Central European Journal of Economic Modelling and Econometrics

Returns to Education and Gender Wage Gap Across Quantiles in Italy

Marilena Furno*

Submitted: 21.10.2019, Accepted: 10.04.2020

Abstract

Various quantile regression approaches are implemented to analyze the characteristics of Italian data on earnings in the tails. A changing coefficients pattern across quantiles shows increasing returns to education along the wage distribution. A quantile decomposition approach shows that higher education grants higher return at all quantiles, thus implying additional, non-linear returns to higher education throughout the entire pattern of the earning distribution. Wage gender gap displays a decreasing pattern across quantiles, and it does not disappear at the higher quantiles. The southern workers penalty decreases across quantiles as well for highly educated workers.

Keywords: quantile regression, decomposition, returns to education, gender wage gap

JEL Classification: C51, G32

*Università degli Studi di Napoli "Federico II"; e-mail: marfurno@unina.it;
ORCID: 0000-0001-6635-8476

1 Introduction

There is a wide literature relating earnings to education and their positive correlation is widely studied and well established (Card, 2001; Blundell et al., 2001; Meghir and Palme, 2000). Higher levels of education usually translate into better employment opportunities and higher earnings. This is an important incentive for individuals to pursue education and training. In many countries, Italy included, gender wage gaps in earnings persist regardless of age, level of education or field of study (OECD, 2019). However, there is no educational gap in Italy: the number of women enrolled in college is greater than the number of men, and there is no sign of segregation in disciplines (Bettio and Villa, 1999). The 2015 OECD country note, for fifteen years old students, states: In Italy, equity in education outcomes is above the OECD average, as 10% of the variation in student performance in mathematics is attributed to differences in students' socioeconomic status (OECD, 2015). For the higher degrees, ISTAT (Italian Statistical Institute) in year 2016 reports 131673 men versus 180118 women with higher degrees, pooling together 3 years and 5 years higher degrees, i.e. Laurea Triennale and Laurea Magistrale. Nevertheless, since women are underrepresented in top-skilled jobs, this implies that women are generally overqualified for the work positions they occupy. Aside from wage differences, the gender gap causes longer job search, longer time to ascend the wage ladder (Sulis, 2007), higher thresholds for promotion – all of which constitute women glass ceiling.
Estimates of returns and of gender wage gap in Italy vary significantly in the literature and yield quite heterogeneous results according to the data set and the methodology. On average, estimates of returns sizably differ according to the selected estimator – ordinary least squares versus instrumental variables, while for the gender wage gap some studies find an opposite sign in Italy – a premium for working women instead of a penalty.
This study investigates the above issues for the Italian economy in year 2014, looking at the behavior of a wage equation not only on average but also across quantiles. We find evidence of increasing returns to education – returns that further increase with the completed degree; of a never vanishing gender gap in earnings; of southern regions penalty that becomes a premium for the highly educated workers at the very top quantile. The different estimators implemented help to pinpoint these issues. A wage equation is computed by ordinary least squares (OLS) and at various quantiles by quantile regressions (Koenker, 2005). Estimating a single rate of return is not appropriate if returns differ by education and/or wage level (Buchinsky, 1998; Dickson and Harmon, 2011). An evaluation on average, like the one provided by OLS, is not particularly informative in this case. Besides the quantile regressions estimator, additional quantile regression-based estimators are here implemented to analyze the data. These approaches, far from being a mere technicality, allow to uncover relevant aspects of the model.
The initial step is to implement the simple quantile regression estimator at different quantiles, analyzing the pattern of returns to education, of gender wage gap and

of all the other coefficients of the equation at the various wage levels. Returns to education increase with the quantile. This can be due to the amount of qualification required: at low wages education might improve earnings only to a limited extent in little payed, low qualified jobs. Vice versa highly rewarded jobs generally require higher qualification, and education can make the difference. In addition, returns to education build up slowly, since it takes some time to progress in a career. This implies that returns to schooling grow with wage quantiles since they increase while moving ahead in a career: education increases wages later more than earlier in life. The direct proportionality between returns to schooling and quantiles could mirror the unobservable ability to bring to fruition education: this ability is more effective in talented workers. Gender wage gap and southern workers penalty decrease at the higher quantiles.

In general, the presence of changing coefficients across quantiles signals heterogeneity in the data, i.e. heteroskedastic errors, and one possible sources of heterogeneity is the individual ability. Inequality may be due to many other elements: the major of the degree, the quality of the institution offering the course, gender, ethnicity, immigration, region of residence and so forth.

