Bachelor degree thesis in Econometrics

Quantile Effects of Education on Earnings: The Italian Case

Candidate
Siria Angino

Thesis coordinator
Giuseppe Ragusa

Academic year 2010/11

Libera Università Internazionale degli Studi Sociali Guido Carli
Faculty of Economics
Course in Business Economics

Abstract

Does education pay? Whom it pays more? And how much?

The aim of this paper is twofold. On one side it provides an overview of the theoretical framework and issues connected to the economics of education, and specifically to the estimation of the causal effect of schooling on earnings.

On the other, practical side, the paper considers how the effect of education evolved in Italy from 1993 to 2008, according to the education level and also to the type of education received. The observed pattern suggests a constant increase from 1998 to 2006 and a decrease in 2008, suggesting a follow-up decline in the next years. Moreover, data show evidence in favor of heterogeneity of the schooling effect. Finally the use of the quantile regression allows us to dig deeper in these results and analyze how the effect of education varies along the earnings distributions.

Keywords: education, earnings, quantile regression.

Introduction. A Practice of Mistitling.

While writing this dissertation, I consulted dozens of papers whose authors were committing themselves to find “the return to education” in a given time and place. But then, analyzing their methodologies, I was confused. Was the return to education the thing which they inquired on? I must disagree.

The concept of ‘return’ is the one that applies to all investments: putting aside the different methods to calculate it, a return is mainly the ratio of the amount of money gained or lost on an investment relative to the amount invested. Relating this definition to education, we should be able to evaluate with reasonable accuracy, and in measurable terms: a) all the benefits and b) all the costs incurred because of education.

The returns to education are of three kinds: private, external and productive. The ‘social return’ is usually defined as the sum of the first two. Let’s try to spot benefits and costs for each category.

A private benefit is, first of all, the one this paper is going to demonstrate and quantify: the increase in wage. But education may also confer a higher position in the workplace organization chart and a higher probability to find a job (because a higher qualification often unlocks requirements and opens new possibilities[1]). Last, it awards the educated with a self-satisfying culture.

Direct private costs are taxes and fees; a serious study should include also school materials (books, calculators, uniforms), transportation and meal expenses. Pushing far, we may also want to consider intangible costs such as studying efforts and other sacrifices undertaken in order to study.

The student will also incur an indirect private cost – the opportunity cost – because he/she will not be working (at least full time) while attending school/university. The opportunity cost should be a proxy for the cost of the student’s time, measured in terms of earning foregone. Generally the approximation used is the average wage of people with the level education immediately lower than the one we are considering: for example, in estimating the opportunity cost of attending university, we will consider the average wage of people who have finished the high school and are now working.

[1] For example, in Italy many occupations in the public sector are available only to people who achieved a certain education level (to be a mailmen, you must have a high school diploma with a minimum score of 70).

When we move to the external benefits, the situation is even harder to disentangle. Society could enjoy a gain because, when the share of educated people increases, if it true that a) a higher education leads to a higher wage, and b) taxes are paid proportionally to wages, then more taxes will be paid, benefiting the whole society.

Furthermore, the positive association between education and health is well established by many studies[2]. Education reduces crime, increases welfare participation, contributes to election participation and good voting behaviors[3], enhances hygiene and safety standards. More educated people are innovative.

All these improvements relapse not just on the single educated individual, but also on his/her family and other people around through an effect that is called "spillover effect" or simply "externality".

Societies with high shares of educated people have better capabilities of understanding, processing and transmitting new information and of discerning interdependencies among different features; as a consequence they undertake better collective decisions.

Moreover, where there is an exchange of ideas among individuals, it is easy that from this stimulating situation innovation emerges, and it is beneficial to the whole society and economy. This positive effect of gathering and confronting ideas is the 'knowledge spillover'.

Direct external costs depend on whether the education system is subsidized or not. In Italy, the State covers the majority of the expenses, including teachers' wages, buildings, equipments and grants. The amount of money spent for each student, however, differs from region to region, especially if we consider university costs.

Indirect external costs are the foregone taxes on the foregone earnings of people deciding to go on with education. Moreover, especially in the last few years, Italy is facing the problem of large masses of students going abroad to find a job; thus Italy is paying for the education of future workers that will not return benefits to the country. We can consider these lost benefits as a lost profit.

Finally, productivity benefits are both in terms of labor force and innovation. Education is said to enhance gross workforce productivity[4]: among other effects, education

[2] "Compared to the poorly educated, well educated respondents are less likely to be unemployed, are more likely to work full-time, to have fulfilling, subjectively rewarding jobs, high incomes, and low economic hardship [...] (that) in turn significantly improve health. The well educated report a greater sense of control over their lives and their health, and they have higher levels of social support (which)[...]are associated with good health. The well educated are less likely to smoke, are more likely to exercise, to get health check-ups, and to drink moderately, all of which, except check-ups, are associated with good health. We conclude that high educational attainment improves health directly, and it improves health indirectly through work and economic conditions, social-psychological resources, and health lifestyle." American Sociological Review © 1995

[3] [31]

[4] Doubts have emerged lately on this point, but we will discuss this later while talking about the 'signaling effect'.

may lead to a higher mental flexibility, thus to a faster adoption of new technologies and processes. A second improvement comes from the research and development efforts, that are tightly linked to universities (as well as to investments in R&D, of course). Third, let's go back to the knowledge spillover and move it in an industrial framework: an innovation in one firm enhances positive competition in the whole sector, thus boosting technological development and higher productivity.

Recall that, in order to achieve an estimation for the educational return we must be able to evaluate the value of these gains in monetary terms. And estimating externalities is not an easy task. You may be tempted to argue that some of these effects are negligible: but how can we say this if we don't investigate them? We cannot progress having prejudices.

Evaluating all these benefits and costs seems a nearly-impossible task. And in fact, counter-intuitively, this is not what most papers whose titles contain the expression "return to education" tried.

One of the pillars of economics of education, David Card, named carefully his fundamental 1999 paper: "The Causal Effect of Education on Earnings". This sounds way more precise and feasible. And it is what many papers in economics of education try: to study the effect on earnings of a change in the education level, without considering the costs associated with this additional education. This is already difficult and full of potential errors without adding also returns.

As in Card 1999,

> *"It is now conventional to refer to b as the 'return to education'. As shown in Willis (1986), [...] b is in fact the internal rate of return to schooling investment, assuming that education is free and that the students earn nothing while in school."*

This paper will consider how the effect of education evolved in Italy from 1993 to 2008. In particular, we think that it is reasonable to assume that an increase in education will have different outcomes if we consider different levels of income: we will analyze this difference using the quantile regression. But first of all let's try to explain the theory behind the practical analysis.

The datasets are taken from the "Indagine dei bilanci delle famiglie italiane" (Survey on Households Income and Wealth - SHIW) by Banca d'Italia.

Part I.

THEORY & ISSUES

1. CAUSALITY

1.1. Cum hoc ergo propter hoc: the false cause fallacy

Despite a strong evidence of a positive correlation between education and wage level, scholars are cautious, even reluctant, to take the step beyond and conclude in favor of a causal effect of schooling.

It is a fundamental logical law: correlation does not imply causation[1]. A common fallacy occurs with spurious relationships, where two events X and Y are correlated and one is considered the cause of the other, while in reality there exists a third factor Z (called 'lurking variable') that causes both X and Y. The number of umbrellas' vendors in the streets is correlated with traffic in Rome, but it does not mean that the vendors cause traffic (or vice versa). Simply, it exists a common cause – the rain – behind both events.

Why do we need to move beyond correlation and try to detect the true causal effect, apart from simple love of truth? Because of policy-making. If there is no causality between education and earnings, then increasing earnings using education would be as useful as decreasing traffic reducing umbrellas' vendors.

Similarly, if the causal effect exists, but it is overestimated, governments would undertake expensive measures to support schooling in the expectation of great future benefits, overvalued as well, with a huge waste of money.

Now let's try to see how this fallacy may enter our regression, and how we can keep it away.

When we create a model, we estimate the causal effect of one or more factors (regressors, or independent variable) on one variable of interest (dependent variable). The sign and the magnitude of the effect are given by the coefficient (or parameter) of the factors.

If we incorrectly leave out one or more important causal factors, the model will compensate these missing factors by under- or over- estimating the coefficient of one or more included factors. The missing variable is usually referred as "omitted variable" (OV), and it must satisfy two conditions to be called so:

a) it has an effect on the dependent variable;

[1]However correlation is a necessary condition for causation to exist.

b) it is correlated with at least one of the included independent variables.

This omitted variable is nothing more than the lurking variable we mentioned before.

In our case, the dependent variable is the logarithm of wages, and the main independent variable is 'education' (later we will see how to measure it). Now we should ask "Is there something that is correlated with wages and that influences education?"

Yes, there are lot of things, indeed. If we think carefully, we may also be able to guess the sign of the bias[2], but we cannot guess its magnitude without more information.

Think to the place an individual studied and works, and assume that they correspond. If this is a poor area, education may be poor as well because of few investments in schooling (correlation between OV and independent variable). And a poor area also gives the individual less opportunities to find a well-paid job (correlation between OV and dependent variable).

If no variable referring to the place the individual lives in is included in the regression, the coefficient of education will account also for the effect of the wealth of the area and it will be overestimated[3].

The coefficient of education everybody is looking for could be explained in this bizarre way. The quantum mechanics 'Many Worlds Interpretation' says that, when a decision between 'a' and 'b' is made, the universe splits into two world: one in which 'a' is chosen, another in which 'b' is chosen. If we could study the same individual in two parallel universes that are absolutely equal, except for the fact that in one universe the individual decided to take a university degree, while in the other he decided to stop with the high school diploma, and we could compare their wages after a certain time, well, we would have the true average causal effect of education (repeating the analysis for many individuals, of course). Let's take this idea as a benchmark: this would be the perfect estimation. In practice, however, some studies try to do something similar matching individuals who are as similar as possible.

To estimate the true causal effect of education on earnings, we must purge the regression from all possible biases arising from omitted variables. How do we obtain

[2]An easy way to estimate if there is an upward or downward bias is to establish if there is a positive or negative correlation between:

- the dependent variable and the independent variable;
- the dependent variable and the omitted variable;
- the independent variable and the omitted variable.

Then we count the positive and the negative correlations within each pair.
+ + + Upward bias
+ + - Downward bias
+ - - Upward bias
- - - Downward bias

[3]Assuming that a dummy variable D is added, with D=0 if the individual lives in a poor area, and D=1 if he lives in a rich area.

this 'purge'? One way is to control for these variables, meaning, we include them in the regression, so that we can study the change in earnings due to a change in the education level all other things being the same. It is the same idea behind controlling identical individuals in universes identical but for the education level choice.

Controlling for omitted variables is a good idea if they are measurable and available in the dataset... But there is one thing that it is as hard to measure as to define: talent, capacity, skill, ability. The bias of biases.

1.2. The Phantom Menace. The ability bias.

The history of economics of education begins in the late 1950s with a harsh debate. Some scholars started to explain the post-war USA productivity growth with the increased average education level, leaving little room to the importance of technological changes. Someone else disagreed: the conclusion was too straightforward. They pointed out that the underlying assumption - the fact that the earnings differences between education levels reflect true productivity differences caused by the education itself, and not by innate capacities - had never been proved.

This innate, probably non-measurable component is at the core of the most-known bias in econometrics: the ability bias.

Most people would agree that the education level a person achieves is correlated with something intrinsic, innate. How many times did we think that a person had a genius for a certain subject? Likewise, success in work is often attributed to the individual's innate capacities rather than to his/her education or other factors. Ability fulfills both the conditions to be an omitted variable. We expect the bias to be upward.

To make things clear, we are trying to understand: if college graduates stopped at the high school diploma, would they have likewise earned more money than the others high school graduates? Because, if it is so, then education accounts for less than what we estimate!

In this sense, the higher earnings observed for better-educated individuals could be a mere indication of their greater ability, instead of a result triggered by education itself. Under this hypothesis, for what concerns our regression, education level and ability would be 'too highly correlated', and including both in the regression could create a problem of collinearity: in this case the coefficients estimated may change unpredictably in response to small changes in data or in the model. The most extreme case is perfect collinearity[4], which is an exact (non-stochastic) linear relation among the regressors[5].

