Quantile Regression Evidence on Italian Education Returns

Pamela Giustinelli*
University of Verona

This study intends to provide some updated empirical evidence on Italian Education Returns through Quantile Regression. Such a methodology enables us to explore the (Quantile Treatment) Effect of Schooling on the (shape of) income conditional distribution (viewed as reflecting the distribution of unobservable ability), and to analyze indirectly the education-ability interaction in the generation of human capital, and its effect on earnings. We obtain estimates displaying a U-*shaped pattern, i.e. higher returns at the highest and lowest quantiles of income, suggesting substitution among human capital factors for low ability individuals, and complementarity for high ability earners.* [JEL codes: C21, I20, J24, J31]

1. - Introduction

The results of several empirical studies on the relationship between the education of individuals and their income show that better educated workers earn higher wages in the labor market (Ashenfelter and Rouse, 1998; Card, 1995). Because of this stylized fact, great interest and effort have been dedicated by labor economists to studying schooling returns, both from a theoretical and empirical point of view.

Although in economic literature the education acquiring de-

---

*<*p-giustinelli@northwestern.edu*>. I am sincerely grateful to three anonymous referees for their comments on this paper, and to Prof. Gustavo Piga for his helpful suggestions. I would also like to thank my thesis Advisor Diego Lubian for always giving me motivation and enthusiasm, and for his encouragement, helpful comments and advice on this work.

cision has been modeled (as an investment increasing individual future income capacity), by addressing the aspects of education endogeneity and heterogeneity across individuals (due to differences in ability, family background *etc.*) which characterize such a choice (Becker, 1967; Card, 1994), there exist many *methodological* problems in estimating returns to schooling as defined in these models.

Common difficulties are represented by the possibility that education is observed with an error, and by the presence of a relevant omitted variables kind of problem (ability is typically unobservable), inducing bias in the estimates. But the main problem in estimating returns to education probably derives from the fact that better educated workers might well earn higher wages not because of the causal effect of additional schooling, but simply because of their greater ability (Ichino, 2001). Thus, the estimated schooling coefficient in the regression of wages would not be a (biased) measure of the causal relation between the two variables but rather of their spurial correlation (*Causality Identification Problem*). Hence the main difficulty in estimating schooling returns concerns the fact that education (i.e. the explanatory variable) and wages (i.e. the dependent variable) are jointly determined, since a real causal relation between the two is lacking.

Copious literature dealing with these issues has developed (see Card, 2001 for a review; Card, 1995; Ichino and Winter-Ebmer, 1999; and for Italy, Flabbi, 1997; Brunello and Miniaci, 1999; Brunello, Comi and Lucifora, 2001; *et* Al.). It mainly comprises empirical studies in which the average marginal return to education is estimated by using *Instrumental Variables Methodology* (IV), under the assumption that the conditional distributions of income simply shift when the level of schooling varies. Actually, this hypothesis might not hold. Moreover, an estimation of the average marginal return to education may not summarize the object of our study efficiently. The idea of estimating returns to schooling by using *Quantile Regression* (Koenker and Bassett, 1978) progresses further to an analysis of individual returns at different quantiles of income distribution. In fact, the effect of education on income may well vary across individuals situated at different "points" of the income distribution itself. By applying quantile regression, it is then possible to check whether the conditional distribution of income simply shifts when education varies or whether there is also a scale effect, such that the entire shape of the distribution changes. Since the conditional distribution of income tends to reflect the distribution of ability which is unobservable, a further important point to stress is that a quantile regression based analysis should be especially informative and explanatory in terms of the relationship between individual ability and education, and the effects that their interaction displays on wages, without imposing any restrictions *a priori* on it. Moreover, the sample quantile regression estimator may be interpreted in terms of the *Quantile Treatment Effect* (QTE) rather nicely: it measures, for each quantile τ, the variation in income required to remain, after treatment (an additional year of schooling), at the τ-th quantile of the conditional distribution itself (provided *Rank Invariance*[1] holds).

The purpose of this study is thus to emphasize the informative power of quantile regression methodology, when compared to a standard OLS estimation for the average individual. On the one hand, our work aims to provide some updated empirical evidence on returns to schooling in Italy (OLS estimates are performed on data drawn from Bank of Italy of 1993, 1995, 1998 and 2000. On the other hand the central aspect is represented by quantile regression estimates on these data.

In Section 2 we briefly present an endogenous model on the education acquiring decision proposed by Card (1994), with a discussion of its implications on the OLS estimator. We also carry out a brief survey of previous empirical studies on returns to schooling, performed with Italian data, applying OLS and IV estimators. In Section 3 the theory on quantile regression is presented, with an interpretation within the returns to schooling framework. We also review certain studies from international literature dealing with conditional income distribution through

[1] The *Rank Invariance* assumption requires that the relative position of individuals in the distribution does not change after the treatment.

quantile regression. Finally, in Section 5 we describe the data used for the estimates, while in Section 6 the results are presented and commented. The conclusions follow.

2. - Education Returns: Economic Theory and Econometric Analysis

2.1 *An Endogenous Model for the Education Acquiring Decision*

An endogenous model for the education acquiring decision was proposed by Card (1994), following Becker (1967). In this model individuals make their educational choice through an optimization problem based on the comparison of the benefits and costs of continuing to study. Earnings and costs are expressed as functions of the years of schooling, and also embody some elements of individual heterogeneity. Each individual makes her schooling choice maximizing a utility function:

$$U(Y, S) = \log(Y) - \phi\ (S)$$

where S is the number of years of schooling, Y is the average annual income of the individual, and $\phi(S)$ is her cost function. More precisely, it is assumed that education induces individuals to accumulate human capital, and that workers endowed with more human capital earn higher wages on the labor market. Thus, the earning function is given by $Y = Y\ (S)$, with $Y' > 0$ and $Y'' < 0$. Human capital revenues increase when education increases, but less than proportionally, as is true for all production factors.

The cost function $\phi(S)$ is strictly convex, i.e. the costs associated with education increase with the years of schooling at an increasing rate. $\phi(S)$ embodies the direct monetary costs (tuition fees, books, transportation *etc.*), the indirect monetary costs or opportunity costs (i.e. what the individual would have earned if she had immediately entered the labor market), and the non-monetary costs (e.g. psychological costs).

Provided that the utility function is globally concave in S, there will exist a (unique) optimal number of schooling years, which is derived by solving the first order condition of the utility maximization problem:

$$\frac{Y'(S)}{Y(S)} = \phi'(S) \tag{1}$$

where $Y'(S)/Y(S)$ is the marginal rate of return associated with an additional year of schooling, and $\phi'(S)$ is the marginal cost of a year of education. The economic significance of the equilibrium condition is standard: the individual reaches her optimal education level when the associated marginal benefits equal marginal costs.

To make the model operational, specific functional forms are attributed to marginal returns and marginal costs respectively. Card (1994) assumes that these two functions are linear and embody in their intercepts elements of individual heterogeneity:

$$\left[\frac{Y'(S)}{Y(S)}\right]_i = \beta_i(S) = b_i - k_b S \quad (k_b \geq 0) \tag{2}$$

$$[\phi'(S)]_i = \delta_i(S) = r_i + k_r S \quad (k_r \geq 0) \tag{3}$$

Therefore the educational choice varies basically for two reasons: 1) differences in ability, b_i, create heterogeneity in the marginal returns to schooling of individuals; 2) differences in the liquidity constraints and in the monetary and cultural conditions of families, r_i, induce heterogeneity in the marginal costs faced by the individuals.

These characterizations of marginal returns and costs associated with education imply specific functional forms for $Y(S)$ and $\phi(S)$ as well. Integrating *(2)* and taking logarithms, the earning function is easily obtained:

$$\log(Y_i) = a + b_i S - \frac{1}{2} k_b S^2 \tag{4}$$

According to this specification, ability affects the slope of the earning function in such a way that, under homothetic preferences, individuals with higher ability will choose a higher level of education[2].

The individual cost function is defined as:

$$\phi(S) = c + r_i S + \frac{1}{2} k_r S^2 \tag{5}$$

where the heterogeneity element constitutes the slope of the function itself.

From the model we obtain the following expression for the optimal education level:

$$S_i^* = \frac{b_i - r_i}{k} \tag{6}$$

with $k = k_b + k_r$.

Equations *(6)* and *(4)* together determine the joint distribution of earnings and schooling[3].

[2] On the contrary, if ability were inserted in the intercept, a_i, the result would be the opposite: in fact, individuals with higher income opportunities on the labor market, for each level of schooling, may well invest less in education, since they would have higher opportunity costs of carrying on studying. In this case, we should then interpret differently this heterogeneity component. CARD D. (1994), following HAUSE J.C. (1972), recalls the concept of "cognitive ability" (the ability of an individual of re-elaborating and applying acquired knowledge; i.e. a form of "working ability" that makes the individual earn higher income, independently of her education and "school ability"). Unifying these two specifications, with an ability component both in the intercept and in the slope of the earning function, the final effect would be uncertain (*a priori*).

[3] Empirical evidence (CARD D. and KRUEGER A.B., 1992; PARK J.H., 1994) shows an approximately linear cross-sectional relation between log earnings and schooling. This circumstance would not seem to be consistent with equation *(4)*, unless $k_b = 0$. In fact, although the relation implied by the model is quadratic, it is reasonable to think that data generated by this model tend to follow a linear pattern. On the one hand, the earnings-education relationship should be concave since once b_i is fixed, individuals with a lower discount rate should choose higher education. On the other hand, since individuals with higher ability tend to study longer, such a relation should display some degree of convexity. In conclusion, the earnings-education relationship for the population as a whole is determined by the combination of these two effects, and it will tend to be more concave the smaller the variance of ability with respect to the discount rate variance.

The marginal return to schooling computed at the optimal education level is given by:

$$\beta_i^* = b_i - k_b S_i^* = \left(1 - \frac{k_b}{k_b + k_r}\right) b_i + \frac{k_b}{k_b + k_r} r_i \tag{7}$$

and represents the causal effect of education on individual earnings[4]. However, because of the Fundamental Problem of Causal Inference (Holland, 1986), it cannot be identified nor measured. Rather, what we are able to identify is the average marginal return in the population, representing a useful parameter which we can compare with the OLS and IV estimators of the education returns:

$$\overline{\beta} = E\left[\beta_i\right] = E\left[b_i - k_b S_i\right] = \overline{b} - k_b \overline{S} = \frac{k_r}{k_b + k_r} \overline{b} + \frac{k_b}{k_b + k_r} \overline{r} \tag{8}$$

Starting from the well known specification:

$$\log(Y_i) = \alpha + \rho S_i + \varepsilon_i,$$

it is possible to show that the probability limit of the OLS estimator can be expressed as

$$p \lim\left(\hat{\rho}_{OLS}\right) = \left(1 - \frac{k_b}{k} + \lambda\right)\overline{b} + \left(\frac{k_b}{k} - \lambda\right)\overline{r} = \overline{\beta} + \lambda\left(\overline{b} - \overline{r}\right) \tag{9}$$

which is clearly inconsistent for the average marginal return to schooling. The bias is positive, since it is the product of two positive elements (λ represents the fraction of the variance of S due to the variability of b; while $\bar{b}$ is surely greater than $\bar{r}$, because the

[4] I am grateful to an anonymous referee for pointing out that it would seem as we faced an aggregation problem, which in fact has never been explicitly arisen in the relevant literature. The specifications that we actually estimate, eq. (31), (both with the years of schooling and the educational dummies as regressors) and the available data allow us to perform a cross-section estimation of the schooling returns, which are average in the OLS case, and correspond to different quantiles in the quantile regression case. As we previously noticed, although in economic literature the education acquiring decision has been modeled by addressing the individual heterogeneity aspects characterizing such a choice (BECKER G.S., 1967; CARD D., 1994), there exist many methodological problems in estimating returns to schooling as defined in these models.

optimal level of schooling cannot be negative). Moreover, it will tend to increase the higher σ_b^2 with respect to σ_r^2 (i.e. the higher λ), and the larger the difference $(\bar{b} - \bar{r})$. The term $\lambda\ (\bar{b} - \bar{r})$ is interpreted as an endogenity bias due to the fact that individuals with higher marginal returns (e.g. with high ability), or lower marginal costs (because less "liquidity constrained"), tend to study longer. In other words, the OLS estimator of the coefficient of schooling in the linear regression of wages is affected by the way in which individuals, characterized by their ability and liquidity constraints, are distributed in the population[5].