Focusing on ability, the quantile regression analysis is complemented by two instrumental variable quantile regression estimators. There is a wide literature on the link between education and the error term of the wage equation, link that causes bias and inconsistency in the results (Card, 2001). The link may be caused by omitted variables, and a typical example of omitted variables is individual ability: talented workers turn out to be more educated and more productive; they may be able to convert schooling into human capital more efficiently. However individual talent cannot be measured and its effect on wages ends up in the error term, thus causing endogeneity of education. The instrumental variables account for the endogeneity of education. A generally accepted choice of instruments is related to family background covariates, to parental education (Blundell et al., 2005). This choice relies on the idea that parents background is correlated with their descendant education but is uncorrelated with offspring earnings. The direction of the bias is unclear: impatient workers have a high discount rate, since they evaluate the direct and indirect costs of education more than the benefit of increased future earnings (the indirect costs of education include foregone income during the school period); vice versa for workers with a low discount rate (Harmon et al. 2003). In our findings the quantile instrumental variable approach increases the reward to education at all quantiles when compared to the initial simple quantile regressions. In the literature, this is a known result when comparing OLS versus the standard instrumental variables estimates. We find that it applies to the quantile regression framework as well: quantile instrumental variables estimate of returns to education increase with respect to the standard quantile regressions results at all quantiles. Endogeneity causes the under-estimation of returns to education in both the least squares and the quantile regression framework. The bias shows a disproportionate relevance of workers with

high discount rates on future earnings. This complements the previous analyses of Italian data (see for instance Brunello et al., 2001).
Finally, following Card and DiNardo (2002) study on technological skill bias, schooling is dichotomized into lower and higher degrees of education and workers are accordingly split in two different groups. The wage differential between these two groups is decomposed into difference in the covariates and discrepancy in the coefficients. A disparity in the covariates of the two groups, workers with higher versus workers with lower education, measures earning variations due to differences in the characteristics of the two groups. For instance, age or region of residence may systematically differ in the two groups and cause wage differentials. Vice versa, a difference in the coefficients provides the sheer compensation to higher education. A quantile decomposition approach compares these two groups of workers at various quantiles to take apart the role of covariates and the role of coefficients at many quantiles. The decomposition shows that the difference in earnings between the two groups of workers is mostly due to the coefficients at all quantiles: wage differentials are mostly due to differences in the returns to schooling. This shows that education premia not only change across quantiles, but they also change according to the completed degree, implying the presence of a non-linear pattern.
A related approach decomposes wage differentials meanwhile accounting for the endogeneity of education, i.e. by including an instrumental variable so that the coefficients effect – the higher education premium – is measured net of endogeneity. These two approaches provide similar results, with the latter yielding a smaller premium with respect to the former. The effect of the instrumental variable is to smooth the pattern of returns across quantiles. In sum, instrumental variable at the quantiles increases return to education with respect to the simple quantile regression and the decomposition at the quantiles shows higher education premia related to the completed degree, signaling non-linearity in returns to education. The gender wage gap reduces across quantiles but does not disappear, while the southern workers penalty becomes a premium at the 90^{th} quantile for the highly educated ones.

2 Estimates of returns to education and gender wage gap in Italy

Several empirical studies analyze the Survey of Household Income and Wealth (SHIW), held by Banca d'Italia, to estimate returns to schooling and gender wage gap in Italy (the English version of the questionnaire and the data set are available from the Banca d'Italia website).
In the empirical analysis of SHIW data, a measure of returns to education and of the gender wage gap is far from being univocal and at times the latter has been even reversed. Cannari and D'Alessio (1998), using 1993 SHIW data, and choosing family background variables as instruments, obtain an estimate of returns to education close

to 7%. Colussi (1997) achieve an estimate of 6.6%, with the same data and similar instrumental variables. Flabbi (1999) estimates the returns to schooling for women and men separately (1991 SHIW data). The estimated coefficients implementing instrumental variables turned out to be higher for men, 0.62 versus 0.56 for women, while the OLS returns were 0.22 for women and 0.17 for men. Brunello and Miniaci (1999), using 1993 and 1995 SHIW data, with instrumental variables (IV) related to family background, compute an OLS estimate of male education returns equal to 4.8%, and an IV estimate of 5.7%. In our model, the IV estimates of returns to education turn out to be higher than the simple OLS coefficients as well. We show that this result holds also in the quantile estimated regressions, with simple quantile regressions providing smaller premia that the IV quantile regressions estimates.

