-0,0001		0,0000	
DUMMY_MALE	0,1124	***	0,1067	***	0,1112	***	0,1169	***	0,1287	***	0,1187	***	0,1178	***	0,1454	***	0,1522	***
DUMMY_MARRIED	0,0104		0,0124		0,0044		-0,0084		-0,0176		-0,0143		0,0025		-0,0042		0,0084	
NCOMP	0,0018		0,0083		0,0115	*	0,0080		0,0093		0,0143	**	0,0118	*	0,0161	*	0,0098	
DUMMY_HOUSEHOLD	0,0032		0,0109		0,0157		-0,0020		-0,0116		-0,0076		-0,0096		-0,0236		-0,0047	
DUMMY_TOWN	0,0035		0,0085		-0,0133		0,0185		0,0061		0,0153		0,0105		-0,0094		-0,0182	
DUMMY_NORTH	0,0500	**	0,0491	***	0,0407	***	0,0431	***	0,0459	***	0,0479	***	0,0332	*	0,0298		0,0178	
DUMMY_SOUTH	-0,0215		-0,0036		-0,0033		0,0152		0,0129		0,0194		0,0265		0,0049		0,0132	
DUMMY_AGRICULTURAL	-0,2705	***	-0,1006		-0,0478		-0,0538		-0,0414		-0,0394		-0,0374		0,0327		0,0588	
DUMMY_PUBLIC	0,0619	**	0,0458	*	0,0527	**	0,0497	**	0,0521	***	0,0583	***	0,0612	***	0,0764	***	0,0357	
DUMMY_OTHER_SECTOR	-0,0171		-0,0190		-0,0108		-0,0107		0,0003		0,0016		0,0119		0,0354		0,0002	
DUMMY_SECT_PARENTS	0,0249		0,0077		0,0095		-0,0033		-0,0063		-0,0036		-0,0071		-0,0182		0,0033	
DUMMY_PART_TIME	-0,1023		0,0022		-0,0112		0,0011		0,0117		0,0203		0,0211		0,0252		0,0876	
Constant	0,8257	***	0,9321	***	1,0655	***	1,1908	***	1,2678	***	1,3136	***	1,3941	***	1,4517	***	1,5753	***

	Year 2006																	
	0.1		0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9	
SCHOOL	0,0553	***	0,0579	***	0,0560	***	0,0614	***	0,0654	***	0,0711	***	0,0731	***	0,0792	***	0,0838	***
EXPERIENCE	0,0259	***	0,0247	***	0,0201	***	0,0200	***	0,0178	***	0,0178	***	0,0186	***	0,0125	***	0,0152	***
EXPERIENCE^2	-0,0004	***	-0,0004	***	-0,0002	***	-0,0002	***	-0,0002	***	-0,0002	***	-0,0002	***	-0,0001		-0,0001	
DUMMY_MALE	0,1149	***	0,1086	***	0,1041	***	0,1132	***	0,1182	***	0,1098	***	0,1202	***	0,1137	***	0,1141	***
DUMMY_MARRIED	0,0259		0,0083		0,0250		0,0164		0,0141		0,0182		0,0205		0,0112		0,0235	
NCOMP	0,0024		0,0144	**	0,0161	**	0,0154	**	0,0174	***	0,0190	***	0,0246	***	0,0286	***	0,0376	***
DUMMY_HOUSEHOLD	0,0214		0,0090		0,0171		0,0142		0,0180		0,0226		0,0299	*	0,0380	*	0,0359	
DUMMY_TOWN	0,0542	**	0,0259		0,0153		0,0062		0,0174		0,0158		0,0044		0,0030		-0,0307	
DUMMY_NORTH	0,0100		0,0318	*	0,0224		0,0153		0,0048		0,0257		0,0031		-0,0123		-0,0245	
DUMMY_SOUTH	-0,0908	***	-0,0529	**	-0,0490	**	-0,0212		-0,0059		0,0124		-0,0089		-0,0184		-0,0017	
DUMMY_AGRICULTURAL	-0,2067	***	-0,1551	***	-0,1166	***	-0,1278	***	-0,1137	***	-0,0685		-0,0391		-0,0124		0,0804	
DUMMY_PUBLIC	0,0488		0,0451	**	0,0583	***	0,0418	**	0,0348	*	0,0333		0,0457	*	0,0343		0,0453	
DUMMY_OTHER_SECTOR	-0,0489	**	-0,0262		-0,0034		0,0015		0,0127		0,0069		0,0216		0,0252		0,0312	
DUMMY_SECT_PARENTS	-0,0185		-0,0071		-0,0048		-0,0160		-0,0032		-0,0106		-0,0040		-0,0269	*	-0,0296	
DUMMY_PART_TIME	-0,0762	*	-0,0542	**	-0,0498	*	-0,0238		-0,0189		-0,0182		-0,0235		-0,0007		0,0544	
Constant	0,9727	***	1,0226	***	1,1388	***	1,1658	***	1,1890	***	1,1731	***	1,1994	***	1,2896	***	1,3068	***