[4]Or multicollinearity, if one regressor is function of two or more regressors

[5]When we estimate coefficients we usually do this using the Ordinary Least Square method (OLS), that minimizes the sum of the squared vertical distances between the observed response in the

However, in order to have collinearity or not, first we should have a good proxy for ability. There have been many proposals, as the IQ level[6], and yet they are deemed to be imperfect and non-satisfying because ability is not just intelligence, but a complex mixture of factors (personality, particular aptitude, intuition…).

We could try to break the link between received education and ability conducting an experiment. We could randomize the education level: when kids start the primary school, we could randomly assign to them the educational level they will achieve, no matter their capacities and ambitions. But as you see, such experiment is probably illegal, surely unethical and, as most ideally useful experiments in economics of education, unfeasible.

Another way to get round the ability measurement problem is to use Instrumental Variables (IV). This instrument should mimic schooling – that is, be correlated with it – but be uncorrelated with ability and hence have no direct effect on earnings.

The solution seems at hand; however, it is not easy to find such instrument. Furthermore, the bias may be exacerbated if the instruments explain a small part of the variance of the endogenous regressor (we are in presence of a 'weak instrument'). Someone proposed the year of birth[7], because it is not correlated with ability, but it is correlated with the level of education: birth year captures the costs and benefits at the time a worker makes his schooling decision. Other proposed as an instrument the number of siblings, the parents' education level, the distance to university at the age of 14, the birthday, the occurrence of educational reforms[8]…

We said that, if more able individuals choose to acquire more schooling, the acquired education level is just the evidence of their capacities. Such evidence can be directed to other individuals in the labor market, a practice called 'signaling'[9].

The idea behind is that potential employees send a signal about their ability level (hence, their productivity) to the employer by acquiring certain education credentials. The employer values this signal because he assumes it to be correlated with a greater ability, or said differently, he assumes high ability employees to pay less for one year of education than low ability employees in terms of time and effort spent.

The firm bases its earning or hiring choices on the conditional expectation of productivity given by the potential employee's qualification. Potential employees with high ability find attractive to reduce the informational gap[10], and also employers are happy with this.

dataset and the response predicted by the linear approximation (the error term). In order to do so, we deal with matrices containing all variables and all observations. At a certain point in the procedure, we must invert this matrix. If there is collinearity, the rank of the matrix (number of linearly independent columns) is different from (smaller than, to be precise) the number of columns, and the matrix is not invertible, ergo, we cannot use the OLS method.

[6] [4]

[7] [24]

[8] See [11]

[9] [33]

[10] In the job market, information is asymmetric because while the individual may know his/her

In this framework, even if education does not contribute anything to an employee's productivity, it has still value both to the employer and to the employee, and employees will buy more education to signal their higher ability[11]. In this case, policy makers should focus on giving to everyone the opportunity to appropriately signal his/her own ability, to reveal his/her own capacities to the market, rather than forcing less able persons to complete degrees and diplomas – this would be a sort of market imperfection.

Some researchers have found evidence for a significant ability bias. Other studies, surprise surprise!, concluded that the ability bias had been overestimated, and that it was even smaller than other biases: as a result, the effect of education seemed to be underestimated (Becker 1964, Griliches 1977). Becker provides us with his witty interpretation of the importance usually attributed to the ability bias:

> *" A more cynical explanation would be that vocal observers are themselves primarily successful college graduates and, therefore, naturally biased toward the view that ability is a major cause of the high earnings received by college graduates."*

The debate has raged since the beginnings of economics of education and is still open to doubt. The scholars have found all kinds of results: downward, upward, small, large, no bias. What does this inconclusiveness suggest? Considerable caution, since we estimate the causal effect of education to back policy decisions.

The ability bias ghost still haunts most econometricians.

level of ability, usually the employer is not able to distinguish among high ability and low ability workers.

[11] Remark: if the world is rigidly credentialist, a comparison between earnings of one education level and the one immediately lower overstates the social benefits of education. The reason is that education is just making you better compared to other people, without actually increasing your productivity that is stable and innate. [8]

2. THE MARBLE BLOCK

2.1. The beginning: Mincer function.

The starting point for most studies in economics of education (and not only) is Mincer's HCEF, acronym for 'human capital earnings function', a true cornerstone for empirical economics. It made its own debut in "Schooling, Experience and Earnings", a 1974 work.

Jacob Mincer discovered a regular dependence of earnings on schooling and experience; a dependence that proved to stay quite stable over time and over space. Because of its solidity, this dependence was used widely by scholars, permitting comparability across studies.

According to the Mincer model, the logarithm of individual earnings in a given period of time can be decomposed into an additive function of a linear education term and a quadratic experience term:

$$logY = \alpha + \beta S + \gamma X + \delta X^2 + \varepsilon$$

where

Y: wage for a period of time.

S: number of completed years of education.

X: number of years the individual has worked since completing his education.

ε : an error term, the statistical residual.

It is often very difficult to have in the dataset the exact information on experience, and so it was for Mincer: indeed he proposed to use the 'potential experience', the number of years an individual of age A could have worked assuming he started school at 6 years and studied for S. From now on, we will call it "Mincer experience". Thus:

$$X = A - S - 6$$

Researchers such as Murphy and Welch (1990) added two higher-order terms in experience and faced a significant improvement in fit. The quadratic term is not flexible enough to capture some features of the data; in particular it understates the growth in wages for the first 10-15 years of experience. We will 'fine-tune' Mincer equation using a quartic function as well.

Even if incomplete, the Mincer model provides a good starting point to build more complex models about the causal effect of education, together with a useful benchmark. We will expand the Mincer model including a vector of socio-demographic characteristics as gender, geography, school type that will help us to eliminate biases and to get and a better vision of the problem.

2.2. Measurement of earnings

Earnings are almost always measured in logarithmic form. First of all, the distribution of earnings (especially if measured in hours) is incredibly close to a normal distribution. Second, the log transformation is convenient for interpretation, since it helps us to see the percentage effect of the independent variables on earnings.

Generally, the choice of the unit measure of time (hour, week, year) is subject to the data available in the dataset. Nevertheless it is important to recall that Mincer shows how some key results for our regression are sensitive to the choice of the earnings measure. For example, using annual earnings, Mincer rejects the hypothesis of linear education, but fails to do so using other measures. Card suggests that individuals with higher education level tend to work more. Under this scheme, the coefficient of education is higher for weekly and annual earnings than for hourly earnings.

Mincer uses hourly wages also because most workers in surveys report hourly wage rate, limiting measurement problems linked with defining hourly wages as earnings divided by hours.

In Italy most datasets, and ours does not make an exception, contain annual wages. And, most important, it contains net rather than gross wages, differently from the majority of other European countries, so that an international comparison would require a careful adjustments according to the tax level.

Unfortunately the complexity of the Italian taxation system inhibits a proper estimation of gross wages starting from the net ones. The main implication on our work is that a change in the causal effect of education on net wages over time may result as a consequence of a different tax pressure, and not of a different value/reward of education in the labor market.

2.3. Concavity - Additive separability

One of the factors that enabled the Mincer model to fit the data remarkably well in many contexts is the introduction of experience instead of age.

Earnings are a concave function of age: earnings rise at a decreasing rate throughout one's life. Look at Figure 1a, with age–earnings profiles, a graph of the earnings-age function with different lines, one for each different group of years of completed education.

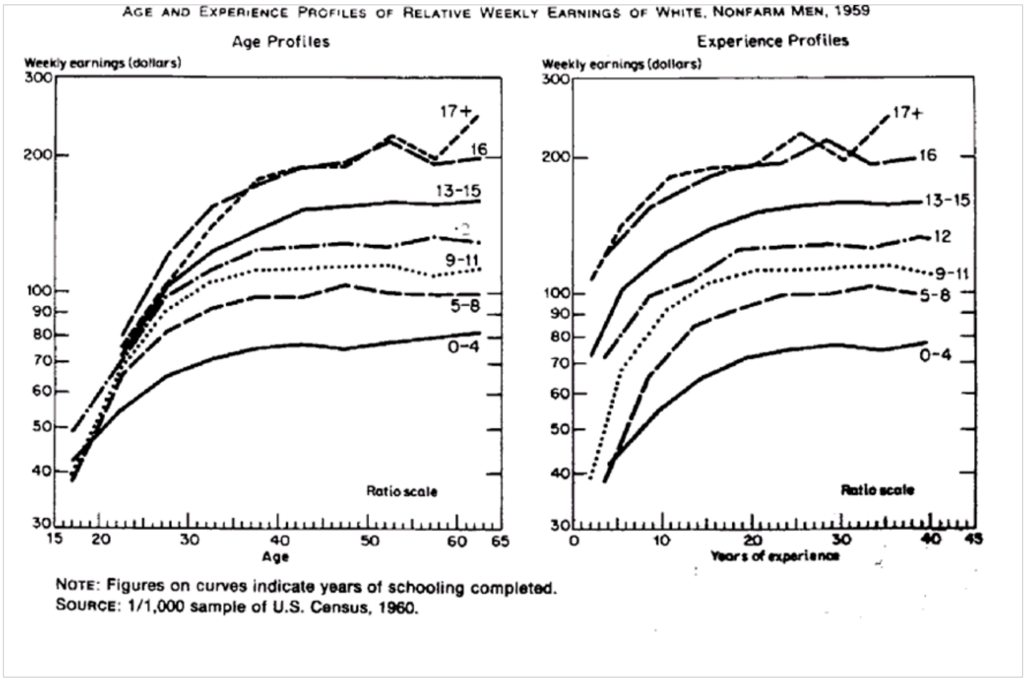

Figure 1[1] - Age and experience profiles of earnings

In his early works, Mincer noted that the age-earnings profile was steeper for more educated persons than for less educated ones, so that log earnings are not a strictly separable function of education and age[2]. In other words, the different lines in Figure 1a are not parallel. This means that it does not exist a single causal effect to education, but rather one different causal effect for each age group.

If we consider experience-earnings profiles, as in Figure 1b, these are relatively parallel for different education groups (even if using annual earnings this hypothesis is rejected). This was Mincer's intuition: introducing potential experience instead of age in the earnings equation, he was able to capture both the shape of the age-earnings profile and the differential slope of the age-earnings profile across education group. In this way, conditional on experience, there's a single effect of education in the labor market.

Are experience-earnings profiles as parallel for different education groups as they were in the 60s? Studying recent data, Card & Lemieux (2001), Heckman (2003)

[1]Source:[28]

[2]A function of two variables F(X,Y) is additively separable if it can be written ad f(x) + g(y) for some single-variable functions f(x) and g(y).

and again Lemieux (2003) concluded that it's not like that. In particular, the university-high school wage gaps are much larger for less experienced than more experienced workers.

2.4. How do we measure education?

In the Mincer model the logarithm of earnings is a linear function of years of completed education. This specification embeds two related hypothesis:

1. the correct measure of education is the number of years of completed education.
2. each additional year of schooling has the same proportional effect of earnings, other things being equal.

However the linearity assumption, Mincer (1997) suggests, could be incorrect in a Becker type model, where individuals are heterogeneous in their preferences and earnings opportunities; in this case, the log earnings could be a convex or a concave function of years of schooling[3].

Despite some studies, the issue of linearity or not remains. Card and Krueger (1992) and Heckman (1996) reject the hypothesis of non-linearity for most levels of education. Mincer (1997) and Deschenes (2001) showed instead that log earnings have become an increasingly convex function of years of education since 1980.

Some studies found that this departure from the standard Mincer model is mainly a consequence of a movement from the equilibrium in the already-quoted Becker model: an abrupt growth in the demand for more educated workers has not yet been balanced by the increase in their supply. This results in a more convex education-earnings relationship since the causal effect of schooling is just the slope of the education-earning relationship.

Probably we are no longer at equilibrium after the dramatic changes in the wage structure that happened in the late 1970s: this distance from the equilibrium quantity of better-educated workers caused both the convexity of the education-earnings relationship and the increase in the causal effect of education.

It would be interesting to see if in the next future the Mincer model will go back to its original good fit of data, now that a stabilization in the wage structure seems to be close in time.

Linearity seems to work very well in stable labor environments except for the 15 and 16 years of schooling, where it respectively over-predicts and under-predicts log earnings. There could be a simple reason behind this imperfection of the model: there exists a wage premium for fulfilling the final years of a school level and getting a credential. This effect is known as the "sheepskin effect" (effetto pergamena) or "credential effect".

[3] We will analyze Becker model later in details.