2.2 *Ability Bias*

Another problem in estimating education returns stems from the presence of unobserved heterogeneity in the earning levels. In literature the ability bias problem is addressed by including a specific individual component in the earning function (Griliches, 1977). Card (1994) introduces this aspect by adding an individual intercept a_i to the earning function itself:

(10)
$$\log(Y_i) = a_i + b_i S - \frac{1}{2} k_b S^2$$

The expression for the probability limit of the OLS estimator of the schooling coefficient becomes then:

(11)
$$\text{plim}(\hat{\rho}_{\text{OLS}}) = \frac{Cov[\log(Y_i), S_i]}{Var(S_i)} =$$
$$= \frac{E\left[a_i(S_i - \bar{S}) + b_i S_i (S_i - \bar{S}) - \frac{1}{2} k_b S_i^2 (S_i - \bar{S})\right]}{\text{Var}(S_i)}$$

[5] Provided that b_i and r_i are negatively correlated, we will observe more frequently individuals with high ability and not very stringent liquidity constraints, and viceversa, i.e. individuals with high education and high income, or with low education and low income. Consequently, the regression line will be positively sloped, with a gradient increasing with the relative frequency of such observations.

A correlation between a_i e S_i may basically exist for two reasons: because a_i is correlated with b_i or with r_i, i.e. the part of income that individuals potentially earn because of their peculiar (cognitive) ability is correlated with the determinants of (schooling) ability and of the discount rate, such as the financial and cultural family background. Analytically, since:

$$E\left[a_i\left(S_i - \overline{S}\right)\right] = E\left[a_i \frac{\left(b_i - \overline{b}\right) - \left(r_i - \overline{r}\right)}{k}\right] = \frac{1}{k}\left(\sigma_{ab} - \sigma_{ar}\right) \qquad (12)$$

the expression of the bias can be written as:

$$\frac{E\left[a_i\left(S_i - \overline{S}\right)\right]}{\mathrm{Var}\left(S_i\right)} = k\frac{\left(\sigma_{ab} - \sigma_{ar}\right)}{(\sigma_b^2 + \sigma_r^2 - 2\sigma_{br})} \qquad (13)$$

It can be reasonably assumed that a_i and b_i are positively correlated ($\sigma_{ab} > 0$), while a_i and r_i are negatively correlated ($\sigma_{ar} < 0$). In particular, a positive correlation between a_i and b_i can exist if a_i represents a measure of the "cognitive ability" of the individual, and if this kind of ability is correlated with the level of education. On the other hand, a negative correlation between a_i and r_i may result from the circumstance that individuals from the richest families have lower discount rates, or a natural aptitude to study longer, and that these individuals tend in general to earn higher incomes because of more favorable access to the labor market.

2.3 *Explanatory Variable Observed with Error*

Let us finally consider the possibility that the explanatory variable "years of schooling" is observed with an error, i.e. the observed schooling years (S_i^0) differ from the true education level (S_i) due to an additive error term:

$$S_i^0 = S_i + \varepsilon_i$$

where ε_i has mean 0, variance σ_ε^2, and it is uncorrelated with earnings Y_i. In this case the probability limit of the OLS estimator is:

$$(14) \qquad p\lim\left(\hat{p}^0_{OLS}\right) = \frac{\text{Cov}\left[\log\left(Y_i\right), S^0_i\right]}{\text{Var}\left(S^0_i\right)}$$

Since it holds that:

$$\frac{\text{Cov}\left[\log\left(Y_i\right), S^0_i\right]}{\text{Cov}\left[S^0_i, S_i\right]} = \frac{\text{Cov}\left[\log\left(Y_i\right), S_i\right]}{\text{Var}\left(S_i\right)}$$

the probability limit *(14)* can be rewritten as:

$$(15) \qquad p\lim\left(\hat{p}^0_{OLS}\right) = \frac{\text{Cov}\left[\log\left(Y_i\right), S_i\right]}{\text{Var}\left(S_i\right)} \cdot \frac{\text{Cov}\left[S^0_i, S_i\right]}{\text{Var}\left(S^0_i\right)}$$

where the first factor is simply the probability limit of the OLS estimator in the standard case, while the second one represents the bias. It is then immediately possible to verify that such a bias leads to underestimation of the parameter ρ; in fact:

$$\theta_0 = \frac{\text{Cov}\left[S^0_i, S_i\right]}{\text{Var}\left(S^0_i\right)} = \frac{\text{Var}\left(S_i\right)}{\text{Var}\left(S_i\right) + \sigma^2_\varepsilon} < 1$$

Taking into account all the problems concerning estimation procedures, the expression for the probability limit of the OLS estimator will be:

$$(16) \qquad p\lim\left(\hat{p}^0_{OLS}\right) = \theta_0\left[\bar{\beta} + \lambda\left(\bar{b} - \bar{r}\right) + k\frac{\left(\sigma_{ab} - \sigma_{ar}\right)}{\left(\sigma^2_b + \sigma^2_r - 2\sigma_{br}\right)}\right]$$

where we can distinguish three sources of bias of the OLS estimator with respect to the average marginal return to schooling

$\bar{\beta}$:[6] 1) the endogenity bias, $\lambda(\bar{b}-\bar{r})$; 2) the ability bias due to the presence of unobserved elements of heterogeneity in income levels,

$$k\frac{\left(\sigma_{ab}-\sigma_{ar}\right)}{\left(\sigma_b^2+\sigma_r^2-2\sigma_{br}\right)}$$

3) the downward bias generated by measurement errors in the years of schooling, θ_0.

2.4 *The Returns to Schooling in Italy: Previous Works*

In the second half of the 1980s and during the 1990s, several empirical studies were carried out with the aim of estimating the returns to schooling in Italy (in Table 1 the results from the main estimates are presented). As Brunello and Miniaci (1999) observe, the first estimates were based on heterogeneous, and not always representative, data[7].

More recent studies, starting from the second half of the 1990s, make wider use of the SHIW data and perform IV estimates of the returns to schooling for this country. Cannari and D'Alessio (1995), using the SHIW data of 1993 (Banca d'Italia, 1993), and choosing family background variables as instruments, obtain an estimate close to 7%, higher than the previous ones. Colussi (1997) achieve an estimate of 6.6%, with the same data and similar instrumental variables.

---

[6] It is important to recall that the IV estimation procedure should not be affected by these forms of bias. On the other hand, different instruments tend to produce different estimates of the average returns for different subgroups of the population (ICHINO A. and WINTER-EBMER R., 1999). In fact, an IV estimate measures the returns of those individuals that, in the context of the natural experiment considered, are compliers, overestimating the average marginal return to schooling in the population, since such individuals have typically higher returns (discount rate bias) (CARD D., 1994).

[7] For instance, ANTONELLI G. (1985) estimates a standard Mincer equation by applying OLS on a regional data set; he obtains a return of 4.6%. The same result was found by CANNARI L. - PELLEGRINI G. and SESTITO P. (1989), by using a larger sample drawn by the SHIW of 1986 (BANCA D'ITALIA, 1986). LUCIFORA C. and REILLY B. (1990) (on the basis of the ENI-IRI data on individual incomes) perform an OLS estimation by gender, finding out that education returns for females were substantially higher than for males.

TABLE 1

ITALIAN EDUCATION RETURNS IN PREVIOUS STUDIES

Author	Data	Sample	OLS	IV	Instruments
			Estimates		
Antonelli (1985)	ER	(men)	0.046	–	
Cannari, Pellegrini and Sestito (1989)	SHIW 1986	(men)	0.046	–	
Lucifora and Reilly (1990)	ENI-IRI	(women)	0.040	–	
		(men)	0.036	–	
Cannari and D'Alessio (1995)	SHIW 1993	(men)	0.045	0.070	parents' education
Colussi (1997)	SHIW 1993	(men)	0.062	0.076	parents' education
Flabbi (1997)	SHIW 1991	(women)	0.022	0.056	reforms 1962 and 1969,
		(men)	0.017	0.162	distance from university
Brunello and Miniaci (1999)	SHIW 1993-95	(men)	0.048	0.057	parents' education, reform 1969
Brunello, Comi and Lucifora (2001)	SHIW 1995	(women)	0.077	–	parents' education, and professional condition, reform 1969
	(experience)	(men)	0.062	–	
	(age)	(women)	–	0.077	
		(men)	0.048	0.059	
		(men)	0.048	0.061	+ risk aversion
Martins and Pereira (2004)	SHIW 1995	OLS	0.062		
			%		
		Quantile	10	0.065	
			20	0.063	
			30	0.057	
			40	0.057	
			50	0.056	
			60	0.057	
			70	0.061	
			80	0.065	
			90	0.068	

Flabbi (1997) estimates the returns to schooling for females and males separately (SHIW data of 1991, Banca d'Italia, 1991), using as instruments: the binary variable provincie, an indicator of the presence of universities in the province of residence of individuals when they were 19; the variable riforme, constructed considering as exogenous events the Italian school system reforms of 1962 (compulsory lower secondary school) and of 1969 (open access to university regardless of the kind of secondary education qualifications attained). The IV estimates are 0.56 for women and 0.62 for men. Flabbi's new result consists of a reversal of the "hierarchy" in the estimates. In fact, the IV estimated coefficients turned out to be higher for males, while the OLS returns, obtained by Flabbi (1997), were 0.22 for women and 0.17 for men. Indeed, the aim of Flabbi (1997, 1999) is to shed light on the "hierarchy" issue, i.e. he wants to clarify whether the higher female returns could be considered a stylized fact of the Italian labor market or whether they depend rather on the estimation methodology applied. The author concludes that the usual hierarchy (i.e. higher female returns) holds in general within an OLS estimation framework, but it is not independent of the specification, while the hierarchy is reversed by the IV estimates. One preliminary explanation is that the bias affecting the OLS estimates conceals the true relative ratio of the returns by gender. But a second explanation is still possible, namely the reversal does not apply to the entire sample, but only affects those with higher returns (e.g. coming from poor families). The economic significance of this result would then lead us to assume the presence of returns' heterogeneity across individuals in the population and of by-gender premarket discrimination against women.

Brunello and Miniaci (1999) and Brunello, Comi and Lucifora (2001) use the SHIW data of 1993 and 1995 (Banca d'Italia, 1993; 1995) to estimate the education returns with instrumental variables relating to family background (education and professional position of parents), the school system reform of 1969, and (only in the second study) to a measure of individual risk aversion.

Brunello and Miniaci (1999) arrive at an OLS estimate of 4.8%, and an IV estimate of 5.7% for the male households. Similar values are obtained by Brunello, Comi and Lucifora (2001). Yet the IV estimates maintain the "hierarchy" displayed by the OLS ones, i.e. higher returns for females, even on an IV basis.

3. - Earnings, Schooling and Ability: A Quantile Regression Approach

3.1 *Quantile Regression Estimator*

The quantile regression model, introduced by Koenker and Bassett (1978), extends the notion of ordinary quantiles in a location model to a more general class of linear models in which the conditional quantiles have a linear form.

Quantile regression is then a statistical technique intended to estimate, and draw inferences about, conditional quantile functions. Just as classical linear regression methods based on minimizing sums of squared residuals enable one to estimate models for conditional mean functions, quantile regression methods offer a mechanism for estimating models for the conditional median function, and the full range of other conditional quantile functions. A well known particular case of quantile regression is represented by the LAD estimator, proposed by Koenker and Bassett (1978), which adapts the median to a linear function of regressors, minimizing the sum of residuals' absolute values.

LAD estimation is potentially attractive for the same reasons that the median may be a better measure of location than the mean: it always exist in the distribution, and it summarizes the location of the distribution itself, even in presence of outliers, by taking into account only the frequency with which the elements appear in the ordered distribution. Analogously, quantile regression estimators are robust to the outliers among the observations on the dependent variable, and turn out to be more efficient than the OLS estimators when the error term is not normally distributed. Moreover, from a computational (and aesthetic) point of view, the quantile regression model can be expressed as a Linear Program (LP), facilitating estimation and simplifying computation[8]. Besides, it can be easily fitted into the Generalized Method of Moments (GMM) framework, which turns out to be useful for assessing the asymptotic properties of the quantile regression estimator. Finally, an important aspect of this methodology concerns the fact that the estimates of different quantiles can be interpreted as different responses of the dependent variable, to variations in regressors, at different "points" of the conditional distribution of the dependent variable itself[9].