	Year 2008								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0513 ***	0,0463 ***	0,0565 ***	0,0619 ***	0,0626 ***	0,0673 ***	0,0684 ***	0,0712 ***	0,0754 ***
EXPERIENCE	0,0200 ***	0,0239 ***	0,0229 ***	0,0203 ***	0,0205 ***	0,0173 ***	0,0138 ***	0,0125 ***	0,0139 ***
EXPERIENCE^2	-0,0002 ***	-0,0003 ***	-0,0003 ***	-0,0002 ***	-0,0002 ***	-0,0002 **	-0,0001	0,0000	-0,0001
DUMMY_MALE	0,0992 ***	0,1121 ***	0,1128 ***	0,1127 ***	0,1258 ***	0,1535 ***	0,1552 ***	0,1555 ***	0,1679 ***
DUMMY_MARRIED	-0,0219	-0,0025	-0,0082	-0,0121	-0,0298	-0,0456 **	-0,0348	-0,0251	-0,0285
NCOMP	0,0256 ***	0,0132 **	0,0170 **	0,0153 **	0,0224 ***	0,0276 ***	0,0285 ***	0,0313 ***	0,0326 **
DUMMY_HOUSEHOLD	0,0483 *	0,0237	0,0322 *	0,0189	0,0086	0,0002	0,0051	-0,0105	-0,0183
DUMMY_TOWN	-0,0025	-0,0244	0,0041	0,0163	0,0429 *	0,0335	0,0323	0,0152	0,0176
DUMMY_NORTH	0,0438 **	0,0291	0,0120	0,0030	0,0013	-0,0248	-0,0330	-0,0492 **	-0,0777 **
DUMMY_SOUTH	-0,0302	-0,0206	-0,0109	0,0096	0,0170	0,0067	0,0015	0,0036	-0,0406
DUMMY_AGRICULTURAL	-0,1714 **	-0,1347 **	-0,1021 **	-0,0744	-0,0724 *	-0,1195 ***	-0,1010 **	-0,0996 *	-0,0154
DUMMY_PUBLIC	0,0669 *	0,0683 ***	0,0439 *	0,0235	0,0386	0,0321	0,0364	0,0534 *	0,0511
DUMMY_OTHER_SECTOR	-0,0547 **	-0,0373 **	-0,0223	-0,0261	-0,0175	-0,0175	-0,0296	-0,0059	0,0025
DUMMY_SECT_PARENTS	0,0222	0,0145	0,0169	0,0070	0,0157	0,0150	0,0086	0,0151	0,0356
DUMMY_PART_TIME	-0,0351	-0,0024	0,0257	0,0438	0,0511 *	0,0781 ***	0,1021 ***	0,1250 ***	0,1102 **
Constant	0,9444 ***	1,1248 ***	1,0744 ***	1,1214 ***	1,1532 ***	1,2101 ***	1,2963 ***	1,3588 ***	1,4575 ***