Some scholars started to introduce dummy variables to capture the non-linearity of the sheepskin effect (Solon and Hungerford 1987, Belman and Heywood 1991), and so did Card & Krueger. In their study, the only credential effect reported is the university one, but surprisingly enough there is no jump in log earnings for the high school completion.

Chiswick (1973) tries to reconcile the idea of ability and productivity in the education system, also providing us with a possible explanation for the solitary sheepskin effect on the university completion: since graduates are on average the most efficient learners, they enjoy proportionately a larger increase in productivity that their years of education alone would indicate.

In general economists have sticked with the linear Mincer functional form. Despite this approval of the traditional measure of schooling, in the late 80s the US Census Bureau shifted toward a degree-based system of measuring post-high school education: under this system, people were not asked how many years they spent in college but rather if they had or not a college degree.

Operating with standard schooling years[4], we are not considering neither the years that students who fail have to repeat nor the ones completed without achieving a qualification. On the other side some students need fewer years to complete a school level, or they start primary school at 5 and not at 6; nonetheless we expect the proxy to underestimate the actual number of years.

Our dataset applies the same degree-based system, but this seems acceptable: while the use of years of schooling as a measure for education has a long history in the USA, this is less natural - and less valid - in countries such as Italy, where we expect a strong sheepskin effect.

2.5. Sheepskin effect

As we said, the sheepskin effect is the wage increase above what would normally be attributed to the completion of an extra year of education, and independent from it. This effect can be estimated empirically studying the earnings differences between 'drop-outs' and 'completers' with an equal number of years of education.

The sheepskin effect has been found (Hungerford & Solon 1987), and it was considered as an evidence of signaling in the labor market. The diploma/degree in fact has a value independent from the accumulated years of education – and in addition to the role that schooling plays in making workers more productive – it also indicates workers as more productive.

[4]Standard years are those normally required to complete one level of education. In Italy the standard years variable take value 0 (no education), 5 (primary school – scuola primaria), 8 (middle school – scuola media), 11 (vocational school - diploma professionale),13 (high school – scuola superiore), 16 (bachelor degree – laurea triennale) , 18 (master degree – laurea magistrale), 21 (PhD – dottorato).

It may be interesting to know that, according to some studies (Golbe, 1985), in imperfect signaling model (the ones with statistical discrimination), minorities receive greater return to signals of high productivity than do white males. Minorities in fact face higher costs of achieving a certain degree/diploma than other people (because of different resources available to the different demographic groups) so they achieve high levels of education only if they are unusually productive.

3. EXPANDING THE MINCER MODEL

3.1. Heterogeneity

So far, we did not highlight much that “Mincer equation” is just a nickname for the Human Capital Earnings Function. Human capital is defined as the stock of knowledge, experiences and competences accrued by an individual during his/her life, resulting in the creation of economic value. One way to acquire human capital is exactly through education.

We carefully determined why our β coefficient is not the return to education. Mainly, we are not taking into consideration all benefits and costs arising from education: we assume just wage benefits and zero costs. However, we cannot ignore in our regression that people are not the same – they have not the same background, possibilities, opportunities, abilities.

Said in economics terms, each individual faces a market opportunity locus that gives the level of earnings associate with alternative schooling choices. This is the idea behind the already-cited Becker model (1976), followed by Card (1995a). Consider two functions:

y(S) is the level of annual earnings an individual will receive if he/she acquires a schooling level S[1]. Since it is assumed that individuals with more human capital earn higher wages, and education induces individuals to accumulate it, y(S) is an increasing function. Moreover it is a concave function because, as is true for all production factors, the increase is less than proportional. In more simplified models, y(S) is linear.

h(S) is the cost (direct and indirect, monetary and non monetary) an individual will pay if he/she acquires a schooling level S. It is an increasing convex function, since we assume that the marginal cost of each additional year increases (one year at university costs more than one year at the high school and so on).

An individual i will choose S in order to maximize the utility function U(S, Y(S)):

$$U(S, y(S)) = log(y(S)) - h(S)$$

[1] As Card (1999) remarks, this higher wage could be due to productivity effects of education and/or other forces such as signaling.

choosing S* such that

$$\frac{y'(S)}{y(S)} = h'(S)$$

so that marginal costs are equal to marginal returns – both associated with an additional year of schooling.

In order for the model to be operational, we assume both sides of the equation to be linear functions with individual-specific intercepts and homogeneous slopes[2]:

$$\frac{y'(S)}{y(S)} = b_i - k_1 S$$

$$h'(S) = r_i + k_2 S$$

with r_i and b_i random variables and k_1 and k_2 non-negative constants. Thus the optimal schooling choice S* is linear with respect to the individual-specific heterogeneity variables.

$$S_i* = \frac{b_i - r_i}{k_1 + k_2}$$

Here, the heterogeneity in S* may come from

1. differences in the access to funds (family wealth) or families' tastes for schooling , r_i.
2. differences in ability b_i, as we have already seen with the ability bias.

The characterization of marginal returns and costs imply specific functional forms for both individual functions log(Y(S)) and h(S):

$$log(y(S)) = a + b_i S - \frac{1}{2} k_1 S^2$$

$$h(S) = c + r_i S + \frac{1}{2} k_2 S^2$$

According to this specification, ability affects the slope, and not the intercept of the earnings function.

Here a problem arises: if individuals differ in their relative earnings capacity, if they know their $b_i$and base their education level decision upon their comparative

[2]In case we have a linear function y(S): $\frac{y'(S)}{y(S)} = b_i$

advantage, there may be a 'return bias' whose direction is not clear. In other words, individuals' returns are correlated with the decision of which level of education to achieve; the individual self-select into different educational qualifications according to his/her proper return, which depend on characteristics observable to the individual but not to the researcher.

However individuals do not necessarily know the parameters of their earning function when they make their school choices. The b_i is thus the best estimate of what he/she expect to gain per year of education, and we must also consider that its distribution may change over time because of shifts in labor market conditions, technologies, etc.

In equilibrium, a greater supply of highly-educated workers will lower the mean of b_i and may also affect other characteristics of its distribution.

3.2. Rationality

In spite of the previous economic calculations, the decision to invest in education could be not completely ruled by rationality, meaning, aimed at maximizing the utility function by equating marginal costs and marginal returns[3]. Education has a cultural, social and historical value far beyond economic benefits, and individuals may include this value in their educational decisions.

Furthermore, early educational investment decisions are usually taken not by the individual him/herself, but rather by other agents such as the parents. We may assume that not only the level, but 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 (encouraging a certain career over another, later on)[4].

Parents could push children towards the straightest way for them to find a job – for example, introducing them to their same profession (consider lawyers, doctors, etc.): in this case, it is cheaper for the children of a lawyer to become lawyers because they would not incur initial investment.

An education commitment can be encouraged also by other agents, such as the individuals' relatives, friends, acquaintances, neighborhoods: in other words, by the social environment. We call it 'peer effect'. Having high-educated friends and relatives may push the individual to go on with schooling and vice versa.

[3]

> *"Rational behavior may pursue, besides instrumental ends such as utility, profit and wealth, social ends as well. As a consequence, action is rational if it is aimed 'not only at economic goals but also at sociability, approval, status and powers' (Granovetter, 1985) [...] Social status and economic power perform a valuable role"* [6]

[4]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.

However parent's decisions, undertaken on the basis of their values and thinking, may reveal to be obsolete for the current system, and the same holds for peers' suggestions/influences. Such decisions create a conflict between individual vs. collective interest, thus originating a mismatch between the kind of education needed by employers and the education supplied by the potential employees. In this scenario, adjustments to the labor market are slowed down.

As a final point, we recall the main idea of behavioral economics: it is nearly impossible to gather all the information needed to take an informed decision. Our rationality is limited by the complexity of the world, the inconsistencies of individual preferences and belief, and other factors... 'Bounded rationality' suggests that even if it is the individual him/herself who decides, without being influenced by the social environment, still he/she does not know the exact value of his β, the effect of education on earnings; furthermore, this β may vary in time.

Conclusion: observations and studies report an increase in the level of education without an increase in the wage premium and employment, and vice versa.

3.3. Control variables

Heterogeneity in costs and returns limits our prediction capacity to draw general results. In order to reduce this problem, we must account for this heterogeneity in our regression. We already spoke about the ability bias (heterogeneity in returns) and how to deal with it.

For what concerns heterogeneity in costs, even if we do not estimate them directly, we must take them into considerations. A good way to do that is to introduce some variables accounting for the role of the family or of the region, for example.

We already spoke about omitted variables: we will use them to control for and mitigate heterogeneity.

3.3.1. Gender

Many studies avoid considering women in their dataset because usually women choose to enter the labor market only if they have an adequate return: so their returns are likely to be overestimated, if one does not consider the large part of the female workforce that stays at home.

We cannot exclude that gender is correlated both with education and earnings: reasons can be a lot. Let's take just some examples. The percentage of girls in 'licei' is quite high with respect to the one in technical and professional schools; this is probably due to cultural constraints. Now, students enrolled in a liceo have a much greater probability to go on with schooling, since they have not the preparation needed to enter the job market immediately after the achievement of the high school diploma. Alternatively, parents could be more protective with their daughters and let them achieve a high level of education, preventing them to enter the job market too early.

When we look at earnings, it could be that women are unlikely to get an upgrade and fill responsible positions because they could face a pregnancy or because they are more reluctant than males to accept transfers and stay away from home for long.

Also looking at the results of this and many other studies, controlling for gender seems to be crucial for a good analysis.

3.3.2. Region

As we have seen dealing with omitted variables, the place in which an individual studies and works could be correlated with the quality of school – and Card and Krueger 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.

In our regression we insert two dummy variables, one for the Centre of Italy and one for the South.

3.3.3. Race / Citizenship

Minorities usually face higher or steeper marginal costs for education. They could experience less or worse education because of poverty and segregation and being victims of discrimination in the job market.

The problem is that in our dataset we have no observation for race. We have just a dummy variable concerning the citizenship of the interviewee: but race and citizenship are not the same thing.

However, even if it is true that persons without Italian citizenship of course may come from rich countries, we assume that the most part is made of immigrants fleeing from the poverty of their countries.

Indeed the effect of the race variable will be underestimated by the citizenship proxy because some people who are not "Caucasian" have the Italian citizenship, even if their marginal costs stay the same.

3.4. Different types of education

The accumulation of human capital is not necessary a smooth, linear and almost continuous process and, especially, the effect of education needs not to be same across different types of schools. The demand for labor, in fact, is not the same for each degree/diploma.

According to the results on the 1995 Italian SHIW dataset , obtained by Lucifora, Comi and Brunello, we remark that:

1. Medicine, that requires more years of studies, is the degree that holds the higher returns.
2. Traditionally male dominated degrees, such as engineering, yield to females the highest returns, but the reverse does not appear to be true.

In the second part of the paper we will replicate these estimations for our datasets, observe if the results are consistent with the 1995 ones and how they evolved in time.

4. REGRESSION

The choice of the regression to use for the analysis of data is a function of two things. First, some models may better fit the data, that is the assumptions on their distribution. Second, one regression better than another may allow you to focus on a specific factor or characteristics you want to inquire on.

The OLS estimation is based on some particular assumptions that may not always be satisfying for our purposes, or it results in some outcomes that we would like to explore in depth – and in order to do this we must look for other tools. This section provides a little overview about the two types of regression we will use in the second part: the weighted least square and the quantile regression.

4.1. Weighted least square regression

The OLS method presupposes that each observation conveys equally precise information or, said differently, the standard deviation of the error term is constant over all the estimated values of the coefficients.

Of course this is hardly the case. If it is not reasonable to treat each observation equally, the weighted least square (WLS) regression can be a way to maximize the efficiency of the coefficient estimation. In order to do this, one of the variables in the dataset must be the 'weight', meaning the proper amount of influence of each observation. The higher the weight, the higher the importance/precision/reliability of the information contained in the related observation. In our case the reliability information comes from a particular section of the survey, in which the interviewer is asked to judge and give grades to the interviewee's comprehension of the questions and the truth and easiness of his/her answers. All the members of the same family have the same weight.

Note that the WLS regression does not concern a particular function to describe the relationship between the independent variables and the dependent one; we can use weights both for linear and non linear regressions. The WLS regression just incorporates the weights into the fitting criterion.

The problem is that the weights are assumed to be known exactly, and this is really unlikely. Estimated weight must be used instead. However experience indicates that small variations in the weights do not often badly affect the regression analysis results.