The τ-th quantile regression is defined as:

$$Q_\tau (y_i|X = x) = x'_i\beta(\tau) \tag{17}$$

where $\{x_i|i = 1, ..., n\}$ is a $k \times 1$ vector of the i-th observation on k regressors, and $\{y_i|i = 1, ..., n\}$ is the i-th observation on the dependent variable. The process $u_i = y_i - x'_i\beta$ has a cumulative distribution function $F_u(\cdot)$ and, by definition, $Q_\tau (u_{\tau i}|x_i) = 0$. The quantile regression estimator $\hat{\beta}_\tau$ is defined as a solution of the minimization problem[10]

$$\begin{aligned} \min_{\beta\in\mathbb{R}^k} \sum_{i=1}^{n} \Big(&|y_i - x_i'\beta| \cdot \tau \cdot 1\{y_i > x_i'\beta\} + \\ &+|y_i - x_i'\beta| \cdot (1-\tau) \cdot 1\{y_i \leq x_i'\beta\}\Big) \end{aligned} \tag{18}$$

[8] The representation of the minimization problem for the quantile regression estimator (eq. *18*) is reported in the appendix. For a more detailed treatment of linear programming's application to the context of quantile regression we refer the interested reader to BUCHINSKY M. (1995), and to KOENKER R. and BASSETT G. (1978) for a simple but enlightening bivariate example with five observations. Moreover, all the relevant theory on LP can be found in HILLIER F.S. and LIEBERMAN G.J. (1990).

[9] To clarify the statement we refer to some application settings quantile regression has been used in (see KOENKER R., 2003). For instance, quantile regression methods have been used in pediatric medicine to study the effect of demographic characteristics and mothers' behavior on children' weight, particularly because they allow one to focus the analysis on the lower tail of the distribution. In fact, low birthweight is known to be associated with a wide range of subsequent health problems, and has even been linked to educational attainment and eventual labor market outcomes. Another example is constituted by recent studies modelling the performance of public school students on standardized exams as a function of socio-economic characteristics (like their parents' income and educational attainment), and policy variables (like class size, school expenditures, and teacher qualifications), since it seems rather implausible that such covariate effects act so as to shift the entire distribution of test results by a fixed amount.

[10] In the appendix we recover the representation of quantile regression es-

In light of such a representation, it is worth to underline the quantile regression estimator's property of being robust to the outliers in the observations on the dependent variable. The geometric interpretation of this result goes as follows: given a solution $\hat{\beta}_\tau$ based on some observations $\{y, X\}$, as long as we move these observations up or down leaving them on the same side of the original τ-th quantile regression line (i.e. without letting the sign of residuals $\hat{u}_\tau = y - X\hat{\beta}_\tau$ change), the solution itself is not altered. Only the signs of residuals enter the estimation process; therefore, the extremum observations affect the result depending on their position relative to the estimated hyperplane, but how far above or below it they are is not relevant. Moreover, like the OLS estimator, the quantile regression estimator satisfies a certain number of equivariance properties. In addition, the quantile estimator is invariant to monotonic transformations[11]. This property turns out to be very useful in dealing with estimation of transformed models (which is not true for the OLS estimator)[12].

The minimization problem *(18)*, we formulated above, can be also rewritten as:

(19)
$$\min_{\beta_\tau \in B\tau} \frac{1}{n} \sum_{i=1}^{n} \left(\tau - 1/2 + 1/2\,\mathrm{sgn}\left(y_i - x_i'\beta_\tau\right)\right)\left(y_i - x_i'\beta_\tau\right)$$

where $\mathrm{sgn}(\alpha) = I(\alpha \geq 0) - I(\alpha < 0)$, from which the first order conditions follow:

(20)
$$\sum_{i=1}^{n} \left(\tau - 1/2 + 1/2\,\mathrm{sgn}\left(y_i - x_i'\hat{\beta}\right)\right) x_i = 0$$

As we previously anticipated, considering them within the Generalized Method of Moments (GMM) framework is useful for assessing the asymptotic properties of the quantile regression es-

---

timator as a solution to an optimization problem (eq. *18*), starting from the analogous definition of quantile in the location model.

[11] Given a monotone function $h(\cdot)$, it holds $Q_\tau\,(h(Y)|x) = h(Q_\tau(Y|x))$, while in general $E(h(Y)|x) \neq h(E(Y|x))$.

[12] For a detailed treatment of these properties see KOENKER R. and BASSETT G. (1978).

timator $\hat{\beta}_\tau$. In particular, following Buchinsky (1995), we define the moment function:

$$(21) \qquad \psi(x_i, y_i, \beta) = (\tau - 1/2 + 1/2\text{sgn}(y_i - x'_i\beta_\tau)x_i$$

which satisfies, under the assumptions of model *(17)*, the orthogonality conditions:

$$E[\psi(x_i, y_i, \beta_\tau)] = 0$$

Although this moment function does not satisfy the differentiability condition, it is possible to obtain the asymptotic distribution ($n \to \infty$) of $\hat{\beta}_\tau$, summarized by the following proposition (Buchinsky, 1995):

PROPOSITION 1 If β_τ is in the interior of B_τ, where B_τ is a compact parameter set in $\mathbb{R}^k$; $F_{u\tau}(0|x) = \tau$ with probability 1, and $F_{u\tau}(\cdot|x)$ is continuous with density $f_{u\tau}(0|x) > 0$; and $n^{-1}\Sigma_{i=1}^{n} x_i x'_i \xrightarrow{p} E[x_i x'_i]$, a positive definite matrix; then:

$$(22) \qquad \sqrt{n}(\hat{\beta}_\tau - \beta_\tau) \xrightarrow{d} N(0, \Lambda_\tau)$$

where:

$$(23) \quad \Lambda_\tau = \tau(1-\tau)\,(E\,[f_{u\tau}(0|x_i)x_i x'_i])^{-1}\,E\,[x_i x'_i]\,(E\,[f_{u\tau}(0|x_i)x_i x'_i])^{-1}$$

Sometimes in the context of quantile regression it is assumed that the distribution of $u_{\tau i} = y_i - x'_i\beta_\tau$ does not depend on x_i (so that $E(u_{\tau i}|x_i) = 0$). In this case *(23)* reduces to:

$$\Lambda_\tau = \frac{\tau(1-\tau)}{f_{u\tau}^2(0)}\left(E\left[x_i x'_i\right]\right)^{-1}$$

like in Koenker and Bassett (1978).

A possible method that can be applied to arrive at a consistent estimate of this matrix is the bootstrap (in the empirical part of this work we use this method in its non parametric version). Briefly, let us consider a random sample $(\tilde{y}_i, \tilde{x}_i)$, with $i = 1, ..., n$, generated by the empirical distribution function $F_n(x, y)$. Let be $\hat{\hat{\beta}}_\tau$ the bootstrap estimate of a quantile regression of $\tilde{y}_i$ on $\tilde{x}_i$. If the procedure is repeated M times, obtaining $\hat{\hat{\beta}}_{\tau 1},..., \hat{\hat{\beta}}_{\tau B}$, we have that:

$$\hat{V}^M\left(\hat{\beta}_\tau\right) = \frac{n}{M}\sum_{j=1}^{M}\left(\hat{\hat{\beta}}_{\tau j} - \hat{\beta}_\tau\right)\left(\hat{\hat{\beta}}_{\tau j} - \hat{\beta}_\tau\right)' \tag{24}$$

is the bootstrap estimator of the asymptotic variance of $\hat{\beta}_\tau$.

3.2 *Quantile Regression and Unobserved Ability*

An important aspect of this methodology concerns the fact that the estimates of different quantiles can be interpreted as different responses of the dependent variable, to variations in regressors, at different "points" of the conditional distribution of the dependent variable itself. More precisely, the sample quantile regression estimator can be interpreted in terms of the Quantile Treatment Effect (QTE): it measures, for each quantile τ, the income variation required for staying, after the treatment, at the τ-th quantile of the conditional distribution (provided the Rank Invariance assumption is met). Hence, quantile regression methodology constitutes a more flexible approach for characterizing the effect of education on different quantiles of the income distribution, for analyzing the relation between ability and education, and the effects that their interaction displays on incomes.

In Section 2 we showed that unobserved ability induces heterogeneity in the conditional distribution of income, by affecting both the intercept and the slope of the earning function *(10)*. Let consider now a more general specification for the Mincer equation, following Arias, Hallock and Sosa-Escudero (2001):

$$\ln(Y_i) = \alpha X_i + \beta_0 S_i + \phi(S_i, v_i) + v_i \tag{25}$$

where X_i is a vector of control variables (such as the age of individual, her experience *etc.*), the function φ expresses the education-ability interaction, while v_i is an idiosyncratic term including also the unobserved ability: $v_i = \gamma A_i + \varepsilon_i$. This specification corresponds to *(10)*, with $\varphi = b_i S_i - 0.5 k_b S_i^2$ and $\gamma A_i = a_i$.

Let consider now the case in which the interaction term is simply $\varphi = \delta A_i S_i$, so that returns are given by:

$$\partial \ln(Y_i)/\partial S_i \equiv \beta_i = \beta_0 + \delta A_i \tag{26}$$

where δ captures the effect of ability on education returns. If $\delta < 0$, the returns decrease when ability increases, and viceversa. This leads to a model with random parameters, such that the OLS estimator, applied to *(25)*, provides a consistent estimate of:

$$\partial E\ (\ln(Y_i)|X_i,\ S_i)/\partial S_i = \beta_0 + \delta \bar{A} \tag{27}$$

which is the return to schooling for an individual with average ability, or the so called *Average Treatment Effect*. A drawback of this approach is that it relies on restrictive parametrizations of the schooling-ability interaction. For instance, model *(26)* presupposes that β_i is a monotonic function of ability.

By considering S_i as exogenous, and the restriction on the error term $Q_\tau\ (v_i|S_i) = 0$, the education effect on the τ-th conditional quantile of Y_i is given by:

$$\partial Q_\tau\ (\ln(Y_i)|X_i,\ S_i)/\partial S_i = \partial \ln\ (Q_\tau\ (Y_i)|X_i,\ S_i)/\partial S_i \tag{28}$$

$$= \beta_0 + \partial Q_\tau(\varphi(S_i,\ v_i)|X_i,\ S_i)/\partial S_i \tag{29}$$

$$= \beta_0 + G_v^{-1}\ (\tau|X_i,\ S_i) \equiv \beta_\tau \tag{30}$$

where G_v is some transformation of the ability distribution in the population, and β_τ can be considered as a measure of the QTE of education on incomes, given $\tau \in (0, 1)$[13]. Quantile regression for different values of τ leads to the estimation of an entire family of

[13] In the case of the specification *(26)*, equation *(30)* becomes $\partial \ln\ (Q_\tau\ (Y_i)|X_i,\ S_i)/\partial S_i \equiv \beta_\tau = \beta_o + \delta Q_\tau\ (A_i)$.

education returns reflecting the distribution of ability across individuals. The interaction between education and ability can then be analyzed comparing β_τ for different quantiles τ_k and τ_s, with $k \neq s$.

3.3 *Quantile Regression and Income Distribution in International Literature*

Many studies have been recently conducted with the aim of analyzing the conditional income distributions of different countries by using quantile regression: for instance, Arias, Hallock and Sosa-Escudero (2001); Buchinsky (1994); Mwabu and Schultz (1996); Martins and Pereira (2004); Fitzenberger and Kurz (2003).

In particular, Buchinsky (1994) analyzes changes occurred in the structure of wages in the US during the last decays through quantile regression, by using the *Current Population Survey* (CPS) data (finding higher inequality of incomes on the one hand, even after controlling for individual characteristics; and higher returns on schooling and experience on the other hand). According to the author, it turned out to be essentially important to examine the dynamics at different points of incomes' distribution, since, while wages had generally increased with education at all quantiles of the distribution, the differences in changes had been much higher at some specific quantiles (for instance, the increase in education returns was stronger at the higher quantiles of the distribution).

Such a pattern is found also by Martins and Pereira (2004), who estimate returns to schooling for 16 European countries, with the purpose of analyzing the behavior of income variations at different points of its conditional distribution, when education increases. In other words, by applying quantile regression they compare the returns for individuals with different ability, attempting to shade light on the effect of education on wage inequality. The stylized fact that emerges from their study is that education returns are higher at the highest quantiles of the conditional distribution of income (in line with Buchinsky, 1994). Sweden is taken as a typical example of this evidence while Greece represents an exception, since for the latter country estimates follow an opposite pattern. Martins and Pereira (2004) also estimate education returns for a sample drawn from the 1995 SHIW data (Banca d'Italia, 1995). Their estimates (which are quite similar to ours for that year, given an identical specification and similar criteria for building the sample) show a *U* pattern which however the authors fail to mention. Hence the results show that more able workers (who earn higher hourly wages, conditionally on their individual characteristics) are associated with higher increases in income due to a higher level of education. Martins and Pereira (2004) basically propose three explanations for this evidence.