	Year 2010								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0502 ***	0,0433 ***	0,0454 ***	0,0632 ***	0,0564 ***	0,0567 ***	0,0666 ***	0,0777 ***	0,0852 ***
EXPERIENCE	0,0268 ***	0,0201 ***	0,0161 ***	0,0185 ***	0,0162 ***	0,0145 ***	0,0126 ***	0,0081 *	-0,0013
EXPERIENCE^2	-0,0004 ***	-0,0003 ***	-0,0002 **	-0,0002 **	-0,0001 *	-0,0001	0,0000	0,0001	0,0003 *
DUMMY_MALE	0,1290 ***	0,1144 ***	0,1289 ***	0,1532 ***	0,1536 ***	0,1455 ***	0,1578 ***	0,1649 ***	0,1663 ***
DUMMY_MARRIED	-0,0301	-0,0196	-0,0221	-0,0156	-0,0039	0,0004	-0,0067	0,0188	-0,0212
NCOMP	0,0265 ***	0,0181 **	0,0169 **	0,0146 *	0,0131 *	0,0155 *	0,0158 *	0,0135	0,0275 **
DUMMY_TOWN	0,0189	0,0626 **	0,0650 ***	0,0499 *	0,0826 ***	0,0593 **	0,0704 ***	0,0414	-0,0062
DUMMY_NORTH	0,0626 **	0,0443 **	0,0494 ***	0,0532 ***	0,0138	0,0129	0,0176	0,0116	-0,0163
DUMMY_SOUTH	0,0137	0,0145	0,0231	0,0225	0,0130	0,0139	0,0312	0,0229	0,0078
DUMMY_AGRICULTURAL	-0,2385 ***	-0,2334 ***	-0,1628 *	-0,0731 *	-0,1095 ***	-0,1140 **	-0,1034 *	-0,1242 *	0,0195
DUMMY_PUBLIC	0,0437	0,0800 ***	0,0892 ***	0,0469 *	0,0641 **	0,0666 **	0,0424	0,0230	0,0273
DUMMY_OTHER_SECTOR	-0,0415 *	-0,0206	-0,0104	-0,0125	-0,0008	0,0107	0,0035	-0,0166	0,0158
DUMMY_SECT_PARENTS	0,0162	0,0092	0,0200	0,0132	0,0049	0,0157	0,0136	0,0129	-0,0113
DUMMY_PART_TIME	-0,0423	-0,0601 *	-0,0173	0,0049	0,0298	0,0184	0,0158	0,0561	0,0175
Constant	0,9201 ***	1,1953 ***	1,2302 ***	1,0599 ***	1,2187 ***	1,2832 ***	1,2611 ***	1,2874 ***	1,4307 ***

	Year 2012								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
SCHOOL	0,0488 ***	0,0542 ***	0,0530 ***	0,0593 ***	0,0630 ***	0,0630 ***	0,0648 ***	0,0738 ***	0,0771 ***
EXPERIENCE	0,0332 ***	0,0289 ***	0,0241 ***	0,0218 ***	0,0218 ***	0,0192 ***	0,0208 ***	0,0165 ***	0,0102
EXPERIENCE^2	-0,0005 ***	-0,0004 ***	-0,0003 ***	-0,0002 ***	-0,0002 **	-0,0002 *	-0,0002 **	-0,0001	0,0000
DUMMY_MALE	0,1366 ***	0,1208 ***	0,1080 ***	0,1270 ***	0,1162 ***	0,1115 ***	0,1099 ***	0,0989 ***	0,1345 ***
DUMMY_MARRIED	0,0469	0,0126	0,0159	-0,0038	0,0052	0,0046	0,0066	-0,0284	-0,0535
NCOMP	-0,0008	0,0087	0,0180 **	0,0256 ***	0,0270 ***	0,0307 ***	0,0320 ***	0,0449 ***	0,0520 ***
DUMMY_TOWN	-0,0746	-0,0359	-0,0281	-0,0261	-0,0046	-0,0260	0,0140	0,0157	0,0152
DUMMY_NORTH	0,0540 *	0,0401 *	0,0340 *	0,0623 ***	0,0563 ***	0,0431 *	0,0294	-0,0079	-0,0278
DUMMY_SOUTH	-0,0057	-0,0199	-0,0080	0,0265	0,0298	0,0336	0,0480 *	0,0255	-0,0177
DUMMY_AGRICULTURAL	-0,1481 ***	-0,1867 ***	-0,1975 ***	-0,1881 ***	-0,1758 ***	-0,1328 **	-0,1153 **	-0,1112 **	-0,1284
DUMMY_PUBLIC	0,0283	0,0474	0,0333	0,0212	0,0254	0,0235	0,0167	-0,0133	0,0222
DUMMY_OTHER_SECTOR	-0,0639 **	-0,0239	-0,0162	-0,0329	-0,0211	-0,0302	-0,0199	-0,0128	0,0424
DUMMY_SECT_PARENTS	0,0318	0,0183	0,0165	0,0159	0,0118	0,0281	0,0255	-0,0181	-0,0053
DUMMY_PART_TIME	-0,1156 **	-0,0727 **	-0,0641 **	-0,0494 *	-0,0389	-0,0183	-0,0048	-0,0144	0,0384
Constant	0,8023 ***	0,8794 ***	1,0093 ***	0,9822 ***	1,0008 ***	1,1028 ***	1,1346 ***	1,2214 ***	1,3513 ***