A second concern is that also WLS, as any other least squares method, is sensitive to the presence of outliers; but if the WLS increases the influence of an outlier with a wrong weight, the results stemming from it will be much worse than those of an non-weighted least square regression.

Nevertheless there's a type of regression that is more robust to outliers: the quantile regression. Combining the WLS and the quantile regression we could find more accurate results.

4.2. Quantile regression[1]

The coefficient we estimate using the OLS method is the effect of a change in the regressor on the conditional mean value of the response variable. This approach is used under the implicit assumption that the possible differences in the impact along the conditional distribution are negligible, or unimportant for that research agenda.

On the contrary, the quantile regression looks at the impact of the explanatory variable at quantiles[2] of the conditional distribution[3] of the response variable other than the mean. The snapshots at different points of the distribution allow a full characterization of the conditional distribution of the dependent variable, resulting in a more flexible, comprehensive and robust analysis.

Mathematically, while the OLS method can be reduced to a problem of numerical linear algebra, QR is a problem of linear programming[4] that can be solved using the 'simplex method'. Standard errors and confidence intervals are obtained using the bootstrap method[5].

In our case, β – the effect of one year/level of education more, ceteris paribus – on the tails of the earnings distribution, meaning on the highest or lower wages, could be significantly different from the β on the mean, estimated using the OLS method. Said differently, the quantile regression allows comparing how a certain level of education may affect more some quantiles[6] of earnings than others.

[1] The quantile regression method was introduced for the first time by Roger Koenker and Gilbert Bassett in their paper "Regression Quantiles", 1978.

[2] Recall that a quantile q is a point in the cdf of a variable that divides it in two parts proportional to and 1- . To make things clear, $q_{0,25}$is the point dividing the cdf of a certain variable such that the 25% of values is smaller than that.

[3] The cumulative distribution function (cdf) describes the probability that a random variable X will be found at a value less than or equal to a certain value x.

[4] A linear programming problem consists in maximizing or minimizing a linear function defined on the set of solutions of another system of non-linear inequalities, called boundaries.

[5] Bootstrapping is a resampling practice used to estimate the empirical distribution and thus the properties of an estimator (to approximate its mean and variance, to build confidence intervals and to calculated p-values).

[6] There's a common error dealing with quantiles when a value of X is said to be in a certain quantile. As we said, the quantile is a point, not an interval. However, we interpret a value "in the kth quartile" as being between the kth-1 and kth quantile .

This potential difference may prove to be crucial in particular analysis, such as the ones investigating wage inequality. If β is actually not constant over the earning distribution, policies designed to increase the poorest wages simply investing in the attainment of higher education may not be as effective as one thinks having in mind the OLS results. There may be other factors to investigate and to invest on, obtaining more vigorous results.

To make things concrete, consider this data from a 2001 paper by Martins and Pereira: the OLS β for Portugal is 12.6%, but considering the first and the ninth deciles[7] the β are 6.7% and 15.6%. It means that more education reward especially people already gaining high earnings. This result is confirmed for all European countries considered, except for Greece[8]. Reasons for this non-constant effect of education will be explored in the second part of this paper.

Using QR it is possible to check whether the conditional distribution of earning simply shifts when education changes (difference in location) or whether there is also a scale effect, in a way that the entire shape of the distribution changes (difference in dispersion).

Another reason to prefer QR to OLS is linked to a technical peculiarity of the wage equation. Since the lowest wages are probably self-excluded from the interviews, the available dataset is likely to be truncated. Such truncation at the lower tail of the distribution increases the mean of the dependent variable, which in turn causes a bias in the OLS estimates since OLS computes the line/plane passing through the conditional mean. The QR is not affected by truncation since quantiles are more robust than the mean.

[7] 10% and 90%

[8] Pereira and Martins explained this exception with the fact that for some countries (Austria, Greece and Italy) only net wages were available. And progressive taxes are likely to have a stronger impact in eroding returns to education at the top of the distribution than at its bottom.

Part II.

GET ONE'S HANDS DIRTY

5. Sample description

The data we are going to analyze are those from the Survey of Household Income and Wealth (SHIW), a representative sample of the Italian resident population usually held each two years. We consider the last eight waves: 1993, 1995, 1998, 2000, 2002, 2004, 2006 and 2008[1]. Our dataset will be restricted to employees aging between 25 and 65.

In this nearly twenty-year period, both the data provided by the survey and the external framework have changed. Vocational school and bachelor degree are not present in the 1993 dataset. The “diploma universitario”, lasting 3 years, was introduced in 1995 and was replaced by the 1999 Berlinguer’s educational reform with the bachelor degree, “laurea triennale”.

Here there are some statistics to sum up the characteristics of our samples.

As you see in the first table, the percentage of women in the sample increased regularly.

Table 1 - Gender distribution

Gender (%)	2008	2006	2004	2002	2000	1998	1995	1993
Male	57.19	57.10	58.07	58.56	59.89	60.40	62.40	64.31
Female	43.89	42.90	41.93	41.44	40.11	39.60	37.60	35.69
	5331	5310	5333	5135	5417	4889	5184	5138

Data on the citizenship of the interwievee are available only in the last two datasets.

Table 2 - Citizenship distribution

	2008	2006
No citizenship (%)	6.30	4.07

Table 3 - Geographic area distribution

Area (%)	2008	2006	2004	2002	2000	1998	1995	1993
North	50.87	51.56	50.07	50.75	49.61	46.19	48.82	45.85
Center	20.13	20.34	22.67	21.52	21.69	22.11	21.12	23.37
South	29.00	28.10	27.26	27.73	28.70	31.70	30.05	30.77

[1] Data on 2010 are not yet available.

For what concerns the educational level, as you see year 1993 does not contain neither the vocational school diploma nor the bachelor degree. In general we can spot a sort of symmetric behavior around the middle school, whose numbers stay constant over time: less people achieved the lowest levels of education, a higher percentage enjoyed higher education. Particularly important is the role of vocational school: it seems to have filled the gap between the middle school and the high school diploma, and its numbers increase over time. Finally, the PhD experienced a jump in the last dataset.

Table 4 - Educational level distribution

	2008	2006	2004	2002	2000	1998	1995	1993
No education	0.36	0.24	0.56	0.80	0.78	0.84	1.04	1.15
Elementary school diploma	5.38	5.59	7.26	8.55	8.55	9.67	12.94	14.83
Middle school diploma	31.42	31.88	31.61	31.04	30.07	28.62	29.01	32.43
Vocational school diploma	10.15	9.34	8.16	8.49	8.44	7.10	6.64	NA
High school diploma	37.03	37.93	38.01	37.33	36.62	38.37	36.28	38.38
Bachelor degree	1.35	1.41	1.28	1.21	1.47	1.15	0.93	NA
Master degree	13.28	13.30	12.77	12.37	13.90	14.03	12.96	12.81
PhD	1.03	0.30	0.36	0.21	0.17	0.22	0.19	0.41

6. Results

6.1. Education

Finally, here are the results we obtain using the weighted least square regression. This is the model we built, regressing against the logarithm of hourly wages:

$$ln(Y_h) = \alpha + \beta_p PrimaryS. + \beta_m * MiddleS. + \beta_v * VocationalS. + \beta_h * HighS + \beta_b * Bachelor + \beta_u * Master + \beta_p * PhD + \gamma * Female + \delta * Center + \lambda * South + \pi MincerExp + \pi_2 MincerExp^2 + \pi_3 MincerExp^3 + \pi_4 MincerExp^4$$

Please remember that:

- coefficients are conditional. We are not considering all the graduates/holders of a diploma population, but just those who are working, and working as employees, not self-employed persons.
- we have no proxy for ability, thus the estimated effects are more predictive rather than causal. In other terms, the coefficients do not represent the causal effect of education on earnings, but rather an instrument to predict what is the likely increase in wages.
- changes in the effect of education in reality could be due to changes in the tax pressure: a general tax relaxation could have increased the predicted effect of education while the actual effect stayed unchanged and vice versa for high taxation. Even more complex patterns are to be considered if we think that a tax policy may affect more some levels of wage rather than others.
- people in the survey are asked to report the net wage written in the pay packet, so usually off-the-book workers do not declare their earnings for the fear that their responses are checked by authorities. This results in an incorrect estimation of actual wages.

Having these limits in mind, the following graph helps us in pointing out some features of the effect of education on earnings. Recall that the numbers on the y-axis are the percentage increase resulting by the achievement of a given education level with respect to the wage of a person without education.

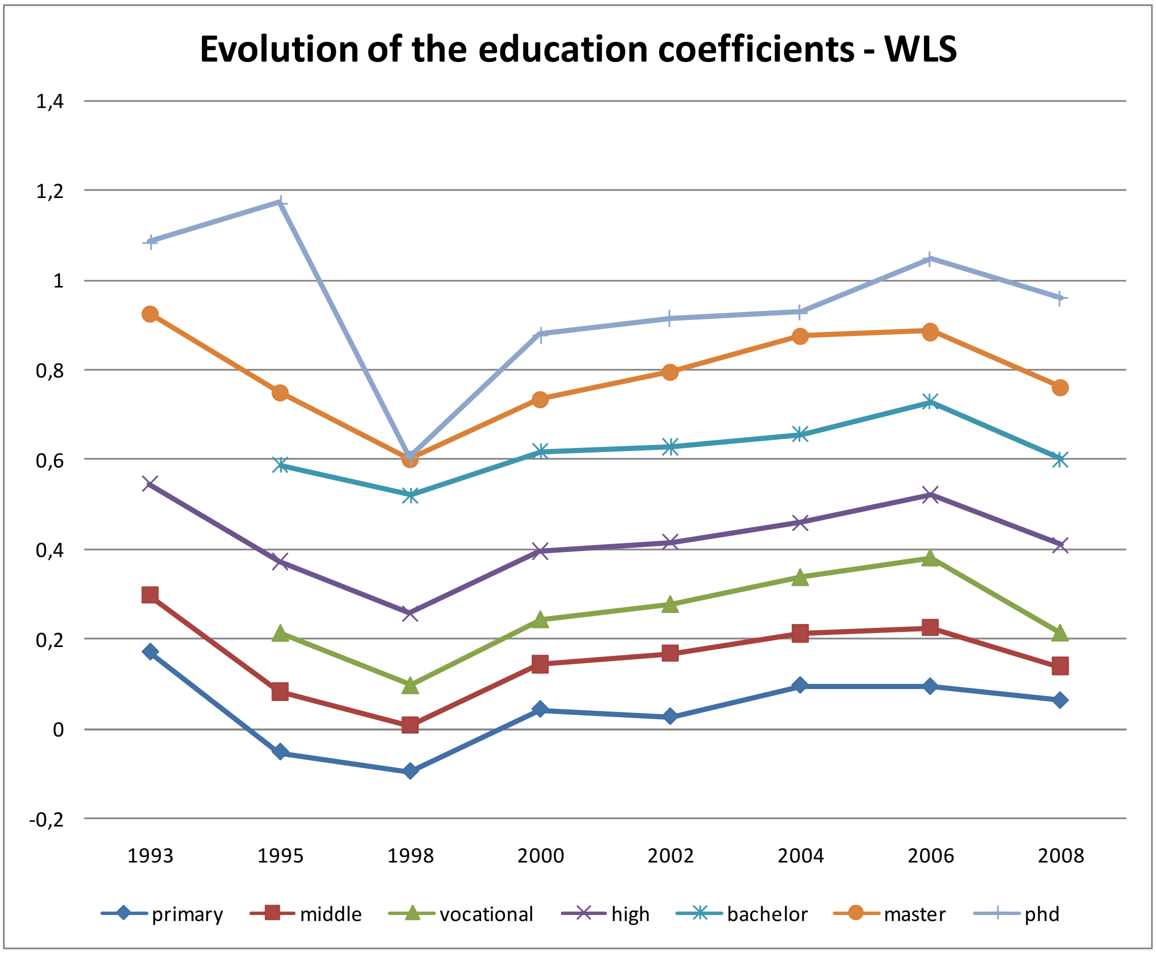

Figure 2

First, as you can see, the lines stay almost parallel over time. Schooling effects for each level of education move together and suggest that they all depend by some external variable, to be looked for in the economic scenario or in the labor market.

Second, the coefficients decrease until 1998, increase until 2006 and now they are decreasing again, quite uniformly with an average drop of 10.5%, going back to the situation of 2002. Having longer period results would allow us to see if this is a recurrent path (a sort of sinusoid moving up and down) or if the trend is different.