The first is related to the so called "over-education" phenomenon, due to the presence of situations in which highly educated workers are employed in low ability jobs, and are consequently paid low wages. Hence the more the lower tail of the income distribution of high ability workers is populated by over-educated individuals (i.e. high-skilled individuals for low qualified jobs), the lower the returns will be at the low quantiles of the income distribution.

A second explanation concerns ability and its interaction with education. In fact, the role played by ability heterogeneity, given a certain level of education, tends to become more important in wage terms when education increases.

The last explanation relies on the presence of quality heterogeneity within school systems or across different educational paths (actually the Mincer equation represents a natural instrument with which to control for the quantity of education, but not for its quality). It may be the case that individuals positioned in the lowest part of the distribution are those who received a low-quality education or that attended (*ex-post*) school courses leading to low-paid jobs. Moreover, such differences should prevail at higher educational levels, characterized by a wider variety of educational paths and, eventually, types of degree.

Mwabu and Schultz (1996) apply quantile regression in order to estimate schooling returns by race in South Africa and try to explain the existing differences. In particular, if we consider the quantile regression residuals as reflecting unobserved ability, they interpret increasing/decreasing returns as indicators of complementarity/sostitutability between ability and education as worker productivity increases. Mwabu and Schultz (1996) find that among highly educated whites the returns to schooling increase with income deciles and interpret this evidence in terms of complementarity between ability and education. Conversely returns on attending high school turn out to be negatively correlated with income deciles (i.e. education and ability appear to be substitutes, at least for less able individuals). Returns for Africans with low education (primary school) decrease with income deciles, indicating substitutability; while the high school estimates tend to follow a non-linear (actually convex) pattern. Finally, an irregular pattern is displayed by estimates relating to tertiary education.

4. - Quantile Regression Evidence for Italy

4.1 *The Data*

The data we used to build our samples have been drawn from several waves of the *Survey of Household Income and Wealth* of the Bank of Italy (SHIW of 1993, 1995, 1998 and 2000). Tables 2, 3, 4 and 5 show descriptive statistics for the main variables used in the estimation, and for the other variables of interest. This applies to each year considered; women and men are examined separately.

The samples are composed of employees between 14 and 60, not employed in the agricultural sector. Only men working full-time are considered, while both full-time and part-time female workers are included.

Some preliminary observations on the variables:

1) The years of schooling of individuals are reconstructed starting from the highest educational qualification. (A 0 is assigned to those who failed to complete primary school). From 1995 the classification of educational degrees became more precise, and the following categories were added to the survey: vocational high school (3 years) and first degree (3 years). More-

TABLE 2

DESCRIPTIVE STATISTICS FOR FEMALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	Minimum				*Mean*				*Maximum*			
Age	15	15	16	16	36.96 (10.06)	37.16 (10.13)	38.36 (10.07)	38.74 (10.07)	60	60	60	60
Years of School	0	0	0	0	11.22 (3.83)	11.28 (3.66)	11.94 (3.53)	11.83 (3.53)	20	21	21	21
Junior High	–	–	–	–	0.28 (0.45)	0.26 (0.44)	0.22 (0.42)	0.23 (0.42)	–	–	–	–
High School	–	–	–	–	0.45 (0.49)	0.48 (0.50)	0.52 (0.50)	0.52 (0.50)	–	–	–	–
Tertiary Education	–	–	–	–	0.14 (0.35)	0.14 (0.35)	0.18 (0.39)	0.18 (0.38)	–	–	–	–
Pot. Experience	0	0	0	0	19.74 (11.08)	19.88 (11.06)	20.42 (10.91)	20.91 (10.94)	54	50	53	53
Blue-Collar	–	–	–	–	0.33 (0.47)	0.37 (0.48)	0.32 (0.47)	0.33 (0.47)	–	–	–	–
Part-time	–	–	–	–	0.11 (0.31)	0.13 (0.34)	0.15 (0.35)	0.16 (0.37)	–	–	–	–
Weekly Worked Hours	4	1	5	5	34.38 (9.06)	34.23 (9.48)	34.06 (9.11)	34.82 (9.43)	70	70	65	100
Net Annual Earnings	400	600	500	500	18,677.32 (7849.34)	19,034.68 (8247.59)	21,482.52 (9384.24)	22,787.49 (10100.93)	64,000	63,900	180,000	150,000
Net Hourly Wage	0.80	1	1.20	1.74	12.98 (7.33)	13.39 (7.93)	15.13 (11.41)	15.44 (18.06)	102.08	150	264.29	750
Father Educ.	0	0	0	0	6.27 (4.17)	6.41 (4.10)	6.98 (4.07)	7.09 (4.43)	17	17	17	20

TABLE 2 *(continues)*

DESCRIPTIVE STATISTICS FOR FEMALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	Minimum				*Mean*				*Maximum*			
Mother Educ.	0	0	0	0	5.39 (3.64)	5.60 (3.50)	6.09 (3.59)	6.32 (4.14)	17	17	17	20
Blue-Collar Father	–	–	–	–	0.37 (0.48)	0.42 (0.49)	0.41 (0.49)	0.40 (0.49)	–	–	–	–
Not Employed Mother	–	–	–	–	0.64 (0.48)	0.58 (0.49)	0.58 (0.49)	0.55 (0.50)	–	–	–	–
Self-Employed Fahter	–	–	–	–	0.24 (0.43)	0.27 (0.45)	0.21 (0.41)	0.22 (0.41)	–	–	–	–
No. of observations					2,212	2,355	2,128	2,299				

* Standard deviations in brackets. Earnings in thousands of lire.

TABLE 3

DESCRIPTIVE STATISTICS FOR FEMALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	I Quartile				*Median*				*III Quartile*			
Age	29	29	30	31	37	37	38	39	45	45	46	47
Years of School	8	8	8	8	13	13	13	13	13	13	13	13
Junior High	–	–	–	–	–	–	–	–	–	–	–	–
High School	–	–	–	–	–	–	–	–	–	–	–	–
Tertiary Education	–	–	–	–	–	–	–	–	–	–	–	–
Pot. Experience	11	11	12	12	18	19	20	21	28	28	28	29
Blue-Collar	–	–	–	–	–	–	–	–	–	–	–	–
Part-time	–	–	–	–	–	–	–	–	–	–	–	–
Weekly Worked Hours	30	28	30	30	36	36	36	36	40	40	40	40
Net Annual Earnings	14,000	14,000	16,000	17,000	19,000	19,500	22,000	23,000	24,000	24,000	26,000	28,000
Net Hourly Wage	8.68	9.25	10.42	10.42	11.28	11.57	13.02	13.20	15.05	15.62	16.67	17.36
Father Educ.	5	5	5	5	5	5	5	5	8	8	8	8
Mother Educ.	5	5	5	5	5	5	5	5	8	8	8	8
Blue-Collar Father	–	–	–	–	–	–	–	–	–	–	–	–
Not Employed Mother	–	–	–	–	–	–	–	–	–	–	–	–
Self-Employed Fahter	–	–	–	–	–	–	–	–	–	–	–	–

* Standard deviations in brackets. Earnings in thousands of lire.

TABLE 4

DESCRIPTIVE STATISTICS FOR MALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	Minimum				*Mean*				*Maximum*			
Age	14	14	15	15	39.29 (10.79)	38.90 (10.83)	39.79 (10.42)	39.73 (10.47)	60	60	60	60
Years of School	0	0	0	0	9.88 (3.79)	10.31 (3.78)	10.74 (3.62)	10.73 (3.65)	20	21	21	21
Junior High	–	–	–	–	0.41 (0.49)	0.36 (0.48)	0.35 (0.47)	0.36 (0.48)	–	–	–	–
High School	–	–	–	–	0.33 (0.47)	0.39 (0.48)	0.44 (0.50)	0.43 (0.49)	–	–	–	–
Tertiary Education	–	–	–	–	0.09 (0.28)	0.10 (0.30)	0.11 (0.31)	0.12 (0.32)	–	–	–	–
Pot. Experience	0	0	0	0	23.41 (11.71)	22.59 (11.65)	23.05 (11.13)	22.99 (11.16)	54	54	54	54
Blue-Collar	–	–	–	–	0.50 (0.50)	0.52 (0.50)	0.49 (0.50)	0.50 (0.50)	–	–	–	–
Weekly Worked Hours	5	5	5	5	40.22 (6.86)	40.56 (7.56)	40.55 (7.40)	40.95 (7.67)	70	98	150	100
Net Annual Earnings	800	700	1,000	800	24,300.06 (11,621.83)	25,109.26 (12,557.43)	28,465.28 (13,201.74)	29,687.22 (15,053.04)	110,000	150,000	180,000	350,000
Net Hourly Wage	0.56	1.67	0.65	0.69	13.42 (7.20)	13.90 (7)	15.51 (8.35)	15.74 (8.08)	166.67	93.75	156.25	191.89
Father Educ.	0	0	0	0	5.27 (3.90)	5.50 (3.92)	6.01 (4.02)	6.41 (4.45)	17	17	17	20

TABLE 4 *(continues)*

DESCRIPTIVE STATISTICS FOR MALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	Minimum				*Mean*				*Maximum*			
Mother Educ.	0	0	0	0	4.58 (3.43)	4.81 (3.55)	5.41 (3.62)	5.53 (3.97)	17	17	17	20
Blue-Collar Father	–	–	–	–	0.43 (0.50)	0.50 (0.50)	0.46 (0.50)	0.44 (0.50)	–	–	–	–
Not Employed Mother	–	–	–	–	0.71 (0.46)	0.66 (0.47)	0.63 (0.48)	0.60 (0.49)	–	–	–	–
Self-Employed Father	–	–	–	–	0.24 (0.43)	0.24 (0.43)	0.21 (0.41)	0.19 (0.39)	–	–	–	–
No. of observations					3,618	3,568	2,996	3,265				

* Standard deviations in brackets. Earnings in thousands of lire.

TABLE 5

DESCRIPTIVE STATISTICS FOR FEMALE SAMPLES*

	1993	1995	1998	2000	1993	1995	1998	2000	1993	1995	1998	2000
	I Quartile				*Median*				*III Quartile*			
Age	31	30	31	31	40	39	40	40	48	48	48	49
Years of School	8	8	8	8	8	8	11	11	13	13	13	13
Junior High	–	–	–	–	–	–	–	–	–	–	–	–
High School	–	–	–	–	–	–	–	–	–	–	–	–
Tertiary Education	–	–	–	–	–	–	–	–	–	–	–	–
Pot. Experience	14	13	14	14	23	22	23	23	32.75	32	32	32
Blue-Collar	–	–	–	–	–	–	–	–	–	–	–	–
Weekly Worked Hours	38	38	38	38	40	40	40	40	40	40	42	42
Net Annual Earnings	18,000	18,200	21,000	22,000	22,500	23,800	26,000	27,000	28,000	30,000	32,000	33,000
Net Hourly Wage	9.40	9.90	10.94	11.51	11.98	12.50	13.89	14.17	15.36	16.03	17.36	18.06
Father Educ.	5	5	5	5	5	5	5	5	8	8	8	8
Mother Educ.	0	5	5	5	5	5	5	5	5	5	8	8
Blue-Collar Father	–	–	–	–	–	–	–	–	–	–	–	–
Not Employed Mother	–	–	–	–	–	–	–	–	–	–	–	–
Self-Employed Fahter	–	–	–	–	–	–	–	–	–	–	–	–

* Standard deviations in brackets. Earnings in thousands of lire.

over, from 1995 the survey gathered information relating to the type of high school diploma and university degree that individuals obtained. The latter information allowed us to assign the exact number of years of study to graduates in different subjects; 2) The variables for educational levels (primary school, secondary school, tertiary studies), the occupational status of the individual (blue collar) and that of her father, part-time employment, self-employed status of the father and the non-working status of the mother are binary variables. For the years 1995, 1998 and 2000 those who attained a vocational high school degree are included in the variable high school, while all individuals with a university degree (laurea breve, laurea and post-graduate degree) are grouped in the variable tertiary education; 3) Potential experience is computed as age minus years of schooling minus 6 (since in Italy compulsory school starts at 6 years of age); 4) The average number of hours worked per week includes overtime; 5) Annual income is net of taxes and social security contributions; 6) The hourly net wage is defined as (net annual income)/(months worked * hours worked per week * 4).