Comi and Lucifora (2001), in the same data set find higher returns for women even on an IV basis, thus reversing the sign of the wage gap. Giustinelli (2004), using SHIW data from 1990 to 2000, finds higher women's returns as well. These results are at odds with the general findings on gender wage gap in the literature. For instance, the European Commission (2019) states that: "There are considerable differences between EU countries. The gender pay gap ranges from less than 8% in Belgium, Italy, Luxembourg, Poland and Romania to more than 20% in Czechia, Germany, Estonia and United Kingdom." Indeed, our results find a premium for male workers in all the quantile regression estimators implemented, generally ranging from 8% to 12% at the lower quantiles and dropping to 5% or 7% at the top quantiles according to the selected approach, while our standard IV estimated (average) male premium is 6%, not too far from Colussi (1997) and Brunello and Miniaci (1999) results. The 8% and 20% European Commission benchmark rates are average (OLS based) rates. We find a range of estimated gender pay gap depending upon the quantile estimator implemented and upon the selected quantile.

Ciccone et al. (2006) in the SHIW waves 1987, 1995, 1998, 2000 find higher returns for higher degrees, further increasing in the southern regions, and we further analyze this issue. Our findings show increasing returns with the degrees as well, combined with a southern region penalty that is sizable at the lower wages but that decreases across quantiles. Zizza (2013), analyzing 1995 to 2008 SHIW data, estimates a raw gender wage gap around 6%, that increases to 11%-12% in the extended version of the model. We find a male premium in all the estimators implemented, higher at the lower wages but decreasing across quantiles. Furno (2014), using 1989-2010 SHIW data, provides evidence of a gender wage gap amplified by regional differences, with southern regions offering less favorable conditions to women. The regional penalty for the southern regions is confirmed in the following analysis.

3 The quantile regression approach

The general linear regression model $y_i = x_i\beta + e_i$ – where y_i is the dependent variable, x_i is the row vector of explanatory variables, β the vector of regression coefficients and e_i is the i.i.d error term – can be estimated at different points of the conditional distribution. In case of OLS, the equation is estimated at the conditional mean. If the goal is to estimate the model in the tails, as in the quantile regression estimator, an asymmetric weighting system is introduced to drive the estimated regression above or below the conditional mean. The quantile regression objective function at the θ quantile (Koenker, 2005) is given by

$$\sum_{y_i \le x_i\beta} (1-\theta)\,|y_i - x_i\beta| + \sum_{y_i \ge x_i\beta} \theta\,|y_i - x_i\beta| = \sum_i \{\theta - 1(y_i \le x_i\beta)\}\,|y_i - x_i\beta|\,. \quad (1)$$

For instance, to estimate the 75^{th} quantile regression, i.e. the regression explaining the 3rd quartile of the dependent variable, the estimated equation will be characterized by 75% of the residuals below the 3rd quartile, and 25% of the residuals above it. It represents the regression passing through the third quartile of the conditional distribution of the dependent variable, given the selected covariates. This result is achieved by using asymmetric weights that assign the value $\theta = 0.75$ to the larger observations to attract the estimated equation upward, and $1 - \theta = 0.25$ to the remaining data. If the conditional distribution of the dependent variable shows constant variability in the sample, the regression coefficients do not change across quantiles – with the sole exception of the intercept – and for any pair of quantiles θ_j and θ_k it is $\beta(\theta_j) = \beta(\theta_k)$. The intercept computes the chosen quantile of the dependent variable when all the other coefficients are set to zero, like in the OLS case where the intercept computes the sample mean of the dependent variable. When the dispersion of the dependent variable conditional to the covariates is not constant in the sample, the errors are heteroskedastic. This causes the regression coefficients estimated at a given quantile, $\beta(\theta)$, to change from one quantile to another: looking at two different quantiles θj and θ_k the estimated coefficients will differ, $\beta(\theta_j) \neq \beta(\theta_k)$. The implication is that the explanatory variables have a different impact on the dependent variable, an impact that changes with the quantiles.

3.1 Instrumental variables at the quantiles

In quantile regression instrumental variables can be implemented following two approaches. Amemiya (1982) discusses a two-stage median regression. In the first step the endogenous explanatory variables are regressed on a set of instruments in order to compute the fitted values of the regression. In the second stage these fitted values replace the endogenous explanatory variable (IVQ from now on).
An alternative method is discussed in Chernozhukov and Hansen (2005, 2006, 2008) (CH_IVQ from now on). The first step is to compute the fitted values of the

regression just as in the previous IVQ estimator. Next a grid search on the coefficient of the endogenous explanatory variable allows to compute the variable *diff* $= y - \alpha($*fitted values*$)$. The variable *diff* is regressed on the exogenous variables and the instrument. A Wald test verifies the null that the instrumental variable does not have any additional overlooked impact. The null implies that at the true parameter the impact of the instrumental variable is zero. Failure to reject the null validates the CH_IVQ estimates.