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*, Vol. 70, pp. 91-117.

Acemoglu, D. and Angrist, J. D. (2001). How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws. In B. Bernanke and K. Rogoff, NBER Macroeconomics Annual (9-74). Cambridge: National Bureau of Economic Research.

Angrist, J. D., Chernozhukov, V. and Fernandez-Val, I. (2006). 'Quantile regression under misspecification, with an application to the U.S. wage structure', *Econometrica*, Vol. 74, pp. 539-563.

Angrist, J. D., G. W. Imbens and Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables, *Journal of the American Statistical Association*, Vol. 91, pp. 444-455.

Angrist, J. D. and Newey, W. (1991). Over-Identification Test in Earnings Function with Fixed Effects, *Journal of Business and Economic Statistics*, pp. 317-323.

Angrist, J. D. and Krueger, A. B. (1991). Does Compulsory Schooling Affect Schooling And Earnings?, *Quarterly Journal of Economics*, pp. 979-1014.

Angrist, J. D. and Krueger, A. B. (1992). Estimating the Payoff to Schooling Using the Vietnam Era Draft Lottery, NBER Working Paper No. 4067.

Antonelli, G. (1985). Risorse Umane e Redditi da lavoro. Franco Angeli, Milan.

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

Ashenfelter, O. and Zimmerman, D. J. (1993). Estimates of the Return to Schooling from Sibling Data: Fathers, Sons, and Brothers", NBER Working Paper No. 4491.

Ashenfelter, O. and Krueger, A. B. (1994). Estimates of the Economic Return to Schooling from a New Sample of Twins", *American Economic Review*, Vol. 84, pp. 1157-73.

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

Ashenfelter, O., Harmon, C. and Oosterbeek, H. (1999). A Review of the Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias, *Labour Economics,* Vol. 6, pp. 453-470.

Becker, G. S. (1962). Investment in Human Capital: A theoretical Analysis, *The Journal of Political Economy*, Vol. 70, pp. 9-49.

Becker, G. S. (1981). Altruism in the Family and Selfishness in the Market Place, *Economica*, New Series, Vol. 48, No. 189, pp. 1-15.

Becker, G. S. (1991). Treatise on the Family, Cambridge: Harvard University Press. (first ed. 1981)

Becker, G. S. (1993). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, Chicago: University of Chicago Press (first ed. 1975).

Belzil, C. and Hansen, J. (2002). 'Unobserved ability and the return to schooling', *Econometrica*, Vol. 70, pp. 2075-2091.

Belzil, C. (2004). On the Specification of Mincerian Wage Regressions with Heterogeneity, Non-linearity, Non-separability, and Heteroskedasticity, IZA Discussion Paper No. 1083, IZA.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings, *Journal of Political Economy*, Vol. 75, pp. 352-65.

Bjorklund, A. and C. Kjellstrom (2002). Estimating the return to investments in education: how useful is the standard Mincer equation?" *Economics of Education Review,* Vol. 21, pp. 195-210.

Blundell, R., Dearden, L. and Sianesi, B. (2001). Estimating the Returns to Education: Models, Methods and Results, Centre for the Economics of Education, London School of Economics.