By examining the Tables 2, 3, 4 and 5 we observe that the average age of women is between 37 and 39 (increasing over time), while it is close to 40 for men. Years of schooling are higher, on average, for females (11-12 years) than for males (10-11 years), and tend to increase over time. The pattern of the educational dummies obviously reflect that of the years of schooling. In particular, the percentage of individuals (both females and males) attaining a high school diploma or a degree has been increasing over the years, and that of individuals with only a primary school education has been decreasing. As far as working hours are concerned, we observe that part-time employment is still a marginal phenomenon, but it is growing (from 11% in 1993 to 16% in 2000). Men tend to work for a higher number of hours, on average, than women, even after excluding part-time female workers. Men have substantially higher net annual wages than women, but the gap tends to decrease when considering hourly wages; this gap has furthermore been decreasing over the years.

Information about family background has been available since

1993. On average, we note the presence of a certain degree of intergenerational persistency both of education and professional status, particularly for women: in fact, conditional on being (on average) more educated than men in the samples, they tend to have more educated parents, while men have in general a higher percentage of blue collar fathers and non-working mothers.

4.2 *OLS Estimates of the Education Returns (Years of Schooling)*

The first estimate we perform on the data presented is an OLS estimation of the following Mincer equation (Mincer, 1974):

$$\ln(w_i) = \alpha + \beta S_i + \gamma_1 X_i + \gamma_2 X_i^2 + \varepsilon_i \qquad (31)$$

where $\ln(w)$ is the logarithm of the net hourly wage, S are the years of schooling, X is potential experience and ε is the idiosyncratic error.

The OLS estimates of education returns for all the years considered (1993, 1995, 1998 and 2000) are shown in the last column of Table 6, taking women and men separately. The regression for the female samples also includes the dummy partime, which takes value 1 if the woman works part-time and 0 otherwise. In Table 6 the Heteroskedasticity Standard Errors (HCSE) are reported (given the results from the Breusch-Pagan test of heterokedasticity that lead us to reject the null hypotesis of homoskedasticity; see Table 7).

The estimated coefficients are as follows: 8.7% (1993), 7.5% (1995), 6.2% (1998) and 6.1% for women; for men we obtained: 6.7% (1993), 6.6% (1995), and (approximately) 6% (1998 e 2000). Hence we notice that the estimates for women's returns are higher than those for men. The same result was found by Brunello, Comi and Lucifora (2001) for a similar sample of individuals. Actually, the intercepts, which represent the log-wage of a full-time worker with no education and no potential experience, are higher for men; moreover, the gender gap tends to reduce over the years. Flabbi (1997), (1999) focuses on the gender difference issue, proposing some possible interpretations for the empirical evidence

TABLE 6

RETURNS TO SCHOOLING ESTIMATES (YEARS)*

Quantile	0.05	0.10	0.25	0.50	0.75	0.90	0.95	OLS
				Females				
1993	0.0919 (0.0072)	0.0829 (0.0053)	0.0784 (0.0030)	0.0826 (0.0030)	0.0893 (0.0025)	0.0955 (0.0032)	0.0973 (0.0041)	0.0867 (0.0024)
1995	0.0792 (0.0057)	0.0698 (0.0043)	0.0692 (0.0031)	0.0735 (0.0025)	0.0828 (0.0031)	0.0879 (0.0040)	0.0876 (0.0046)	0.0749 (0.0026)
1998	0.0615 (0.0053)	0.0527 (0.0040)	0.0526 (0.0030)	0.0571 (0.0027)	0.0729 (0.0032)	0.0751 (0.0037)	0.0741 (0.0046)	0.0622 (0.0028)
2000	0.0549 (0.0072)	0.0575 (0.0062)	0.0539 (0.0025)	0.0596 (0.0030)	0.0685 (0.0028)	0.0778 (0.0033)	0.0766 (0.0052)	0.0612 (0.0028)
				Males				
1993	0.0612 (0.0052)	0.0567 (0.0027)	0.0539 (0.0017)	0.0629 (0.0018)	0.0715 (0.0024)	0.0790 (0.0023)	0.0829 (0.0040)	0.0676 (0.00177)
1995	0.0647 (0.0039)	0.0651 (0.0024)	0.0609 (0.0022)	0.0598 (0.0016)	0.0671 (0.0018)	0.0732 (0.0028)	0.0769 (0.0039)	0.0660 (0.00174)
1998	0.0523 (0.0060)	0.0522 (0.0045)	0.0514 (0.0021)	0.0534 (0.0019)	0.0624 (0.0026)	0.0741 (0.0030)	0.0830 (0.0038)	0.0599 (0.00215)
2000	0.0640 (0.0042)	0.0562 (0.0037)	0.0491 (0.0025)	0.0532 (0.0016)	0.0632 (0.0022)	0.0721 (0.0027)	0.0719 (0.0042)	0.0603 (0.00194)

* Bootstrap standard errors in brackets. Estimates performed in *R*.

TABLE 7

TESTS FOR HOMOSKEDASTICITY AND SIMMETRY

	Females	Males	Females	Males	Females	Males	Females	Males
Years	*1993*		*1995*		*1998*		*2000*	
R^2 (OLS)	0.44695	0.42595	0.35140	0.43385	0.26911	0.35439	0.27804	0.36538
B-P Test*	0.0003	0.00021	0.0000	0.00257	0.00000	0.00004	0.00002	0.07050
(*p*-value)								
Heterosk.**	119,315	163,041	200,300	71,160	185,342	108,299	295,589	121,678
p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Simmetry**	1,789,723	3,828,609	1,981,360	5,021,722	1,739,939	3,404,891	1,689,244	4,510,693
p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
R^2 (OLS)	0.45033	0.43232	0.35056	0.43183	0.27216	0.34902	0.27619	0.37004
B-P Test*	0.00227	0.00001	0.00000	0.00261	0.00000	0.00017	0.00004	0.06823
(*p*-value)								
Heterosk.**	160,972	176,297	212,283	88,517	210,779	157,659	338,073	195,671
p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Simmetry**	1,691,700	3,933,010	2,061,016	4,608,752	1,241,258	3,462,240	1,425,513	4,136,447
p-value	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
n° of obs.	2,212	3,618	2,365	3,568	2,128	2,996	2,299	3,265

* Breusch-Pagan Test. ** BUCHINSKY M. (1995).

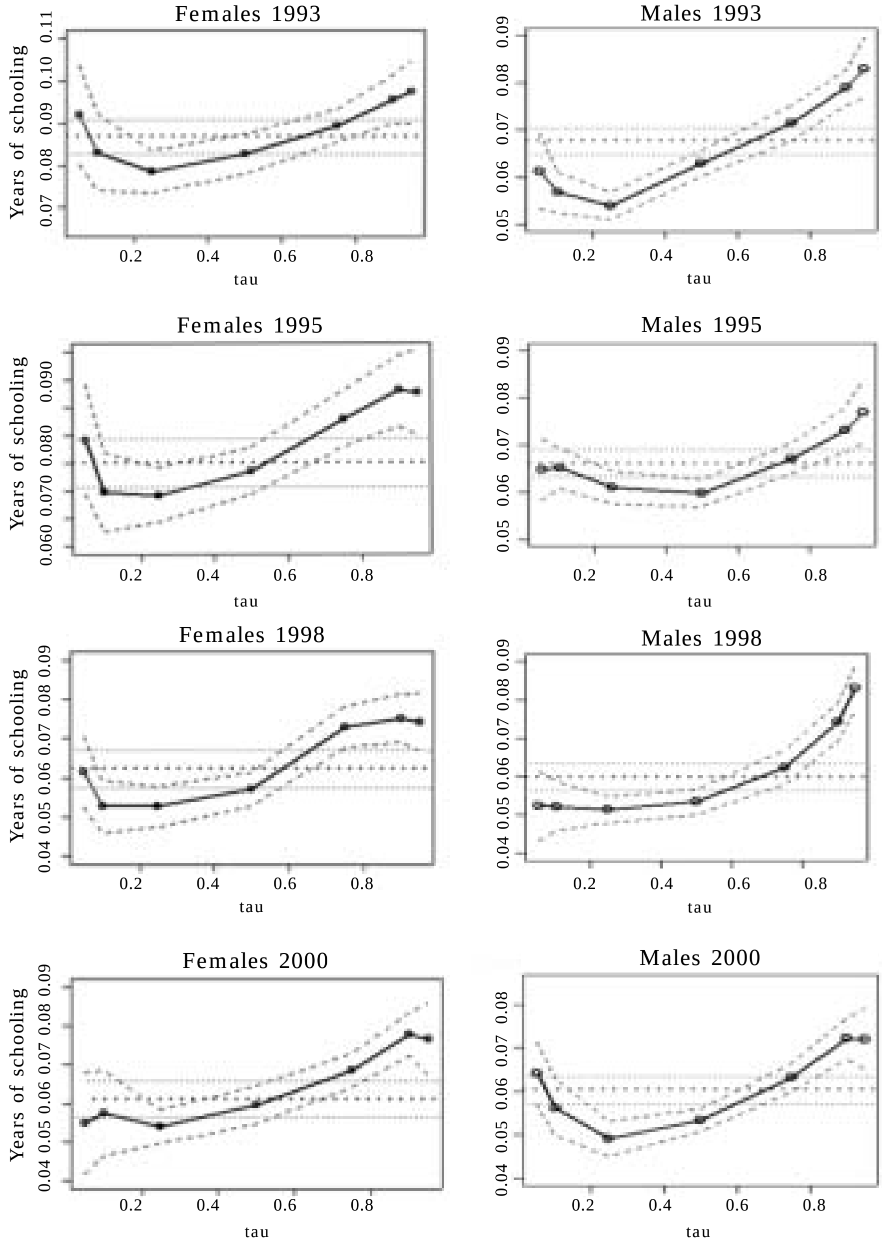
he finds and outlining the dependence of the results on the applied estimation technique (OLS, OLS with Heckman's correction, IV). A possible interpretation of the higher female returns (in the light of Card's model) may depend on the presence of a larger ability bias of the OLS estimator in the female samples. Assuming that the measurement errors on the explanatory variable "years of schooling" affect women and men in the same way, and that $\bar{b}$ e $\bar{r}$ do not differ, in principle, for females and males, we ask whether it is reasonable to think that there exists a stronger negative correlation between abilities (a_i and b_i) and liquidity constraints for women, and a stronger positive correlation between a_i and b_i themselves. A negative correlation between individual ability and liquidity constraints can be easily justified, provided that there exists a certain degree of intergenerational persistence of ability. In such a situation more able parents choose higher levels of education, earn higher incomes and have more able daughters on average. Hence, in the daughters' generation the most skilled individuals tend to have parents with the highest incomes. In so far as the parents' higher incomes reduce the daughters' liquidity constraints, in the former's generation the most able individuals will have lower marginal costs of education. Looking at the descriptive statistics this seems a plausible assumption, since women have higher education on average, and (on average) higher educated parents. Moreover, the percentage of women with a blue-collar father and/or a non-working mother is lower than the correspondent percentage in the male samples.

5. - Quantile Regression Estimates of Education Returns (Years of Schooling)

In Table 6 the quantile regression estimates for the mincerian specification are presented. More precisely, they are the estimated coefficients for the female and male education returns at 7 quantiles of income distribution $\tau = \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$, for the four considered years. The values for the bootstrap standard errors are also presented. In Graph 2, each plot

GRAPH 2

RETURNS TO SCHOOLING ESTIMATES AT DIFFERENT QUANTILES
(YEARS OF SCHOOLING)



depicts the estimated education coefficient of the relative sample in the quantile regression model. In particular, the dotted solid line represents the point estimates, with the hatched lines tracing out a 90% pointwise confidence interval. Superimposed on the plot is a dotted line, representing the OLS estimate of the mean effect, with its 90% confidence interval. In Table 7 the results for the heteroskedasticity and symmetry tests (Buchinsky (1995, 1998)) are shown. The *p-values* lead the null hypoteses of homoskedasticity and symmetry to be rejected for all samples.