3.2 Decomposition at the quantiles

Finally, we are interested in checking if the model changes between subgroups of the data. Consider two subsets – non-educated/educated, or past/present, or more in general untreated/treated – indexed respectively by 0 and 1. Having estimated the model in the two different subsets, 0 and 1, the changes from one group to the other, $\hat{y}_0 - \hat{y}_1$, where $\hat{y}_0$ and $\hat{y}_1$ are the fitted values computed in group 0 and 1, can be computed by

$$\hat{y}_0 - \hat{y}_1 = x_0\hat{\beta}_0 - x_1\hat{\beta}_1 \tag{2}$$

with $\hat{\beta}_0$ and $\hat{\beta}_1$ being the estimated OLS coefficient. Next, by adding and subtracting the term $x_1\hat{\beta}_0$, which multiplies the group 1 covariates by the group 0 estimated coefficients, the difference can be decomposed into

$$\hat{y}_0 - \hat{y}_1 = x_0\hat{\beta}_0 - x_1\hat{\beta}_1 = x_0\hat{\beta}_0 - x_1\hat{\beta}_0 + x_1\hat{\beta}_0 - x_1\hat{\beta}_1 = (x_0 - x_1)\hat{\beta}_0 + x_1\left(\hat{\beta}_0 - \hat{\beta}_1\right) \tag{3}$$

which shows how the difference between two groups can be split into changes in the covariates $(x_0 - x_1)$ – i.e. changes in the independent variables from one group to the other – and changes in the coefficients $(\hat{\beta}_0 - \hat{\beta}_1)$ – that are not explained by the variables since they are caused by a change in the model coefficients. In the above decomposition, the term $x_1\hat{\beta}_0$ is the counterfactual: it measures the hypothetical value of y_1 in case the regression coefficients do not change from group 0 to group 1. It is usually estimated on average using the OLS estimates of β_0 multiplied by the sample means of the covariates in group 1, $Ex_1\hat{\beta}_0$. By replacing the fitted values in each subset $\hat{y}$ with their sample averages, the Oaxaca-Blinder decomposition at the mean is given by

$$Ey_0 - Ey_1 = (Ex_0 - Ex_1)\,\hat{\beta}_0 + Ex_1\left(\hat{\beta}_0 - \hat{\beta}_1\right). \tag{4}$$

The average difference of the dependent variable in the two groups is decomposed into the average difference in covariates times the group zero OLS estimated coefficients, plus the difference in coefficients as computed by the OLS/average regressions in each group and multiplied by Ex_1.
In the quantile regression decomposition, the above model is computed at different

quantiles θ to identify changes in the covariates and in the coefficients at the center and in the tails of the distribution.

$$\begin{aligned}\hat{y}_0(\theta) - \hat{y}_1(\theta) &= x_0\hat{\beta}_0(\theta) - x_1\hat{\beta}_1(\theta) = \\ &= x_0\hat{\beta}_0(\theta) - x_1\hat{\beta}_0(\theta) + x_1\hat{\beta}_0(\theta) - x_1\hat{\beta}_1(\theta) = \\ &= (x_0 - x_1)\hat{\beta}_0(\theta) + x_1\left(\hat{\beta}_0(\theta) - \hat{\beta}_1(\theta)\right).\end{aligned} \tag{5}$$

Chernozhukov et al. (2013) discuss the conditions required for valid first stage quantile regression estimates to compute the counterfactuals $x_1\hat{\beta}_0(\theta)$. They show that bootstrap is a valid approach to estimate standard errors and to make inferences about the counterfactuals; this allows us to assess the statistical relevance of the quantile regression decomposition. Quantile regression decomposition confirms, at each quantile, whether a discrepancy between actual and counterfactual values is statistically relevant, and if such a discrepancy is stable or changes across quantiles. In a sample divided into groups 0 and 1, the quantile decomposition approach can be summarized as follows:

i) $m = 100$ values are drawn from a uniform distribution $(0,1)$; for each draw from this distribution the corresponding quantile regression is computed at the quantile θ_j for that particular draw, $j = 1, ..., m$; 100 vectors of estimated coefficients $\beta_0(\theta_j)$ are provided by minimizing the quantile regression objective function $\sum_{n_0}\{\theta_j - 1(y_{0i} \leq x_{0i}\beta)\}\,|y_{0i} - x_{0i}\beta|$ in group 0; 100 estimated coefficients $\beta_1(\theta_j)$ are separately computed in group 1, by minimizing $\sum_{n_1}\{\theta_j - 1(y_{1i} \leq x_{1i}\beta)\}\,|y_{1i} - x_{1i}\beta|$ (Machado and Mata, 2005);