To this regard, consider Figure 3. It is taken from the work by Brunello, Comi and Lucifora (2000). They consider years of education instead of level of education, but we think that results are overall parallel. Notice that in the last part of the graph the schooling effect is decreasing as it is in our analysis. The trend suggests non-stationarity, with the average effect of education increasing over the decades. If this is correct, data from the 2010 survey (and those later) should give as result a follow-up decrease, but not such that we go back to the 1998 situation.

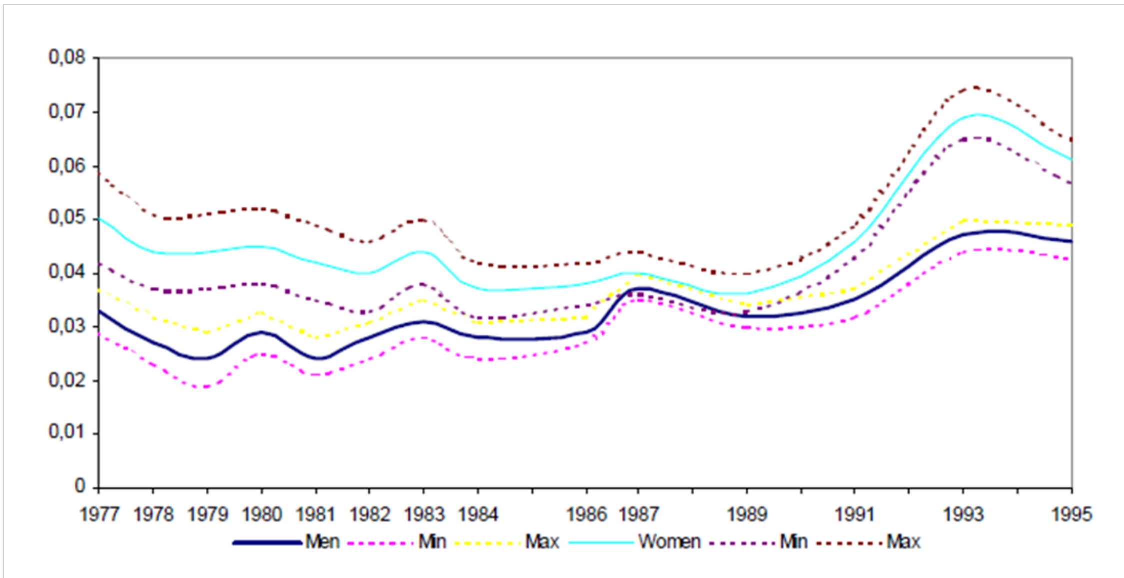

Figure 3 - Evolution of the education coefficient in Brunello, Lucifora and Comi (2000)

Third result, the PhD pattern is very irregular. In 1998 and in 2004 it was not statistically different from the master and the bachelor degree: the t-test[1] is 0.97 in the first case and 0.47 in the second. This could be due to the small amount of observations for PhD; otherwise PhD in Italy would seem to be still a risky investment.

Fourth, the biggest gap is between the high school diploma and the bachelor degree (an average difference around 20%), suggesting the presence of the sheepskin effect for graduates in the Italian job market. If we consider instead the average gap between two contiguous levels of education and divide it by the year between the two, we find a measure for the wage improvement per year of studying. The highest improvement is the one of the 2 years to complete high school after the vocational school diploma and the other 2 years to achieve a master degree (both around 7.5%). At the third place the bachelor degree years bears around 7% return each. PhD completion returns around 5% per year of studying, very similar to the return for each year of high school considering a non-vocational school (that does not give to the student any intermediate diploma, thus completed in at least 5 years).

The elementary school coefficient is statistically different from 0 only in 1993, probably due to a deterioration of the marginal importance of such low education. The middle school 'pattern of significance' is much more ambiguous.

Here are the tables with all the results. For each year, the first column in the estimated coefficient, the second one its standard error and the third one the level of

[1]The t-test is a statistical hypothesis test that assesses the probability for the means of two different groups to be statistically equal.

significance: * for 95% level and ** for 99% level. The "0" coefficients and standard errors for the degrees of the potential experience are actually very small numbers, below our treshold of significance.

Table 5a - Results of the WLS regression

		2008			2006			2004			2002	
	β	S.E.	CI	β	S.E.	CI	β	S.E.	CI	β	S.E.	CI
elementary	**0,062**	0,079		**0,094**	0,103		**0,096**	0,069		**0,027**	0,062	
middle	**0,139**	0,078		**0,227**	0,103	*	**0,212**	0,070	**	**0,167**	0,062	**
vocational	**0,214**	0,079		**0,380**	0,104	**	**0,336**	0,072	**	**0,277**	0,064	**
high	**0,410**	0,078	**	**0,521**	0,103	**	**0,460**	0,071	**	**0,416**	0,062	**
bachelor	**0,600**	0,086	**	**0,728**	0,109	**	**0,656**	0,080	**	**0,627**	0,074	**
master	**0,761**	0,079	**	**0,885**	0,104	**	**0,875**	0,071	**	**0,795**	0,063	**
phd	**0,959**	0,092	**	**1,048**	0,131	**	**0,928**	0,114	**	**0,915**	0,144	**
mincer exp	**0,049**	0,011	**	**0,054**	0,011	**	**0,035**	0,010	**	**0,061**	0,012	**
mincer exp^2	**-0,001**	0,001		**-0,001**	0,001		**0,000**	0,001		**-0,002**	0,001	**
mincer exp^3	**0,000**	0,000		**0,000**	0,000		**0,000**	0,000		**0,000**	0,000	*
mincer exp^4	**0,000**	0,000		**0,000**	0,000		**0,000**	0,000		**0,000**	0,000	
central area	**0,013**	0,012		**0,013**	0,012		**-0,007**	0,011		**-0,046**	0,012	**
south area	**-0,077**	0,011	**	**-0,072**	0,011	**	**-0,073**	0,011	**	**-0,026**	0,012	*
female	**-0,110**	0,009	**	**-0,109**	0,009	**	**-0,095**	0,009	**	**-0,091**	0,010	**
intercept (α)	**1,377**	0,094	**	**1,223**	0,117	**	**1,328**	0,085	**	**1,181**	0,083	**

Table 5b - Results of the WLS regression

		2000			1998			1995			1993	
	β	S.E.	CI	β	S.E.	CI	β	S.E.	CI	β	S.E.	CI
elementary	**0,043**	0,047		**-0,096**	0,057		**-0,052**	0,051		**0,171**	0,056	**
middle	**0,144**	0,047	**	**0,008**	0,058		**0,082**	0,051		**0,298**	0,057	**
vocational	**0,243**	0,049	**	**0,097**	0,060		**0,213**	0,053	**	**NA**	NA	/
high	**0,396**	0,048	**	**0,258**	0,058	**	**0,374**	0,051	**	**0,545**	0,057	**
bachelor	**0,618**	0,058	**	**0,520**	0,072	**	**0,587**	0,070	**	**NA**	NA	/
master	**0,734**	0,048	**	**0,601**	0,059	**	**0,750**	0,052	**	**0,924**	0,058	**
phd	**0,877**	0,139	**	**0,605**	0,123	**	**1,172**	0,100	**	**1,085**	0,099	**
mincer exp	**0,065**	0,010	**	**0,072**	0,011	**	**0,052**	0,011	**	**0,050**	0,012	**
mincer exp^2	**-0,003**	0,001	**	**-0,003**	0,001	**	**-0,002**	0,001	*	**-0,002**	0,001	*
mincer exp^3	**0,000**	0,000	**	**0,000**	0,000	**	**0,000**	0,000		**0,000**	0,000	
mincer exp^4	**0,000**	0,000	**	**0,000**	0,000	**	**0,000**	0,000		**0,000**	0,000	
central area	**-0,051**	0,010	**	**-0,032**	0,012	**	**-0,022**	0,011	*	**0,003**	0,012	
south area	**-0,053**	0,010	**	**-0,013**	0,011		**-0,001**	0,010		**0,035**	0,011	**
female	**-0,074**	0,008	**	**-0,062**	0,009	**	**-0,054**	0,009	**	**-0,065**	0,010	**
intercept	**1,148**	0,069	**	**1,217**	0,080	**	**1,099**	0,075	**	**0,972**	0,081	**

6.2. Control variables results

6.2.1. Experience

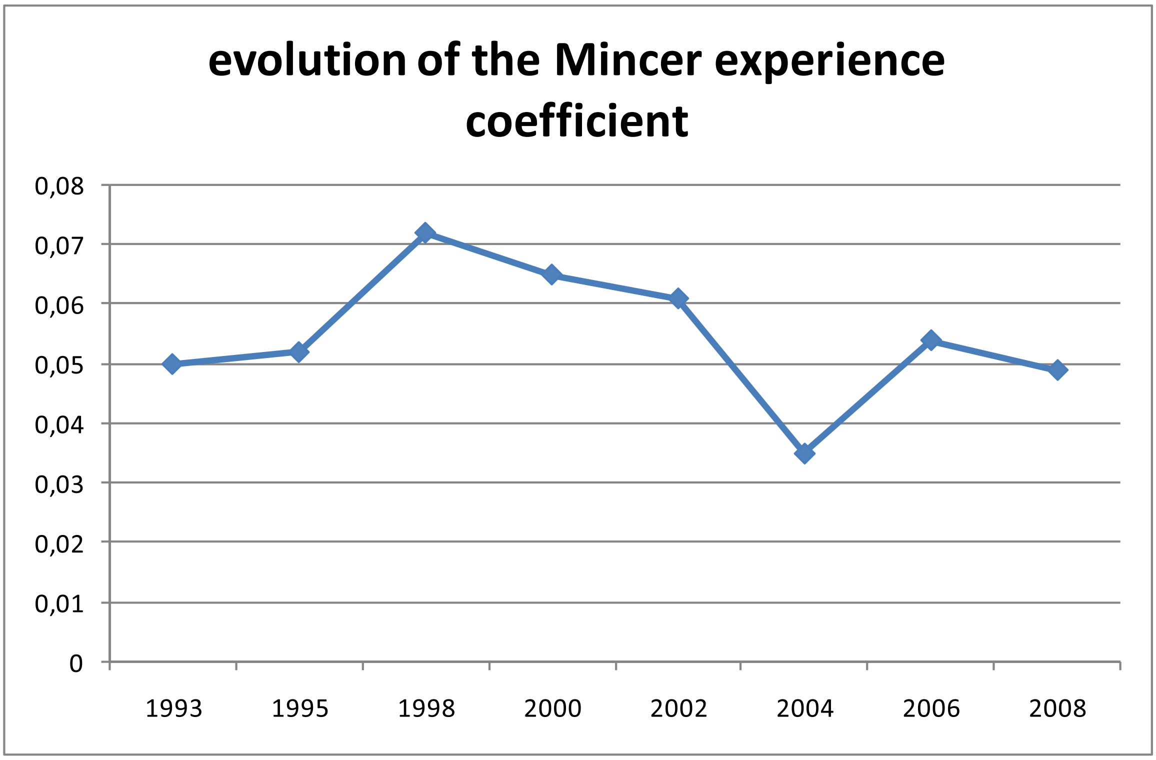

Figure 3

The Mincer experience coefficient is nearly a mirror image of the education pattern. Basically when the education coefficients decrease (from 1993 to 1998), the importance of experience increases, followed by an opposite behavior from 1998 to 2006, when the experience importance decreases and education one increases.

6.2.2. Gender

The coefficient for gender is always highly significant and negative across all years. The interesting feature is that it is increasingly negative as approaching 2008, getting to – 11% in 2008.