The estimates, initially high at the lowest quantiles ($\tau = 0.05$), tend to decrease for those individuals staying below the first quartile of income distribution ($\tau = 0.25$), then they continue to increase, finally stabilizing at the highest quantiles of the distribution ($\tau = 0.90$, $\tau = 0.95$). Moreover, the OLS estimate of the average effect is generally a value corresponding to the estimate for $0.50 < \tau < 0.75$, i.e. the average marginal return to schooling is usually higher than the estimated return of the median income earner. Higher returns for women are confirmed. Moreover, we notice different behavior of the estimates for female and male samples. On the one hand, the estimates for men that stay at the lowest quantiles of the distribution are, in fact, much further below the OLS level, and outside its bandwidth, than the estimates for females; on the other hand, the estimates for male samples increase faster than the female ones, particularly at the highest quantiles. Let us recall the meaning of a quantile regression estimate of education returns. For each quantile, the years of schooling coefficient tells how much the individual log-wage should increase in order for the individual to maintain that quantile in the wage distribution after studying one year longer (i.e. after receiving the treatment). For instance, looking at the second row of Table 6 (sample of 1995) we read that the log-wage of a worker at the median quantile of the log-wage distribution itself ($\tau = 0.50$) should increase by 7.35% in order to maintain the median quantile in the distribution of workers with an additional year of education. Thus, estimates tell us that increments in income for individuals located at the lowest and the highest quantiles of the distribution should be higher for them to maintain their income quantile after studying one more year. In other words, this means that the shape of the conditional distribution of income changes; or, to be more precise, there is not only a shift effect on earning distribution, but its shape also varies with education.

It is also possible to interprete the results in terms of substitutability/complementarity between education and ability. By inspecting of Graph 2, the decreasing returns for the lowest quantiles (i.e. for the low ability individuals) show a certain degree of schooling-ability substitutability. But the relationship between the two factors tends to become of complementarity at the high quantiles (i.e. for high ability individuals).

5.1 *OLS Estimates of Education Returns (Educational Dummies)*

In some schooling systems — including the Italian system — additional investment in education that does not lead to the award of a degree might not grant additional labor market returns. Thus, in order to address this issue we use educational dummies rather than years of schooling in the earnings specification. More precisely, we estimate the (OLS and quantile) regressions of the log net hourly wage on the binary variables "junior high", "high school" and "tertiary education", controlling, as before, for potential experience and its square, and maintaining the variable "partime" for females. The results for the OLS estimates are shown in the last column of Table 8. The estimated coefficients of the educational dummies should be interpreted as differentials with respect to the baseline return accruing to individuals with no school or with only primary school. For example, a male employee with a high school degree earns, on average, 48% more than a male employee with the same potential experience belonging to the reference group in 1995. The pattern of the estimated returns by educational level confirms that there is a monotonic (positive) relationship that links returns to education to the highest level of education attained.

Moreover, we find that in 1993 returns to education are higher for female employees than for males; in 1995 they tend to de-

TABLE 8

RETURNS TO SCHOOLING ESTIMATES
(educational *dummies*)*

Quantile	0.05	0.10	0.25	0.50	0.75	0.90	0.95	OLS
				Females				
				Junior high				
1993	0.2947	0.3075	0.1997	0.18656	0.2236	0.2347	0.2595	0.2352
	(0.0909)	(0.0691)	(0.0446)	(0.0324)	(0.0286)	(0.0320)	(0.0422)	(0.0302)
1995	0.3034	0.1821	0.1692	0.14204	0.1982	0.1818	0.1993	0.1802
	(0.1212)	(0.1004)	(0.0437)	(0.0345)	(0.0318)	(0.0659)	(0.0736)	(0.0361)
1998	0.1074	0.1302	0.1053	0.16207	0.1493	0.2039	0.2138	0.1140
	(0.0928)	(0.0764)	(0.0390)	(0.0453)	(0.0369)	(0.0823)	(0.1948)	(0.0444)
2000	0.3883	0.3262	0.1311	0.10079	0.0737	0.0117	0.0480	0.01198
	(0.1046)	(0.0865)	(0.0583)	(0.0409)	(0.0456)	(0.0684)	(0.1111)	(0.0455)
				High school				
1993	0.7535	0.6807	0.5577	0.5352	0.6229	0.6987	0.7730	0.6265
	(0.0749)	(0.0734)	(0.0495)	(0.0328)	(0.0329)	(0.0349)	(0.0518)	(0.0316)
1995	0.7319	0.5414	0.4855	0.4669	0.5587	0.5905	0.6316	0.5193
	(0.1238)	(0.0971)	(0.0422)	(0.0358)	(0.0301)	(0.0602)	(0.0734)	(0.0367)
1998	0.4373	0.3930	0.3621	0.3861	0.4284	0.5072	0.5502	0.3799
	(0.0785)	(0.0755)	(0.0393)	(0.0474)	(0.0400)	(0.0812)	(0.1950)	(0.0439)
2000	0.6927	0.5866	0.3663	0.3377	0.3674	0.3721	0.4273	0.3919
	(0.0995)	(0.0892)	(0.0623)	(0.0452)	(0.0475)	(0.0683)	(0.1078)	(0.0456)
				Tertiary education				
1993	1.1565	1.0444	0.9746	1.00354	1.1148	1.1693	1.1827	1.0816
	(0.0825)	(0.0833)	(0.0601)	(0.0392)	(0.0366)	(0.0404)	(0.0508)	(0.0362)
1995	1.0370	0.8113	0.8055	0.87338	1.0294	1.0499	1.0866	0.9032
	(0.1272)	(0.1041)	(0.0533)	(0.0488)	(0.0367)	(0.0662)	(0.0776)	(0.0422)
1998	0.5880	0.5965	0.6260	0.70640	0.8786	0.9239	0.9422	0.7159
	(0.0846)	(0.0792)	(0.0455)	(0.0536)	(0.0418)	(0.0831)	(0.2098)	(0.0503)
2000	0.8360	0.8331	0.6146	0.66452	0.7639	0.7783	0.8193	0.7028
	(0.1172)	(0.0997)	(0.0622)	(0.0487)	(0.0520)	(0.0699)	(0.1244)	(0.0504)

* Bootstrap standard errors in brackets. Estimates performed in *R*.

(cont.) TABLE 8

RETURNS TO SCHOOLING ESTIMATES
(educational *dummies*)*

Quantile	0.05	0.10	0.25	0.50	0.75	0.90	0.95	OLS
				Males				
				Junior high				
1993	0.2344	0.1923	0.1424	0.1782	0.1762	0.2321	0.2423	0.1825
	(0.0622)	(0.0271)	(0.0171)	(0.0157)	(0.0233)	(0.0255)	(0.0282)	(0.0162)
1995	0.1754	0.2040	0.2006	0.2054	0.2059	0.2258	0.2075	0.2096
	(0.0608)	(0.0354)	(0.0246)	(0.0197)	(0.0251)	(0.0273)	(0.0507)	(0.0189)
1998	0.1047	0.1421	0.1689	0.1513	0.1976	0.1702	0.1973	0.1649
	(0.0587)	(0.0418)	(0.0256)	(0.0187)	(0.0187)	(0.0274)	(0.0492)	(0.0244)
2000	0.3485	0.3047	0.1803	0.1417	0.1461	0.1409	0.1424	0.1720
	(0.0939)	(0.0546)	(0.0355)	(0.0236)	(0.0242)	(0.0661)	(0.0492)	(0.0267)
				High school				
1993	0.5188	0.4658	0.4025	0.4562	0.5083	0.5801	0.6065	0.4950
	(0.0573)	(0.0291)	(0.0174)	(0.0174)	(0.0274)	(0.0295)	(0.0340)	(0.0179)
1995	0.5087	0.4947	0.4550	0.4346	0.4699	0.5227	0.5756	0.4810
	(0.0589)	(0.0351)	(0.0246)	(0.0197)	(0.0274)	(0.0287)	(0.0612)	(0.0201)
1998	0.3670	0.3496	0.3498	0.3472	0.4203	0.5164	0.6055	0.4023
	(0.0578)	(0.0444)	(0.0276)	(0.0225)	(0.0225)	(0.0367)	(0.0512)	(0.0259)
2000	0.6178	0.4922	0.3473	0.3379	0.4123	0.4812	0.4859	0.4106
	(0.0870)	(0.0561)	(0.0376)	(0.0256)	(0.0285)	(0.0678)	(0.0514)	(0.0278)
				Tertiary education				
1993	0.8136	0.7654	0.7460	0.8541	0.9309	1.0440	1.1173	0.8845
	(0.0757)	(0.0470)	(0.0362)	(0.0299)	(0.0324)	(0.0501)	(0.0518)	(0.0267)
1995	0.8016	0.8187	0.8125	0.8325	0.9242	0.9902	1.0201	0.8784
	(0.0720)	(0.0538)	(0.0369)	(0.0330)	(0.0347)	(0.0396)	(0.0682)	(0.0264)
1998	0.6293	0.6718	0.6898	0.7207	0.8686	0.9218	1.0447	0.7763
	(0.0825)	(0.0570)	(0.0389)	(0.0296)	(0.0296)	(0.0534)	(0.1164)	(0.0357)
2000	0.8686	0.8082	0.6656	0.6980	0.8296	0.8940	0.9292	0.7774
	(0.1028)	(0.0715)	(0.0424)	(0.0320)	(0.0372)	(0.0732)	(0.0673)	(0.0334)

* Bootstrap standard errors in brackets. Estimates performed in *R*.

crease; while in subsequent years estimated returns for males become higher than those obtained for the female samples. We also notice that the estimated returns of female workers follow a decreasing pattern over time.

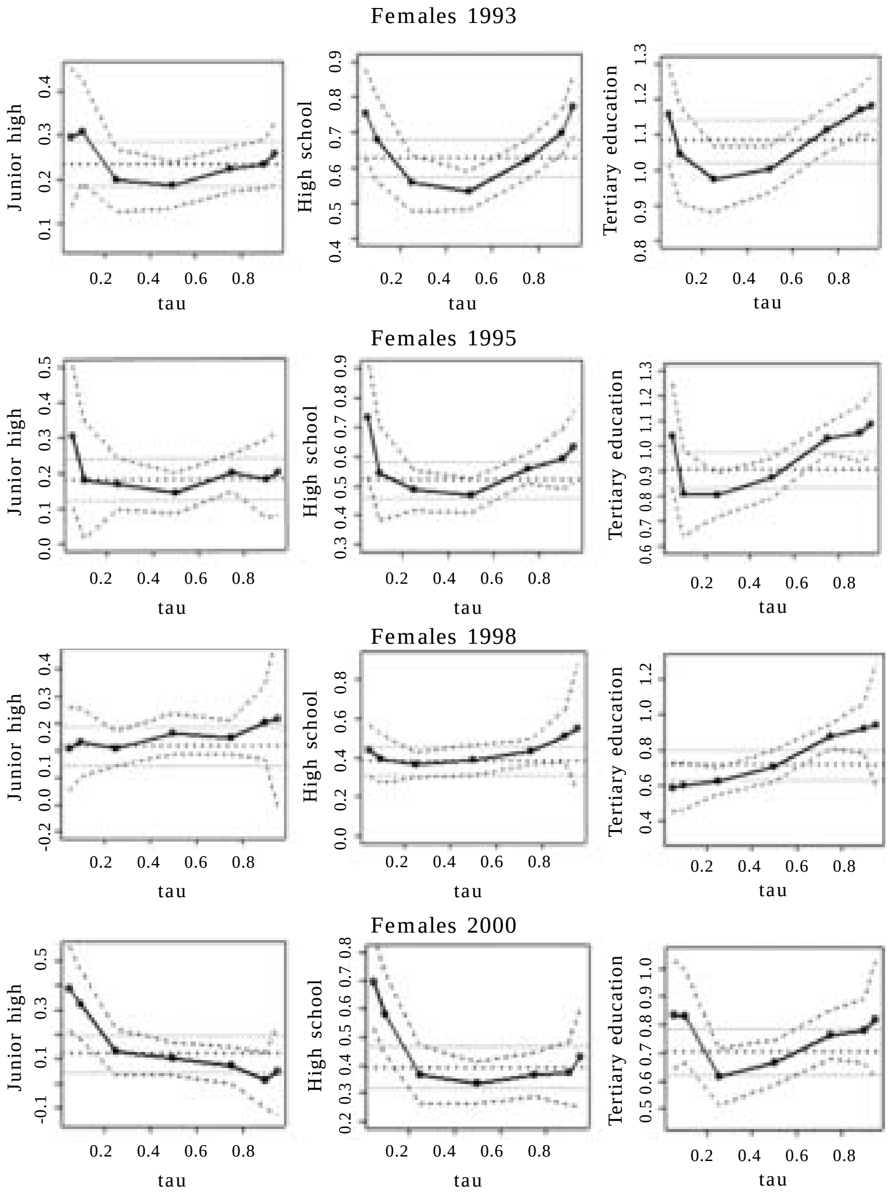
5.2 *Quantile Regression Estimates of Education Returns (Educational Dummies)*

The quantile regression estimates of the educational dummies are presented in Table 8 and plotted in Graphs 3 and 4. Within this specification the estimated coefficients should be interpreted (for each quantile) as the percentage increment of the log-wage that an individual with at most primary education should receive to maintain her quantile on the wage conditional distribution, if she attained a junior high school diploma, a high school diploma or a degree, respectively. Upon inspection of the tables we notice that, in general, education returns are higher for females than for males, at each educational level, although the gap in favor of females tends to vanish between 1998 and 2000, when we find higher returns for male employees, at least at the highest quantiles. Moreover, new aspects come out by controlling for educational levels, that we were not aware of from the previous specification. In fact, from the figures it immediately emerges that the estimated coefficients of junior high fall within the OLS bandwidth, for almost every quantile and sample, i.e. these estimates display little variability across quantiles. In terms of the education-ability relationship, the two factors do not seem to display any particular pattern across quantiles. Only in 2000 do the estimates tend to decrease when the quantile increases, calling for substitutability.