Blundell, R. and Powell, J. L. (2001). Endogeneity in Nonparametric and Semiparametric Regression Models, Cemmap Working Paper CWP09/01, The Institute for Fiscal Studies; Department of Economics, UCL.

Blundell, R. and Powell, J. L. (2003). Advance in Economics and Econometrics, chapter Endogeneity in Nonparametric and Semiparametric Regression Models, pp. 312-357. Cambridge University Press.

Blundell, R., Dearden, L. and Sianesi, B. (2005). Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey, *Journal of the Royal Statistical Society*, Series A (Statistics in Society), Vol.168, Issue 3, pp. 473-512.

Bonjour, D., Cherkas, L., Haskel, J., Hawkes, D. and Spector, T. (2003). Returns to Education: Evidence from U.K. Twins, *The American Economic Review*, Vol. 93, pp. 1799-1812.

Bound, J., Jaeger, D. and Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak, *Journal of the American Statistical Association*, pp. 443-450.

Bowman, A.W. and Azzalini, A. (1997). Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations. Oxford University Press, Oxford.

Brandolini, A. and Cannari, L. (1994). “Methodological appendix: the Bank of Italy's Survey of Household income and wealth”, in Ando A., Guiso L., Visco I. (a cura di), Savings and the accumulation of wealth. Essays on Italian households and government saving behaviour, Cambridge University Press.

Brunello, G. and Miniaci, R. (1999). The Economic Returns to Schooling for Italian men. An Evaluation Based on Instrumental Variables, *Labour Economics*, Vol.6, pp. 509-519.

Brunello, G., Comi, S. and Lucifera, C. (1999). The Returns to Education in Italy: A Review of the Applied Literature, in R. Asplund and P.T. Pereira, Returns to Human Capital in Europe. A Literature Review, ETLA, Helsinki.

Brunello, G., Comi, S. and Lucifera, C. (2000). The Returns to Education in Italy: A new look at evidence, IZA DP No. 130.

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

Buchinsky, M., (1994). Changes in the US wage structure, 1963- 1987: application of quantile regression, *Econometrica*, Vol. 62, pp. 405- 458.

Cannari, L., Pellegrini, G. and Sestito, P. (1989). Età, Istruzione e Retribuzioni, Contributi all’analisi economica, Bank of Italy, Rome.

Cannari, L., and D’Alessio, G. (1995). Il Rendimento Dell’Istruzione: Alcuni Problemi Di Stima. Temi di Discussione n. 253, Bank of Italy, Rome.

Card, D. and Krueger, A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States, *The Journal of Political Economy*, Vol. 100, pp. 1-40.

Card, D. (1995). Earnings, Schooling and Ability Revisited. In S. Polacheck (ed.), Research in Labor Economics, Vol. 14, 23-48, Greenwich: JAI Press. (NBER, Working Paper No. 4832)

Card, D. (1999). The Causal Effect of Education on Earnings. In Ashenfelter, O. and D. Card, Handbook of Labor Economics, Rotterdam: Elsevier.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems, *Econometrica*, Vol. 69, pp.1127-60.

Carrasco, R., Jimeno, J. F., Ortega, A.C (2014). Returns to Skills and the Distribution of Wages: Spain 1995-2010, *Oxford Bulletin of Economics and Statistics*, Vol. 77, pp. 542-565.

Checchi, D., Ichino, A. and Rustichini, A. (1999). More equal but less mobile? Intergenerational mobility and inequality in Italy and in the US, *Journal of Public Economics*, Vol. 74, pp. 351-393.

Checchi D. (2004). Does educational achievement help to explain income inequality?, in A. Cornia, ed., Inequality, Growth and Poverty in an Era of Liberalization and Globalization, Oxford.

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

Chernozhukov, V., Hansen, C. and Jansson, M. (2007). 'Inference approaches for instrumental variable quantile regression', *Economics Letters*, Vol. 95, pp. 272-277.

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

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

Chiswick, B. (2003). Jacob Mincer, Experience and the Distribution of Earnings, *Review of Economics of the Household*, Vol. 1, pp. 343-361.

Ciccone, A. (2004). Human Capital as a Factors of Growth and Employment at the Regional Level. The Case of Italy. Report for the European Commission, DG for Employment and Social Affairs.