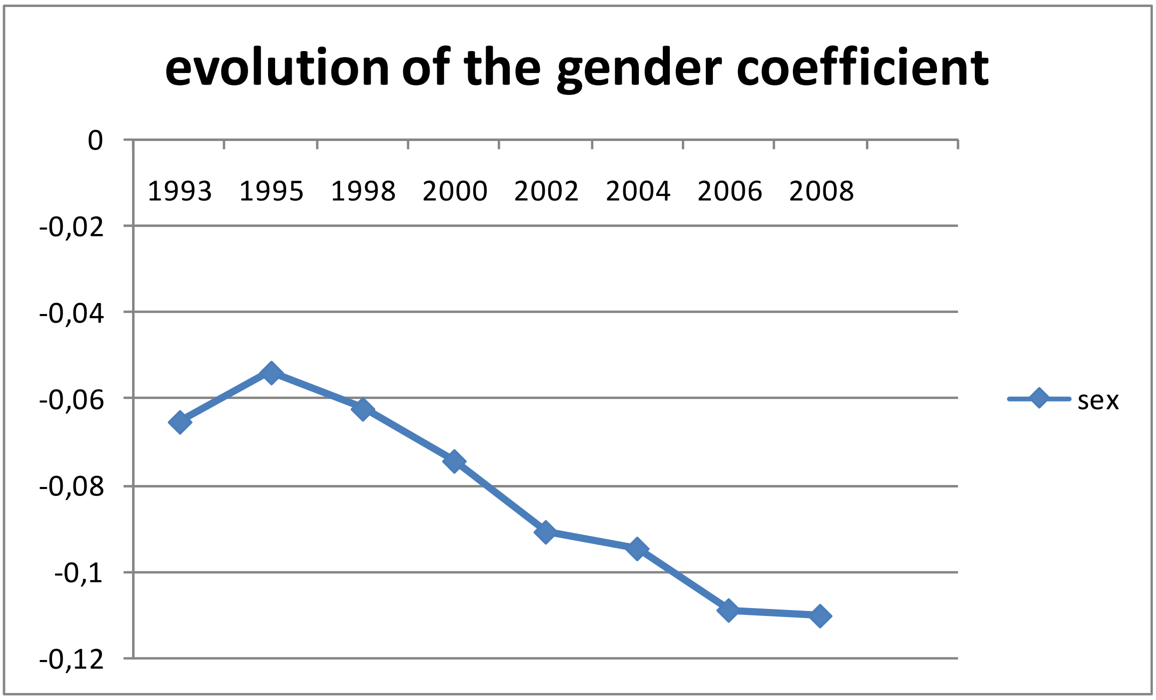

Figure 4

One possible explanation has already been taken into consideration in the previous sections and lies in the fact that in early years women entered the labor market only if they had a sufficient compensation – otherwise they would have decided to stay home and take care of their children, while the burden of working was left to men. Because of this sample selection bias, many analyses do not consider female workers.

This tendency may have decreased over time, due to a more severe struggle for gender parity: as a result, now women may look for job without caring about their partners' decisions, looking at their personal fulfillment, requiring a smaller return than before to start working. A partial proof of this may lay in the fact that the number of employed woman regularly increases in the sample (from 35.69% in 1993 to 43.89%)[2].

However, this could explain only why the coefficient is decreasing steadily, not why it is negative. Is this an evidence of discrimination? Or do women choose less rewarding jobs than males do because of any reason?

6.2.3. Region

In five datasets out of eight the south coefficient is significant and in four of them it is negative. The most recent datasets shows that, *ceteris paribus*, an employee working in the South of Italy earns about 7% less than one working in the North.

[2] And also the total number of observations increases.

The coefficient for the centre of Italy keeps oscillating between the significance and the insignificance. When it gains significance, its value is negative, between -2 and -5%.

6.2.4. Citizenship

The lack of the Italian citizenship highly penalizes employees. The percentage expected decrease in wage for a non Italian citizen is around 9% in 2006 and 17% in 2008.

6.3. Types of education

In order to compare our results for different types of education with those of 1995 found by Brunello, Comi and Lucifora, and see if they are consistent, we slightly changed our regression.

In particular, Brunello, Comi and Lucifora do not include variables about primary and middle school: this means that the coefficients measure the percentage change in earnings with respect to a person that stays under the soil of the vocational school diploma. Please take care of it while confronting the results in this section with the others previously reported.

Second, they divided the results according to the gender variable.

Consider also that their regression included age and age squared and not experience, and no dummies for the area in which one lives.

Because of the small number of observations for some variables in the 2008 dataset, we decided to join in a single sample data from 2008 and 2006. You find the results in the table. The first column of each year contains the coefficients and the second column the standard errors.

Table 6 - Schooling returns by type of education (hourly wages)

	2006/08		1995		2006/08		1995	
	Male				Female			
Vocational diploma (3y)	0.11	0.015	0.14	0.021	0.13	0.019	0.20	0.031
High-school diploma								
Vocational	0.22	0.021	0.15	0.031	0.21	0.026	0.24	0.037
Technical	0.29	0.011	0.24	0.013	0.26	0.017	0.25	0.023
Lycei	0.30	0.021	0.24	0.030	0.28	0.022	0.31	0.040
Art lyceum	0.32	0.073	0.36	0.067	0.16	0.047	0.31	0.076
Teacher lyceum	0.24	0.054	0.33	0.055	0.40	0.019	0.43	0.023
Other	0.25	0.047	0.21	0.067	0.24	0.050	0.24	0.073
Bachelor degree	0.58	0.088	0.35	0.071	0.74	0.10	0.47	0.077
Master degree								
Sciences	0.61	0.034	0.52	0.048	0.66	0.033	0.66	0.050
Veterinary	0.69	0.064	0.32	0.102	0.45	0.089	0.63	0.369
Medicine	0.81	0.046	0.60	0.097	0.76	0.054	0.75	0.165
Engineering	0.59	0.030	0.44	0.045	0.40	0.090	0.80	0.261
Architecture	0.49	0.079	0.47	0.131	0.56	0.096	0.70	0.213
Economics & Statistics	0.67	0.033	0.45	0.050	0.52	0.042	0.60	0.099
Political Sciences	0.60	0.049	0.33	0.102	0.54	0.048	0.57	0.140
Law	0.65	0.044	0.65	0.074	0.59	0.049	0.56	0.088
Humanities	0.54	0.041	0.62	0.045	0.69	0.022	0.67	0.031
Others	0.60	0.049	0.55	0.078	0.55	0.044	0.50	0.083
PhD	0.92	0.049	NA	NA	0.81	0.069	NA	NA

(The “Sciences” category comprehends Physics, Mathematics, Information Technologies, Biology and Chemistry).

As we were pointing out in a previous section, things are likely to change greatly if we consider monthly wages instead of hourly wages, like in the previous table. One example for all, the male coefficient for a degree in humanities, considering 2006/08 results, changes from 54% under hourly wages to 28% under monthly wages. This may be due to the fact that some degrees lead to works with a high hourly wage but with few working hours per week. The table below shows results when we consider monthly wages.

Table 7 - Schooling returns by type of education (monthly wages)								
	2006/08		1995		2006/08		1995	
	Male				Female			
Vocational diploma (3y)	0.08	0.014	0.13	0.020	0.18	0.020	0.18	0.028
High-school diploma								
Vocational	0.19	0.020	0.12	0.19	0.27	0.027	0.25	0.032
Technical	0.26	0.010	0.13	0.028	0.31	0.018	0.22	0.020
Lycei	0.23	0.019	0.20	0.012	0.33	0.023	0.27	0.035
Art lyceum	0.20	0.067	0.15	0.061	0.12	0.050	0.21	0.067
Teacher lyceum	0.05	0.050	0.13	0.051	0.37	0.020	0.21	0.021
Other	0.22	0.044	0.11	0.061	0.27	0.053	0.12	0.065
Bachelor degree	0.51	0.080	0.20	0.066	0.99	0.11	0.22	0.069
Master degree								
Sciences	0.47	0.032	0.21	0.043	0.66	0.035	0.30	0.044
Veterinary	0.62	0.059	0.21	0.093	0.46	0.095	0.57	0.327
Medicine	0.80	0.042	0.61	0.089	0.92	0.057	0.60	0.146
Engineering	0.54	0.028	0.28	0.041	0.41	0.095	0.85	0.231
Architecture	0.24	0.073	0.18	0.120	0.45	0.102	0.22	0.189
Economics & Statistics	0.61	0.031	0.39	0.046	0.58	0.044	0.30	0.088
Political Sciences	0.54	0.045	0.15	0.093	0.56	0.051	0.45	0.124
Law	0.62	0.041	0.57	0.068	0.73	0.052	0.26	0.077
Humanities	0.28	0.038	0.14	0.042	0.52	0.024	0.25	0.027
Others	0.35	0.045	0.13	0.069	0.50	0.047	0.11	0.079
PhD	0.85	0.045	NA	NA	0.73	0.073	NA	NA

With respect to high school diplomas, hourly returns for males and females are similar unless in the cases of the art and the teaching lyceum. But the coefficients for males decrease considering monthly wages.

Technical schools and 'normal' licei (those with scientific, humanistic and linguistic studies) show a very similar return under hourly wages, but a technical preparation for males slightly prevails over the others. Girls who decide not to go on with studies are more benefited by the teaching lyceum and the 'normal' ones.

In all cases, the coefficient for bachelor degree has clearly increased from 1995, when it was considered just a "diploma universitario". This may be due to the effect of a requalification of the three-year course that came in 1999 with the Berlinguer's reform as a way to align the Italian education system to the European one. Before this reform, the "diploma universitario" had a much fainter academic flavor and a more job-oriented one, similarly to the relationship between the vocational school and other high schools.

Among master degrees, medicine is still the discipline that grants the highest returns, both for males and females. But please remember that the variable on medicine hides the fact that one kind of degree in medicine lasts 6 and not 5 years, the "single-cycle degree" that gives to the students the possibility to become a doctor. The simple bachelor degree in medicine, even accompanied by a master degree, precludes this occupation; instead it opens up different possibilities, as becoming a radiologist, a physiotherapist, a dietician...

Also a degree in law seems quite valuable.

Another remark is that hourly returns for women are smaller than the males' ones in almost all degrees, while in 1995 there was an opposite situation. Actually, while returns for women have decreased (possible explanations have been offered previously), those for men largely increased with the exception of the law degree. Using monthly returns, the portrait emerging is more ambiguous.

Finally, the big engineering coefficient for females has disappeared, going from 80-5% to 40-41%, from the highest return in degrees to the lowest!

As you may see, the coefficients for the simple dummy variables concerning the highest degree of education achieved hide big differences. The type of education really seems to matter. The results in this section are a point in favor of the heterogeneity idea.

6.4. Predictivity of the model

Residuals are the part of the independent variable that is not described by the model, or, in other terms, the differences between the sample and the estimated function value. In our case it is the part of wage that is not explained by the variables in the regression. So, it can be considered as a measure of predictivity of the model.

If we sum the squares of the differences between values of y and the average y (the residuals), and we divide it by the residual degrees of freedom [3] , we get an "average" value for the residuals.

Computing this measure for each year, we can see how the predictivity of the model evolves over time, that is, if the variables included preserve their importance or if other variables, not included, are assuming more and more value in the determination of the dependent variable.

[3] The degrees of freedom is the number of values in the final calculation of a statistic that are free to vary. The degrees of freedom (df) of the model is the number of parameters to be estimated, including the intercept, minus one. The residual degrees of freedom is the number of observations in the sample is the number of observations in the sample minus the degrees of freedom of the model.

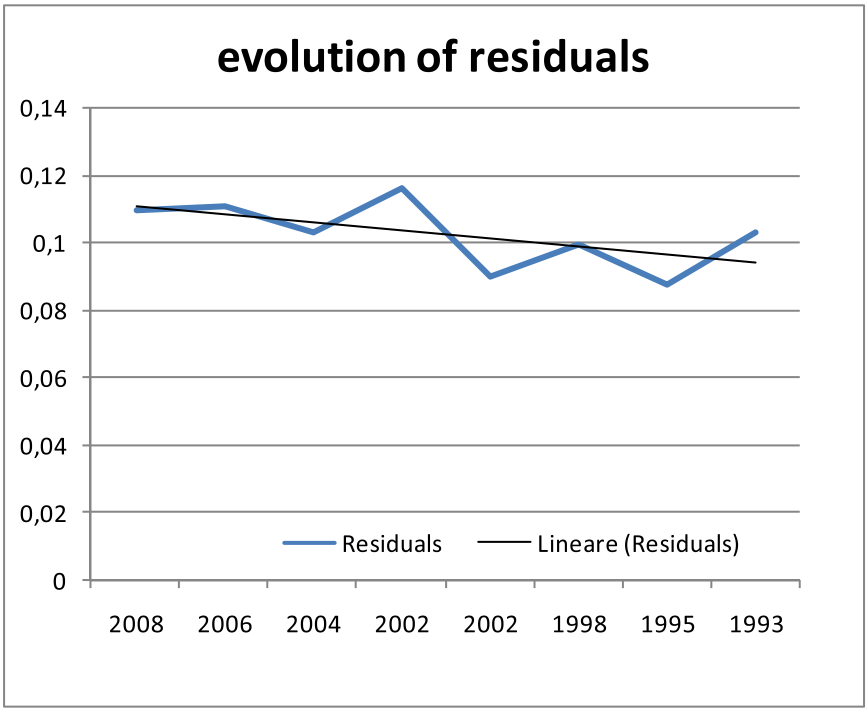

Figure 5

Actually, the slope of the black line shows how the predictivity of the model decreases over time, since average residuals grows. What does it mean? Probably in the last years some new factors emerged, essential in the determination of wages, for example flexibility, knowledge of foreign languages and mastery of information technologies.