The pattern changes however when we analyze the estimated coefficients of high school and tertiary education. As far as the first is concerned, we observe a convex shape of the estimates across quantiles, similar to that of the estimates for years of schooling. In other words, at the lowest and the highest quantiles individuals need (ceteris paribus) a higher increment in wage (with respect to the median earners) to maintain their quantile on

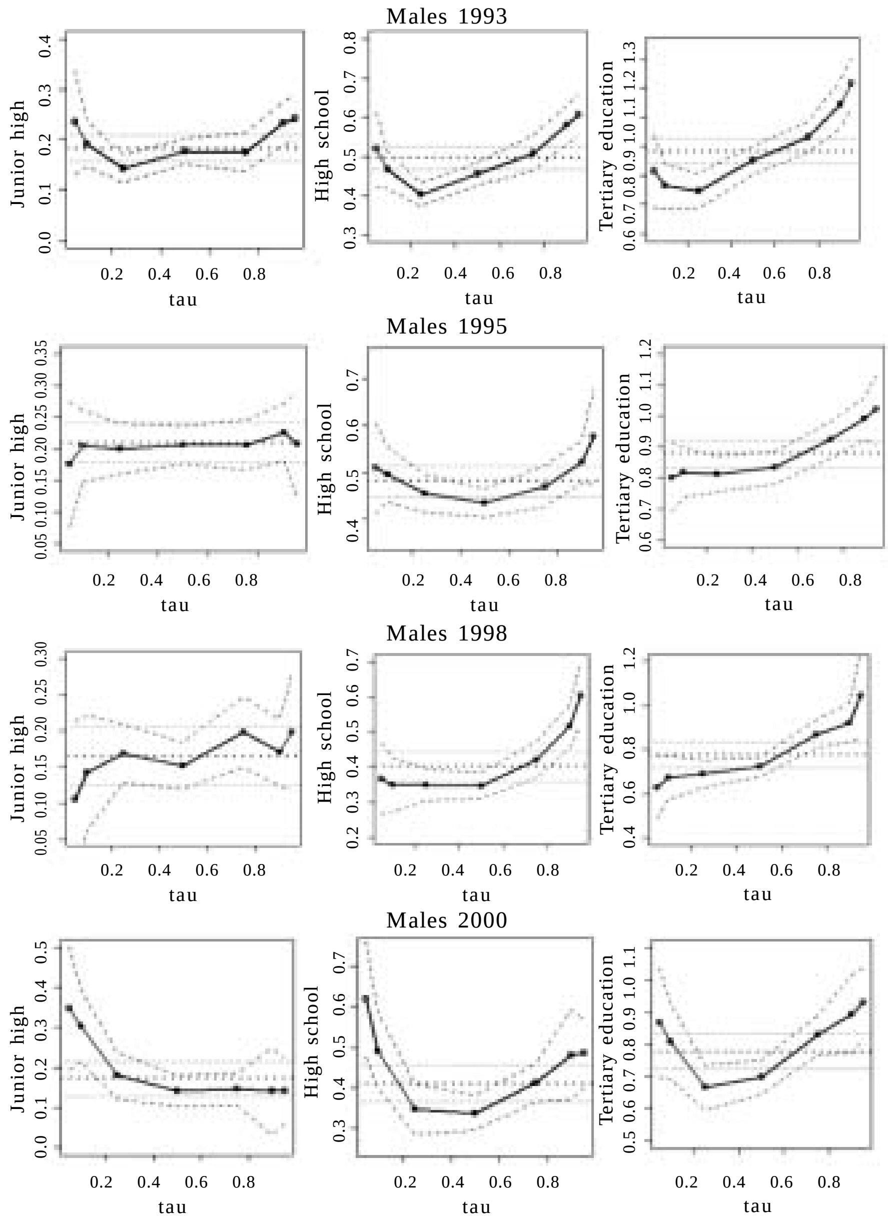
GRAPH 3

RETURNS TO SCHOOLING ESTIMATES FOR FEMALES
AT DIFFERENT QUANTILES
(educational *dummies*)



GRAPH 4

RETURNS TO SCHOOLING ESTIMATES FOR MALES
AT DIFFERENT QUANTILES
(educational *dummies*)



the earning distribution of individuals with a high school diploma. Thus, we conclude again that the shape of the conditional distribution of earnings for individuals who attained a high school diploma is different from the conditional distribution of those with only primary education. Finally, we can interpret this pattern for the estimates in terms of substitutability between education and ability at the first quantiles, and in terms of complementarity at the highest quantiles.

The estimated coefficients of the variable "tertiary education" display variability across quantiles too, following an increasing pattern when the quantile increases. This suggests complementarity of factors. This is particularly evident for 1995 and 1998 and, in general, for male samples. For female employees the pattern of the estimates tends to be more convex, with very high values at the lowest quantiles (τ = 0.05 and τ = 0.10), similar or even higher than the ones corresponding to the highest quantiles (τ = 0.90 and τ = 0.95); while for men an increasing trend prevails. There is an exception in the behavior of estimates in 2000, when the pattern is quite convex for the estimated coefficients of males, and the estimates for women decreasing across quantiles (substitutability).

Another interesting point is that in the female sample in 1998 we observe the flattest behavior of the estimates (see variables "junior high" and "high school") and the highest drop in the marginal returns for women. In 1998 a certain variability across quantiles is displayed only by graduate female workers. In particular, estimates are higher at the highest quantiles (complementarity).

6. - Conclusions

With this study we intend to provide an update of the empirical evidence on returns to schooling in Italy (using the SHIW data of the 1990s and 2000), estimating them both via OLS and Quantile Regression, and emphasizing the informative power of the latter methodology, compared to a standard OLS estimation for the average individual.

The OLS results are very similar to those found in Italian literature on education returns, and they confirm the evidence of higher female returns. We interpret this result as due to a stronger effect of the endogeneity bias and ability bias on female estimates, assuming a stronger intergenerational persistence and a higher correlation between the ability (and liquidity constraints) of women and that of their families.

The new results come from the quantile regression estimates. The estimated coefficients of schooling years display a *U*-shaped pattern with quantiles. Hence we notice that the increment in earnings for individuals at the lowest and the highest quantiles should be higher in order for them to maintain their quantile in the distribution, after studying one year longer. In other words, we do not observe a simple shift of the distribution of earnings when τ varies, but its scale and its shape also change. In terms of substitutability/complementarity between ability and education, we consider the two factors as substitutes at the lowest quantiles, and complements at the highest ones (corresponding to the most skilled individuals).

The most interesting aspect we find in relation to the specification with the educational dummies is the different behavior displayed by the estimated coefficients of the three levels of education. In fact, the returns to attaining a junior high school diploma fall into the bandwidth of the OLS estimate, for almost every quantile and sample. But this pattern changes when we analyze the behavior of estimated coefficients for "high school" and "tertiary education". The first display a convex pattern across quantiles (education-ability substitutability for low ability individuals, complementarity for high ability individuals), similar to our findings for the year-of-schooling estimates. Hence, the shape of the conditional distribution of earnings changes for those who achieved a high school diploma with respect to the distribution of individuals with a primary education. The estimated coefficients of "tertiary education" display variability across quantiles, too. In particular they tend to increase when the quantile increases (complementarity between ability and education).

We can thus appreciate the capacity of the models based on quantile regression to fully represent the impact of education on the entire distribution of incomes, and their resulting attractiveness for further economic studies. Unfortunately, the treatment is often subject to endogeneity or self-selection, so that not even standard quantile regression (and not only the OLS estimator) provides consistent estimates of the treatment effect. This constitutes the main limit of this work (and also a starting-point for further studies), given the endogeneity problem that typically characterizes the educational acquiring decision. In fact, some estimation methodologies have been proposed in literature, applying quantile regression and instrumental variables jointly (Honoré and Hu 2004; Arias, Hallock and Sosa-Escudero 2001; Chernozhukov and Hansen 2001). A future development of this empirical study on Italian data could therefore be an estimation of education returns by applying one of the cited models, in order to combine the information advantages of a quantile regression approach with those of instrumental variables, thereby allowing us to address endogeneity and heterogeneity problems affecting educational choices.

Another direction of development of this work[14] would entail estimating richer specifications, that include for instance, on the one hand final grades of diploma or laurea, on the other hand some variables capturing the sector or the job qualification of individuals. Controlling for the latter kind of variable would allow the researcher to carry on a study focused on specific sectors, distinguishing between public and private sectors. As far as Italy is concerned, this last aspect is dealt with for instance by Brunello, Comi and Lucifora, 2001, that actually perform OLS end IV estimation. The first proposal can be viewed as referring to that literature, following Griliches, 1977, that uses grades reported in standardized exams (e.g. the IQ test) as variables capturing the unobservable ability. A general "contra" of this approach consists of implicitly defining the ability concept, consequently capturing only a certain "kind of individual ability".

---

[14] I wish to thank an anonymous referee for pointing this out.

APPENDIX

1. - From the Location to the Regression Model: Quantiles via Optimization

Any real valued random variable Y may be characterized by its distribution function:

$$F(y) = \text{Prob}(Y \leq y)$$

For any $0 < \tau < 1$, the τ-th quantile of Y is defined as:

$$Quant(\tau) = \inf\{y | F(y) \geq \tau\} \qquad (32)$$

where the median, $Quant(1/2)$, plays the central role. Like the distribution function, the quantile function provides a complete characterization of the random variable Y.

The quantiles may be formulated as the solution to a simple optimization problem. For any $0 < \tau < 1$, define the following piecewise linear check function:

$$\rho_\tau (u) = u(\tau - 1\{u \leq 0\}) \qquad (33)$$

illustrated in Graph 1.

Minimizing the expectation of $\rho_\tau (Y - \xi)$ with respect to ξ, yields solutions $\xi(\tau)$, the smallest of which is Quant(τ), defined in equation *(2)*. Formally, one can write:

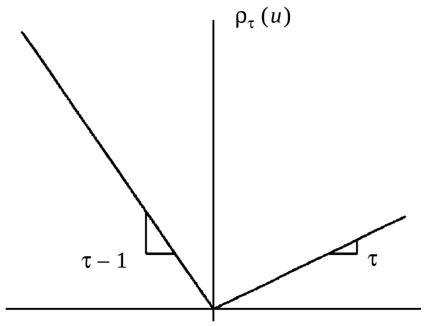
$$\min_{\xi \in \mathbb{R}} E\, [\rho_\tau (Y - \xi)]$$

or, equivalently,

$$\xi^* (\tau) = \text{argmin}_{\xi \in \mathbb{R}}\, E\, [\rho_\tau (Y - \xi)]$$

GRAPH 1

CHECK FUNCTION



with:

(34) $$Quant\,(\tau) = \inf \xi^*(\tau)$$

The sample analogue of $Quant(\tau)$, from a random sample $\{y_1, \ldots, y_n\}$, is called τ-th quantile of the sample, and can be determined by solving:

$$\min_{\xi\in\mathbb{R}} \sum_{i=1}^{n} \rho_\tau\,(y_i - \xi)$$

that is:

$$\hat{\xi}^*\,(\tau) = \text{argmin}_{\xi\in\mathbb{R}} \sum_{i=1}^{n} \rho_\tau\,(y_i - \xi)$$

with:

(35) $$\widehat{Quant}\,(\tau) = \inf \hat{\xi}^*\,(\tau)$$

Then, using:

$$u = -\,|u| \cdot 1\,\{u \leq 0\} + |u| \cdot (1 - 1\{u \leq 0\}) =$$
$$= |u| \cdot (1 - 2 \cdot 1\{u \leq 0\})$$

the check function can be usefully rewritten as follows:

$$\rho_\tau(u) = |u|(1 - 2 \cdot 1\{u \le 0\})(\tau - 1\{u \le 0\})$$
$$= |u|(\tau - 2\tau \cdot 1\{u \le 0\} + 1\ \{u \le 0\})$$
$$= |u|(\tau \cdot 1\ \{u \le 0\} + (1 - \tau) \cdot 1\ \{u \le 0\})$$

from which we get, for instance, that the median corresponds to:

$$\rho_{\frac{1}{2}}(u) = |u|\left(\frac{1}{2}\left(1 - 1\{u \le 0\}\right) + \frac{1}{2} \cdot 1\{u \le 0\}\right) = \frac{1}{2}|u| \qquad (36)$$

while the other quantile are obtained by weighting positive and negative residuals differently (i.e. in an asymmetric way).