Ciccone, A., Cingano, F. and Cipollone, P. (2006). The private and social return to schooling in Italy, Temi di Discussione n. 569, Bank of Italy, Rome.

Cingano, F., and Cipollone, P. (2009). I rendimenti dell'istruzione, Questioni di Economia e Finanza (Occasional Papers) n.53, Bank of Italy, Rome.

Colussi, A. (1997). Il Tasso Di Rendimento dell'Istruzione In Italia: un'Analisi Cross-Section, *Politica Economica*, Vol. 13, pp. 273-294.

Davidson, R. and MacKinnon, J. G (2007). Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis*, Vol. 51, pp. 3259-3281.

De La Fuente, A. and Ciccone, A. (2002). Human Capital in a Global and Knowledge-Based Economy, European Commission: DG Employment and Social Affairs.

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* Vol. 30, pp. 1118-1122.

Dickson, M. (2013). The causal effect of education on wages revisited, *Oxford Bulletin of Economics and Statistics*, Vol. 75, pp. 477-498.

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.

Flabbi, L. (1997). Investire In Istruzione: Meglio per lui o per lei?, mimeo, Bocconi University, Milan.

Flabbi, L. (1999). Returns to Schooling in Italy: OLS, OV and Gender Differences. Working paper of Bocconi University, Serie di econometria ed economia applicata.

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

Giustinelli, P. (2004). Quantile regression evidence on Italian education returns, *Rivista di Politica Economica*, XI-XII, pp. 53-107.

Griliches, Z. (1977). Estimating the Returns to Schooling: Some Econometric Problems, *Econometrica*, Vol. 45, pp. 1-22.

Harmon, C.P. and Walker, I. (1995). Estimates of the Economic Return to Schooling for the UK, *American Economic review*, pp. 1278-86.

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

Harmon, C., Hogan, V. and Walker, I. (2003). Dispersion in the Economic Return to Schooling, *Labour Economics*, Vol. 10, pp. 205-214.

Harmon, C.P., Oosterbeek, H and Walker, I. (2003). The Returns to Education: Microeconometrics, *Journal of Economic Surveys*, Vol. 17, pp. 115-155.

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,* Vol. 33, pp. 1021-1037.

Hastie, T. and Tibshirani, R. (1986) Generalized additive models (with discussion). *Statistical Science*, Vol. 1, pp. 297-318.

Hastie, T. and Tibshirani, R. (1990) Generalized Additive Models. London: Chapman and Hall.

Hastie, T. and Tibshirani, R. (1993) Varying-coefficient models (with discussion). *Journal of Royal Statistical Society B*, Vol. 55, pp. 757-796.

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

Heckman, J. J. (1979). Sample Selection Bias as A Specification Error, *Econometrica*, Vol. 47, pp. 153-161.

Heckman, J. J. and Killingsworth, M. R. (1986). Female Labor Supply: a Survey, in Ashenfelter O. and R. Layard (ed.) Handbook of Labor Economics, Amsterdam: North Holland.

Heckman, J. J., Lochner, L. and Todd, P. (2003). Fifty Years of Mincer Earnings Regressions, NBER, Working Paper no. 9732.

Heckman, J. J., Lochner, L. and Todd, P. (2006). Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond, In E. Hanushek and F. Welch, Handbook of the Economics of Education, pp. 307-458, Rotterdam: Elsevier.

Heckman, J. J., Lochner, L. and Todd, P. (2008). Earnings Functions and Rates of Return, *Journal of Human Capital*, Vol. 2, pp. 1-31.

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

Horowitz, J. and Lee, S. (2002). Semiparametric methods in applied econometrics: Do the models fit the data", *Statistical Modeling*, Vol. 2, pp. 3-22.

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

Ichino, A. and Winter-Ebmer, R. (1999). Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments, *European Economic Review*, Vol. 43, pp. 889-901.

Idson, T. L. and Feaster, D. J. (1991). A Selectivity Model of Employer-Size Wage Differentials, *Journal of Labor Economics*, Vol. 8, pp. 99-122.

Imbens, G. and Angrist, J. D. 1994. Identification and estimation of local average treatment effects, *Econometrica*, Vol. 62, pp. 467-475.