6.5. Quantile effects

According to our results so far, the effect of education on earnings changed over time, and its behavior is similar to the one of a sinusoid. But was the behavior the same for each category of wage? In other terms, were the changes the same over the whole distribution, or did we barely scratch the surface?

We merge datasets from 1993, 2000 and 2008 and build this new regression:

$$ln(Y_h) = \alpha + \alpha_{00} * year00 + \alpha_{080} * year08 + \beta schoolys + \beta_{00} schoolys * year00 +$$

$$\beta_{08} * schoolys * year08 + \gamma * Female + \delta * Center + \lambda * South +$$
$$\pi MincerExp + \pi_2 MincerExp^2 + \pi_3 MincerExp^3 + \pi_4 MincerExperience^4$$

where

- schoolys are standard years of education[4].
- year00 and year08 are dummies for years 2000 and 2008 respectively.
- α is the 1993 base wage for male workers without education and living in the North of Italy.
- α_{00} is the increment of the base wage for year 2000 w.r.t. 1993.
- $\alpha_{08}$is the increment of the base wage for year 2008 w.r.t. 1993.
- β is the (percentage) effect of one year more of education.
- β_{00} is the increase in the effect of one year more of education given that we are considering data from 2000.
- β_{08} is the increase in the effect of one year more of education given that we are considering data from 2008.

We chose to use years of education this time - instead of the different levels of education - because of the dissemination of the observations obtained using the latter. In this case we would need enough data to have a good number of observations for each level of education and for each quantile of wages. Not all the cells of this matrix can be filled with the proper number of observations, thus we prefer to go for the schooling years solution and have more robust results.

Moreover, we focused just on 3 years - we picked 1993, 2000 and 2008 because of the similar time span among the three - so the analysis of the features and behavior of the coefficients is simpler. Afterward we provide the graphs for the regression including all years: you will see that the results are not so different from the ones we are analyzing here, so our reasoning and comments can be extended to them.

Here you have the coefficients for the WLS regression, their standard errors and the levels of significance.

[4]For more details, go back to the discussion about standard school years in Section 2.4.

Table 8 - WLS results			
	Coefficient	S.E.	S.L.
α_{00}	**0.130**	0.018	**
α_{08}	**0.235**	0.019	**
school years	**0.059**	0.001	**
β_{00}	**-0.005**	0.002	**
β_{00}	**-0.002**	0.002	
mincer exp	**0.042**	0.006	**
mincer exp^2	**-0001**	0.000	**
mincer exp^3	**0.000**	0.000	
mincer exp^4	**-0.000**	0.000	
central area	**-0.013**	0.007	*
south area	**-0.026**	0.006	**
female	**-0.083**	0.005	**
intercept	**0.838**	0.036	**

The following graphs are a powerful tool. You have the value of the coefficients for each quantile and their confidence intervals, the grey areas. The red line is the coefficient obtained using the OLS regression (pay attention, OLS, not WLS!) and its 95% confidence interval is indicated with the dotted lines. The red line helps us see how much the OLS regression (but similarly the WLS one, being constant as well) is not able to capture differences of the coefficients along the distribution. Last, the black line represents the 0% effect and it is useful to see whether the coefficients are statistically significant in each quantile (if the grey area does not touch the black line).

Let's focus on the effect of education. As you see, the coefficient β for the years of education is strongly significant.

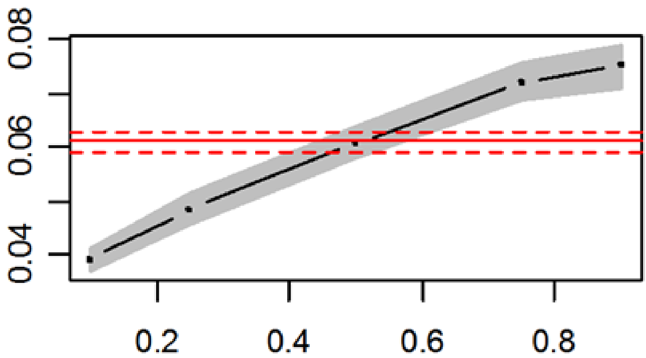

Figure 6 - Effects of years of education by quantiles

But what hides behind the WLS regression coefficient - a simple and constant 5.9% increase given one year more of education? Using the quantile regression we discover that the increase will be smaller (around 4-5%) for those having low wages (LW workers) and higher (7%) for those already earning high wages (HW workers).

Why? One reason could lie in fact that people with low wages usually perform unskilled activities. These jobs require no special training/education, but simple duties which demand the exercise of little or no independent judgment/previous experience. Education here does not help much differentiating wages. Going on with quantiles, the differences in ability become much more relevant in terms of wage. This is one of the explanations that Pereira and Martins (2004) gave, relating it to ability.

They proposed other - possibly complementary - explanations. One reason for the effect of education increasing over the wage distribution could be over-education. In a situation where very high-schooled workers take jobs with low-skill requirements, they will earn a low wage and the coefficient of education in the first quantiles would be particularly low.

The other explanation concerns school quality and fields of study. Individuals in the first quantiles could be precisely those who experienced low-quality schools or who chose types of education with poor returns because of the scarce interest of the labor market for these disciplines.

In any case, assuming that the effect we found is causal and not simply predictive, policy makers willing to increase low wages through education should not expect a result of +5.9% per year, but rather a +4-5%. And of course the effect will be different according to which year more of education the person is achieving (completing the fourth year of the high school is probably different from studying one more for a PhD). If their aim is to reduce wage inequalities, the way they are considering is absolutely the wrongest one.

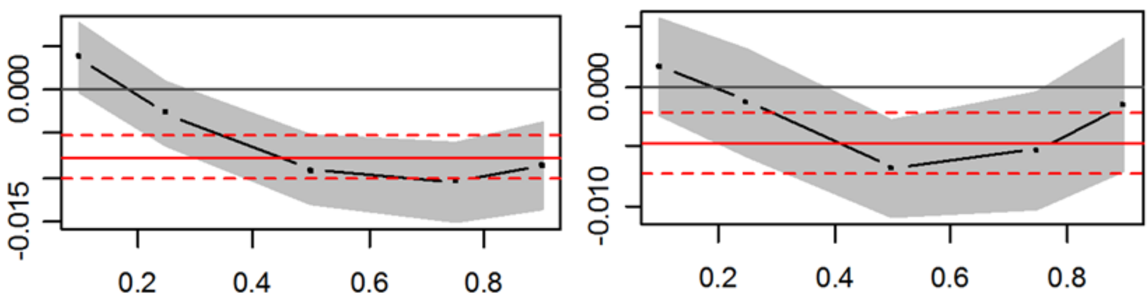

Figure 7 - Change in effect of education, 2000, & Figure 8 - Change in effect of education, 2008, by quantiles

However both β_{00} and β_{08} are negative on average, even if not always significant over the distribution. We see that in 2000 middle wage (MW) workers and HG workers' education was less valuable than in 1993 of about 1%. So, while in 1993 a worker with a high wage gained 7% from one year more of education, in 2000 he/she gained only 6%.

A similar tendency exists in 2008, but just for the median values. Here, instead of a return of 6% for one schooling year more, these workers will get around 5.4%.

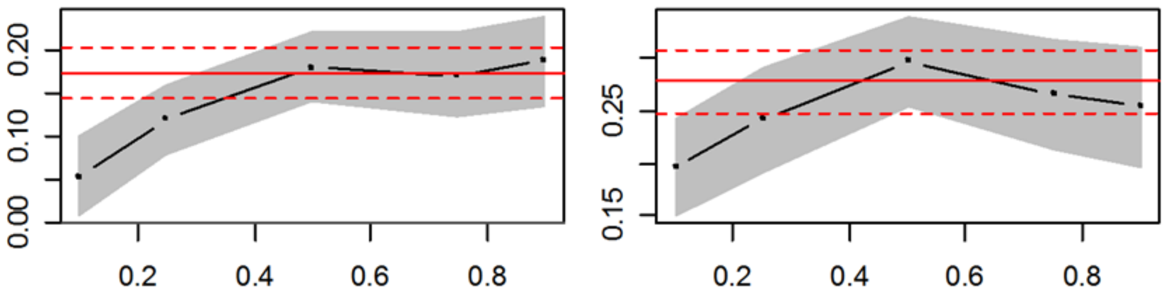

Figure 9 - Change in the intercept, 2000, & Figure 10 - Change in the intercept, 2008, by quantiles

Let's consider base wage, in our case the one for a male worker without education and experience living in the North of the country. In 1993 it was nearly constant to 0.84, thus 2.3 euro per hour[5]. Only HW had a base wage of 2.8 euro (the logarithmic value is 1.037)

The coefficient $\alpha_{00}$is positive and significant, but it is higher for medium/high wages than for low wages. This means that base wages in 2000 grew for everybody, but especially for MW/HW wages. The WLS regression would result in a coefficient around 13%, but in reality the wages grew less for low wage workers (only 5%), while those for high wage workers grew of 17-18%.

The situation of base wages in 2008 sees a big increase with respect to 1993. Here wages increase very much for middle wage workers (30%), less for the extremes of the distribution (19% for LW, 25% for HW workers), resulting in a sort of concave parabola.

So base wages have grown. It could be due to the increasing bargaining power of trade unions and to the coming into existence of more protections and concession for workers. Of course in our calculations we are not taking the effect of inflation, so that this increase in wage could have been not actually perceived.

[5]Since we are considering the logarithm of hourly wages, to obtain the absolute measure we compute the exponential: $e^{0.83} = 2.3$

Some more comments about the other variables we left without a control for the year. Using the quantile regression, a deeper understanding can be achieved.

Women earn from 9% to 9.5% less than men, quite firmly along the distribution. The difference for LW workers is instead 8.4%.

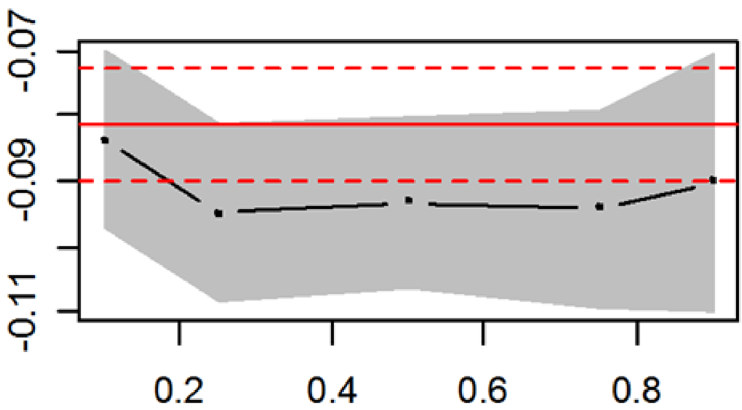

Figure 11 - Effect of the gender by quantiles

The coefficient for the Center of Italy is significant and negative only for LW workers (-6% for the 1st quantile, -2% for the second one). The effect of being in the South changes completely, from a -9.7% for LW and +3 for HW.