To understand the validity of the above formulation, let consider the objective function:

$$Q_\tau(\xi) = \sum_{i=1}^{n}\left(|y_i - \xi| \cdot \tau \cdot 1\{y_i > \xi\} + |y_i - \xi| \cdot (1 - \tau) \cdot 1\{y_i \le \xi\}\right) \qquad (37)$$

The derivative of this function with respect to ξ (except that for $y_i = \xi$, at which the derivative does not even exists) is:

$$\frac{\partial Q_\tau(\xi)}{\partial \xi} = \sum_{i=1}^{n}\left(-\tau \cdot 1\{y_i > \xi\} + (1 - \tau) \cdot 1\{y_i \le \xi\}\right) \qquad (38)$$

If the fraction of the observations with $y_i \le \xi$ is equal to τ, the derivative is zero:

$$\frac{\partial Q_\tau(\xi)}{\partial \xi} = \sum_{i=1}^{n}\left[-\tau\left(1 - 1\{y_i > \xi\}\right) + (1 - \tau) \cdot 1\{y_i \le \xi\}\right] = 0$$

$$\sum_{i=1}^{n}\left(-\tau + 1\{y_i \le \xi\}\right) = 0$$

$$\frac{1}{n}\sum_{i=1}^{n} 1\{y_i \le \xi\} = \tau$$

This last expression represents the condition on residuals. In fact, being ξ_τ the τ-th quantile, the number of negative and zero residuals will be exactly $n\tau$.

While it is more common to define the sample quantiles in terms of the order statistics (constituting a sorted rearrangement of the original sample), their formulation as a minimization problem has the advantage that it yields a natural generalization of the quantiles to the regression context.

Just as the idea of estimating the unconditional mean, viewed as the minimizer:

$$\hat{\mu} = \arg\min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} (y_i - \mu)^2$$

can be extended to estimation of the linear conditional mean function $E(Y|X = x) = x'_i\beta$ by solving:

$$\hat{\beta} = \arg\min_{\beta \in \mathbb{R}^k} \sum_{i=1}^{n} (y_i - x_i'\beta)^2$$

the linear conditional quantile function:

$$Q_\tau(Y|X = x) = x'_i\beta_\tau \tag{39}$$

can be estimated by solving:

$$\hat{\beta}_\tau = \arg\min_{\beta_\tau \in \mathbb{R}^k} \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta_\tau) \tag{40}$$

2. - Quantile Regression and Linear Programming

It can be easily shown that the problem *(18)* is representable as a linear program. As an appealing consequence of this fact, both under a theoretical and practical view point, linear programming can be used to better characterize the solution $\hat{\beta}_\tau$ to the problem *(40)*. In particular, the *Duality Theorem* and the *Equilib-*

rium Theorem of linear programming shed light on the properties of the solution itself and on its sensitivity to the "outliers" (Buchinsky, 1995).

Consider the reference model:

$$y_i = x'_i\beta_\tau + u_i \qquad Q_\tau(y_i|x_i) = x'_i\beta_\tau \tag{41}$$

with $Q_\tau(u_{\tau i}|x_i) = 0$ (by construction), and the minimization problem *(18)*:

$$\min_{\beta_\tau \in \mathbb{R}^k} \sum_{i=1}^{n} \left(\left|y_i - x_i'\beta_\tau\right| \cdot \tau \cdot 1\left\{y_i > x_i'\beta_\tau\right\} + \left|y_i - x_i'\beta_\tau\right| \cdot (1-\tau) \cdot 1\left\{y_i \le x_i'\beta_\tau\right\} \right) \tag{42}$$

Now note that y_i can be rewritten as a function of positive elements alone:

$$y_i = \sum_{j=1}^{k} x_{ij}\beta_{\tau j} + u_{\tau i} = \sum_{j=1}^{k} x_{ij}\left(\beta_{\tau j}^1 - \beta_{\tau j}^2\right) + \left(u_{\tau i}^+ - u_{\tau i}^-\right) \tag{43}$$

where $\beta_{\tau j}^1 \ge 0$, $\beta_{\tau j}^2 \ge 0$ (with $j = 1, \ldots, k$), and $u_{\tau j}^+ \ge 0$, $u_{\tau j}^- \ge 0$ ($i = 1, \ldots, n$).

Therefore, problem *(18)*, in terms of the component of *(43)*, takes the following formulation:

$$\begin{cases} \min \left[\tau\iota' u^+ + (1-\tau)\iota' u^-\right] \\ y = X\beta_\tau + u_\tau^+ - u_\tau \\ \left(\beta_\tau, u_\tau^+, u_\tau^-\right) \in \mathbb{R}^k \times \mathbb{R}_+^{2n} \end{cases}$$

or, equivalently, in a more compact (matrix) notation:

$$\begin{cases} \min_z \; c'z \\ Az = y \\ z \ge 0 \end{cases}$$

where $A = (X, -X, I_n, -I_n)$ is a matrix $n \times (2n + 2k)$; $y = (y_1, \ldots, y_n)'$ is a vector $n \times 1$ collecting the observations on the dependent variable; $z = \beta_\tau^{1'}, \beta_\tau^{2'}, u_\tau^{+}, u_\tau^{-})'$ is a vector $(2k + 2n) \times 1$; $c = (0', 0', \tau\iota', (1 - \tau)\iota')'$; $X = (x_1, \ldots, x_n)'$ is a matrix $n \times k$ with the i-th row given by x_i', with i=1, ..., n; I_n is an identity matrix with dimension n; 0' is a vector $k \times 1$ of zeros and ι is a vector $n \times 1$ of ones. In the LP system of equations, u_τ^{+} and u_τ^{-} are used in such a way to make the constraints binding, so that the program is expressed in its standard form, according to the transformation rules of LP. Generally, the constraints would be inequalities, in the canonic form of the program: in fact, for the observations on the dependent variable lying below the hyperplane, $y_i - x_i'\beta_\tau < 0$; while for the observations above it, $y_i - x_i'\beta_\tau > 0$. In other words, the term $u_{\tau i}$ is not constrained in its sign. Such a program represents a Primal Problem, while its Dual Problem is defined as:

$$\begin{cases} \max\limits_{w} \ w'y \\ w' \ A \leq c' \end{cases}$$

It is approximately equivalent to the FOC's of the minimization problem (eq. *(20)*) Buchinsky (1995).

BIBLIOGRAPHY

ANTONELLI G., *Risorse umane e redditi da lavoro,* Milano, F. Angeli, 1985.

ARIAS O. - HALLOCK K.F. - SOSA-ESCUDERO W., «Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data», *Empirical Economics,* vol. 26, 2001, pp. 7-40.

ASHENFELTER O. - ROUSE C., «Income, Schooling and Ability: Evidence from a New Sample of Identical Twins», *Quarterly Journal of Economics,* vol. 113, 1998, pp. 253-84.

BANCA D'ITALIA, *Indagine sui bilanci delle famiglie italiane,* distribuzione elettronica dei microdati, Roma, Banca d'Italia, 1986, 1991, 1993, 1995, 1998, 2000.

BECKER G.S., *Human Capital and the Personal Distribution of Income,* Michigan, University of Michigan Press, Ann Arbor, 1967.

BRUNELLO G. - MINIACI R., «The Economic Returns to Schooling for Italian Men. An Evaluation Based on Instrumental Variables», *Labour Economics*, vol. 6, no. 4, 1999, pp. 509-19.

BRUNELLO G. - COMI S. - LUCIFORA C., «The Returns to Education in Italy: A New Look at the Evidence», in HARMON C. - WALKER I. - WESTERGARD-NIELSEN N. (a cura di), *The Returns to Education in Europe,* Cheltenham (UK) Edward Elgar, 2001.

BUCHINSKY M., «Changes in the US Wage Structure 1963-1987: Application of Quantile Regression», *Econometrica*, vol. 62, 1994, pp. 405-58.

— —, «Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study», *Journal of Econometrics*, vol. 68, 1995, pp. 303-38.

— —, «Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research», *Journal of Human Resources*, vol. 33, 1998, pp. 88-126.

CANNARI L. - D'ALESSIO G., «Il rendimento dell'istruzione: alcuni problemi di stima», Roma, Banca d'Italia, *Temi di Discussione*, no. 253, 1998.

CANNARI, L. - PELLEGRINI G. - SESTITO P., «Redditi da lavoro dipendente: un'analisi in termini di capitale umano», Roma, Banca d'Italia, *Temi di Discussione,* no. 124, 1989.

CARD D., «Earnings, Schooling, and Ability Revisited», Cambridge, MA, *Working Paper,* no. 4832, NBER, 1994.

— —, «Using Geographic Variation in College Proximity to Estimate the Return to Schooling», in *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, Toronto, University of Toronto, 1995, pp. 201-22.

— —, «Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems», *Econometrica,* vol. 69, 2001, pp. 1127-60.

CARD D. - KRUEGER A.B., «Does School Quality Matter? Return to Education and the Characteristics of Public Schools in the United States», *Journal of Political Economy,* vol. 100, 1992, pp. 1-40.

CHERNOZHUKOV V. - HANSEN C., «An IV Model of Quantile Treatment Effect», *Econometrica*, vol. 73, no. 1, 2005.

COLUSSI A., «Una analisi cross section del tasso di rendimento dell'istruzione in Italia», *Politica Economica*, vol. 13, no. 2, 1997.

FITZENBERGER B. - KURZ C., «New Insights on Earnings Trends Across Skill Groups and Industries in West Germany», *Empirical Economics*, vol. 28, 2003, pp. 479-514.

FLABBI L., «Investire in istruzione: È meglio per lui o per lei? Stima per genere dei rendimenti dell'istruzione in Italia», *Working Paper*, no. 8, Milano, Università degli studi di Milano - Bicocca, 1997.

— —, «Returns to Schooling in Italy: OLS, IV and Gender Differences», *Working Paper*, no. 1, Milano, Università Bocconi, 1999.

GRILICHES Z., «Estimating the Returns to Schooling: Some Econometric Problems», *Econometrica*, vol. 45, 1997, pp. 1-22.

HAUSE J.C., «Earnings Profile: Ability and Schooling», *Journal of Political Economy*, vol. 80, 1972, pp. 108-38.

HILLIER F.S. - LIEBERMAN G.J., *Introduction to Operations Research*, New York, McGraw-Hill, 1990.

HOLLAND P.W., «Statistics and Causal Inference», *Journal of the American Statistical Association*, vol. 81, 1986, pp. 945-60.

HONORÉ B.E. - HU L., «On Performance of Some Robust Instrumental Variables Estimators», *Journal of Business and Economic Statistics*, vol. 22, no. 1, 2004, pp. 30-9.

ICHINO A., «Il problema della causalità. Una introduzione generale e un esempio», in *Manuale di economia del lavoro*, Bologna, il Mulino, 2001, pp. 457-83.

ICHINO A. - WINTER-EBMER R., «Lower and Upper Bounds of Returns to Schooling: An Exercise in IV Estimation with Different Instruments», *European Economic Review*, vol. 43, 1999, pp. 889-901.

KOENKER R., «Short Course on Quantile Regression», *Technical Report*, Urbana-Campaign, University of Illinois, 2003.

KOENKER R. - BASSETT G., «Regression Quantiles», *Econometrica*, vol. 46, 1978, pp. 33-50.

LUCIFORA C. - REILLY B., «Wage Discrimination and Female Occupational Intensity», *Labour*, vol. 4, no. 2, 1990, pp. 147-68.

MARTINS P.S. - PEREIRA P.T., «Does Education Reduce Wage Inequality? Quantile Regression and Evidence From 16 Countries», *Labour Economics*, vol. 11, 2004, pp. 355-71.

MINCER J., *Schooling, Experience and Earnings*, New York, Columbia University Press, 1974.

MWABU G. - SCHULTZ T.P., «Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa», *American Economic Review*, vol. 86, no. 2, 1996.

PARK J.H., «Returns to Schooling: A Peculiar Deviation from Linearity», Princeton, NJ, Princeton University, *Working Paper*, no. 335, 1994.