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

Koenker, R., Basset, G.S. (1978). Regression quantiles, *Econometrica*, Vol. 46, pp. 33-50.

Koenker, R. (2005). Quantile Regression, Cambridge University Press, Cambridge.

Koop, G., Tobias, J. L. (2004). Learning about heterogeneity in returns to schooling, *Journal of Applied Econometric*s, Vol. 19, pp. 827-849.

Lam, D. and Schoeni, R. F. (1993). Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil, *Journal of Political Economy*, Vol. 101, pp. 710-740.

Lemieux, T. (2008). The changing nature of wage inequality, *Journal of Population Economics* Vol. 21, pp. 21-48.

Lucas R.E. (1988). On the Mechanics of Economic Development, *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

Lucifora, C. and Reilly, B. (1990). Wage Discrimination and Female Occupational Intensity, *Labour*, Vol.4, pp.147-168.

Ng, S. and Pinkse, J. (1995). Nonparametric Two-step Estimation of Unknown Regression Functions when the Regressors and the Regression Error are Not Independent. Université de Montréal, Département des Sciences Économiques, Montréal.

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

Meghir, C. and Rivkin, S. G. (2010). Econometric Methods for Research in Education, NBER, Working Paper no. 16003.

Mendolicchio, C. (2006). A Disaggregate Analysis of Private Returns to Education in Italy. Discussion Paper 54, Département des Sciences Économiques de l'Université catholique de Louvain.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *The Journal of Political Economy*, Vol. 66, pp. 281-302.

Mincer, J. (1974). Schooling, Experience and Earnings. New York: Columbia University Press.

OECD (2014) Education at a Glance, Paris.

Naticchioni P., Ricci A. and Rustichelli E. (2009). Far Away from a Skill-biased Change: Falling Educational Wage Premia in Italy, *Applied Economics*, Vol. 42, pp. 3383-3400.

Patrinos, H. A., Ridao-Cano, C. and Sakellariou, C. (2006). Estimating the Returns to Education: Accounting for Heterogeneity in Ability, World Bank Policy Research Working Paper 4040.

Patrinos, H. A., Ridao-Cano, C. and Sakellariou, C. (2009). A Note on Schooling and Wage Inequality in Public and Private Sector, *Empirical Economics*, Vol. 37, pp.383-92.

Peracchi, F. (2004). Educational wage premia and the distribution of earnings: an international perspectives, in E.Hanushek and F.Welch (eds.), *Handbook of the Economics of Education*, North-Holland, Amsterdam.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update, *World Developmen*t, Vol. 22, pp. 1325-1343.

Rosen, S. (1974). Measuring the Obsolescence of Knowledge, in T. Juster.

Schultz, T. W. (1961). Investment in “Human Capital”, *The American Economic Review*, Vol. 51.

Sianesi, B. and Reenen, J. V. (2003). The Returns to Education: Macroeconomics, *Journal of Economic Surveys*, Vol. 17, pp.157-200.

Spence, M. (1973). Job Market Signaling, *Quarterly Journal of Economics*, Vol. 87, pp. 355-373.

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

Vella, F. and Gregory, R. G. (1996), Selection Bias and Human Capital Investment: Estimating the Rates of Return to Education for Young Males, *Labour Economics*, pp. 197-219.

Vieira, J.A.C. (1999). Returns to Education in Portugal, *Labour Economics*, Vol. 6, pp. 535-541.

Willis, R. (1986). Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions. In O. Ashenfelter and R. Layard (eds), *Handbook of Labor Economics*, Amsterdam and New York: North Holland.

Wood, S. N. (2006). *Generalized Additive Models. An Introduction with R*. Chapman and Hall, London, 2006.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B),* Vol. 1, pp. 3-36.

Wooldridge, J.M. (2012). *Econometric Analysis of Cross Section and Panel Data*. MIT press, 2nd Edition.

Zheng, J. X. (2000). A nonparametric analysis of the US earnings distribution, in *Advances in Econometrics: Applying Kernel and Nonparametric Estimation to Economic Topics*, Vol. 14 (Eds) T. Fomby and R. C. Hill, JAI Press, Stamford, Connecticut, pp. 129–54.