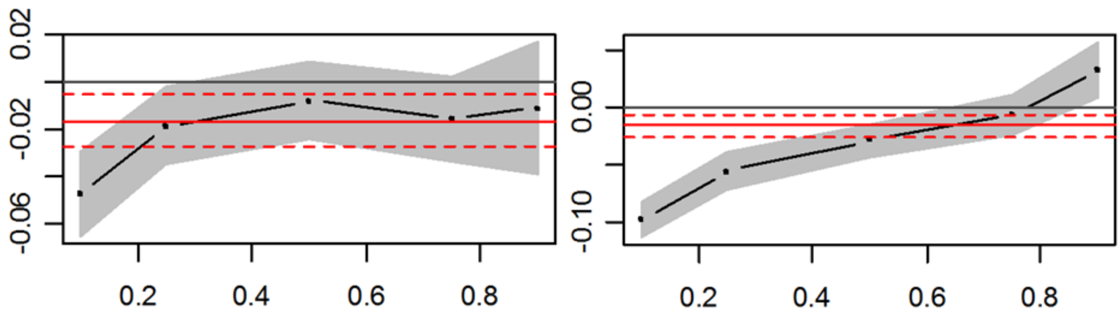

Figure 12 - Effect of Centre area & Figure 13 - Effect of South area, by quantiles

Working in the South of Italy seems to pay especially high-wage, presumably skilled workers. This could be due to the fact that a lot of skilled workers tend to move in the North/Center to find better opportunities and a better quality of life: as a consequence, few skilled ones remain in the South and the demand for them is high.

Behold the strength of the quantile regression!

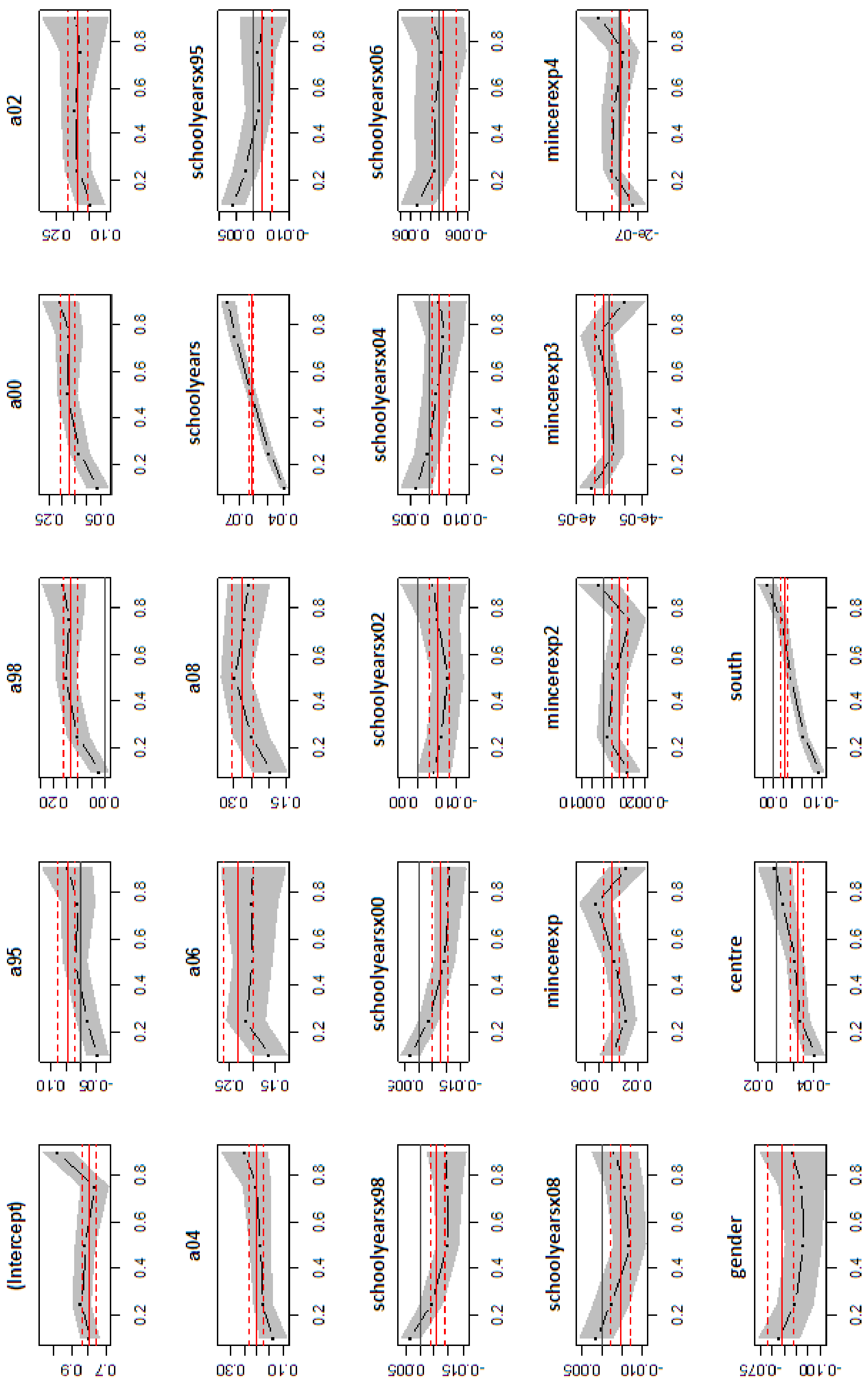

Figure 14 - Effects on education year by year, by quantiles

Conclusions

In the first part of this paper we have defined the true causal effect of education on earnings and we have tried to find out to correct way to estimate it. Surely the obstacles were a lot and hard to overcome, and when going back to the reality of data, some compromises had to be accepted and many limits had to be recognized.

The main issue in economics of education is the ability bias. Without the right proxy for it, or without a good instrumental variable to eliminate the problem - and in our dataset we had none of the two -, we can only resign and accept the fact that the estimate we get is just a predictive effect, not a causal one. We can forecast what the wage of a person will be given his/her education, but we cannot distinguish the importance of education from the one of intrinsic ability.

The other important limitation comes from the way in which surveys are built in Italy. People are asked their net wage, so we cannot separate the effect of a change in tax policy from the effect of education on earnings. Moreover, off-the-book workers do not report their wage for the fear their data are checked and they are prosecuted.

Our analysis shows that the effects for each level of education move nearly together over time. It seems that the same market forces influence all of them, resulting in a sinusoid pattern. The idea of 'waves' in education effect is confirmed by other long term studies as the one by Brunello, Comi and Lucifora (2000).

Another idea finds its validation in our paper: the heterogeneity of the effect of education. As one would expect, different types of education lead to different effects. It is important to notice how results change if we use hourly or monthly wages, as we were describing in the theoretical part.

The quantile regression signals that the effect of education in 2008 has decreased if we look back at 15 years ago, and not uniformly along the distribution. Middle and high wages workers experienced a smaller return, while low wages workers have seen their return to education, already small, quite unchanged.

As the last articles on this issues claims[6], on the wake of the first revelations about the forthcoming Isfol[7] report, education pays, even if not as much as in the past

[6] "In Italia la laurea paga, ma in Europa di più", Claudio Tucci, "Il Sole 24 Ore", 06/06/2011, http://www.ilsole24ore.com/art/economia/2011-06-06/italia-laurea-paga-europa-135739.shtml?uuid=AaChSWdD

[7] Institute for Development and Vocational Training.

and not equally. The Isfol press release, dated June 6th 2011[8], reports that graduated people - including holders of bachelor and master degrees and PhD - earn hourly 38.2% more than the holder of a high school diploma (2010 data). To make a comparison we compute the weighted average of returns, taking as weights the percentage of workers in each educational level. The gap we get is a bit smaller: 32.34%. One reason could be the evolution of the situation from 2008 to 2010 so that graduates experience more comparative advantage.

The Isfol report focuses especially on employment rate, corroborating our idea that more education pays also in terms of possibility to find a job, something that we consciously did not consider in this paper. The employment rate for the holders of a high school diploma is higher than the one for graduates just until 29-30 years. So the degree pays, but

- not as much as in other EU countries.
- not as much as in the past, and we confirmed this outcome

Education seems still to be the most valuable investment available on the market. And the reader shall remember that we are only talking about monetary returns. I firmly believe that education provides people with a much deeper wealth.

[8]http://www.bda.unict.it/Public/Uploads/article/Comunicato%20stampa%20Isfol%20-%2006%20giugno%202011.pdf

Bibliography

[1] J. Angrist, V. Chernozhukov, and I. Fernández-Val. Quantile regression under misspecification, with an application to the us wage structure. 2004.

[2] G.S. Becker and N. Tomes. Human capital and the rise and fall of families. *Journal of Labor economics*, 4(3):1–39, 1986.

[3] D. Belman and J.S. Heywood. Sheepskin effects in the returns to education: An examination of women and minorities. *The Review of Economics and Statistics*, 73(4):720–724, 1991.

[4] M.K.L. Blackburn and D. Neumark. Are ols estimates of the return to schooling biased downward? another look, 1993.

[5] R. Blundell, L. Dearden, and B. Sianesi. Estimating the returns to education: Models, methods and results. 2001.

[6] G. Bottone. Education in italy: is there any return? *ISAE Working Papers*, 2009.

[7] G. Brunello, S. Comi, C. Lucifora, and F.E.E. Mattei. *The returns to education in Italy: a new look at the evidence*. IZA, 2000.

[8] B. Caplan. Two educational heresies: Ability bias vs. signaling. 2006.

[9] D. Card. Earnings, schooling, and ability revisited, 1994.

[10] D. Card. The causal effect of education on earnings. *Handbook of labor economics*, 3:1801–1863, 1999.

[11] D. Card. Estimating the return to schooling: Progress on some persistent econometric problems. 2000.

[12] D. Card and A.B. Krueger. Does school quality matter? returns to education and the characteristics of public schools in the united states, 1990.

[13] D. Card and T. Lemieux. Can falling supply explain the rising return to college for younger men? a cohort-based analysis*. *Quarterly Journal of Economics*, 116(2):705–746, 2001.

[14] D. Checchi, A. Ichino, and A. Rustichini. More equal but less mobile?:: Education financing and intergenerational mobility in italy and in the us. *Journal of Public Economics*, 74(3):351–393, 1999.

[15] B. Chiswick. Schooling, screening, and income. *Does college matter*, pages 151–158, 1973.

[16] O. Deschenes. Unobserved ability, comparative advantage, and the rising return to education in the united states 1979-2000. 2001.

[17] M. Furno. Quantile regression and structural changes in the italian wage equation, 2006.

[18] P. Giustinelli. Quantile regression evidence on italian education returns. *Rivista di Politica Economica*, 94(11/12):49, 2004.

[19] D.L. Golbe. Imperfect signalling, affirmative action, and black-white wage differentials. *Southern Economic Journal*, pages 842–848, 1985.

[20] Z. Griliches. Estimating the returns to schooling: Some econometric problems. *Econometrica: Journal of the Econometric Society*, pages 1–22, 1977.

[21] J. Heckman, A. Layne-Farrar, and P. Todd. Human capital pricing equations with an application to estimating the effect of schooling quality on earnings. *The Review of Economics and Statistics*, 78(4):562–610, 1996.

[22] J.J. Heckman, L. Lochner, and P.E. Todd. Fifty years of mincer earnings regressions, 2003.

[23] T. Hungerford and G. Solon. Sheepskin effects in the returns to education. *The review of economics and statistics*, 69(1):175–177, 1987.

[24] B. Kaymak. Ability bias and the rising education premium in the united states: A cohort-based analysis. *Journal of Human Capital*, 3(3):224–267, 2009.

[25] T. Lemieux. The mincer equation thirty years after schooling, experience, and earnings. *Jacob Mincer A Pioneer of Modern Labor Economics*, pages 127–145, 2006.

[26] L. Lochner. Education and crime. *International Encyclopedia of Education*, 2008.

[27] P.S. Martins and P.T. Pereira. Does education reduce wage inequality? quantile regression evidence from 16 countries. *Labour Economics*, 11(3):355–371, 2004.

[28] J. Mincer. Schooling, experience, and earnings. human behavior & social institutions no. 2. 1974.

[29] J. Mincer. Changes in wage inequality, 1970-1990, 1996.

[30] K.M. Murphy and F. Welch. Empirical age-earnings profiles. *Journal of Labor economics*, 8(2):202–229, 1990.

[31] S.W. Polachek. *Earnings over the life cycle: the Mincer earnings function and its applications*. Now Pub, 2008.

[32] F Sgobbi and Cainarca G.C. How much does it pay? the return to education in italy. 2007.

[33] M. Spence. Job market signaling. *The quarterly journal of Economics*, 87(3):355, 1973.

[34] P. Strawinski. External return to education in europe. *IRISS Working Paper Series*, 2009.

[**?**, 1, 3, 5, 6, 7, 10, 11, 17, 18, 22, 25, 26, 27, 28, 32, 31, 34, 4, 24, 33, 8, 2, 20, 30, 13, 29, 12, 16, 21, 23, 15, 9, 19, 14]

Contents