ii) $m = 100$ random samples with replacement, each of size n, are drawn from the group 0 subset, yielding m random samples of x_0, from now on $\hat{x}_0$, to compute the distribution of the variable $\hat{y}_{0/0} = \hat{x}_0\hat{\beta}_0(\theta_j)$ via the empirical distribution function of $\hat{y}_{0/0}$, $\hat{F}_{0/0}(y) = \sum_j(\theta_j - \theta_{j-1})1(\hat{y}_{0/0} \leq y)$, where θ_j and θ_{j-1} are two adjacent quantiles; $m = 100$ random samples with replacement are drawn from group 1 covariates, from now on $\hat{x}_1$, to compute $\hat{y}_{1/1} = \hat{x}_1\hat{\beta}_1(\theta_j)$ and its empirical distribution $\hat{F}_{1/1}(y)$ (Melly, 2006);

iii) the counterfactual can be computed by $\hat{y}_{0/1} = \hat{x}_1\hat{\beta}_0(\theta_j)$ and its empirical distribution function by $\hat{F}_{0/1}(y) = \sum_j(\theta_j - \theta_{j-1})1(\hat{y}_{0/1} \leq y)$;

iv) the difference in the covariates $\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)$, and the difference in the coefficients $\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)$, can now be estimated at various quantiles;

v) a bootstrap approach allows to compute standard errors and to implement inference.

To compute the decomposition at several quantiles, the comparison between y_0 and y_1 becomes a comparison between their unconditional distributions:

$$\hat{F}_{0/0}(y) - \hat{F}_{1/1}(y) = \left[\hat{F}_{0/0}(y) - \hat{F}_{0/1}(y)\right] + \left[\hat{F}_{0/1}(y) - \hat{F}_{1/1}(y)\right]. \tag{6}$$

The results of the quantile regression decomposition, obtained by analyzing the difference between observed and counterfactual distributions at various quantiles, cannot be revealed by a standard average decomposition analysis. For instance, quantile regression decomposition may reveal the case of discrepancies at low quantiles attributable to the covariates, i.e. explained by changes in the covariates, while at the median and top quantiles a divergence is linked to the coefficients, i.e. is not explained by the regression model. In general, the source of discrepancy in the covariates and/or the coefficients may differ across quantiles. It may be the case that an increasing effect of the covariates across quantiles counterbalances a decreasing coefficients effect, thus providing a stable overall discrepancy. Alternatively, the two effects may cancel one another at some/all quantiles. In sum, the source of discrepancy in the covariates and/or the coefficients may differ across quantiles.

4 Data set and model

The SHIW survey in year 2014 covers 8156 households. It provides detailed information on demographic variables, income, consumption, wealth. It covers a representative sample of the Italian resident population. Sampling is carried out in two stages: the first covers the selection of municipalities, the second the selection of households.

In what follows people younger than 20 or older than 60 are excluded from the sample. The purpose is to analyze individuals that have completed their education and that are not yet retired. This reduces the sample to 9313 individuals. Next, a measure of the hourly wage is computed by the ratio between annual earnings and annual worked hours, and there are 5160 workers earning a wage and 4682 of them have an age comprised between 20 and 60. Finally, those with an hourly wage below 4 euros are excluded from the analysis, and this leads to our sample of size 4267.

The standard Mincer (1974) equation, where the log of wage is function of experience and education, is expanded to include additional explanatory variables – gender, family characteristics, region of residence. The dependent variable is the logarithm of hourly wage in euros, net of taxes and social security contributions. The explanatory variables age and age squared are used to proxy experience; education takes values from 0 to 21 to indicate years of achieved education (respectively: no education, elementary, junior high, vocational school, high school, 3-years university degree, 5-years university degree, doctorate); a regional dummy assumes unit value for 1155 workers living in the South, and reflects the inertia of the southern economy; a dummy for marital status with unit value for 2624 married or cohabiting workers; a dummy for children assuming zero value for 2647 childless workers out of 4267; a dummy for gender assuming unit value for 2350 male workers. Italy presents a wide gap in women work participation, and the lower share of employed women can be ascribed to the subset of women with lower education. In this data set the numbers of working women are 1917 out of 4267, 1289 educated to at least high school degree and 424 with at most the junior high degree, against respectively 1228 and 867 in the working men subset. This may be because less educated, formally unemployed women may be engaged in precarious jobs such as call center or babysitting, that are excluded from the analysis. These variables mostly consider worker's characteristics as opposed to variables for job quality or business cycle, often considered by some extended versions of the Mincer equation. Table 1 provides the summary statistics of the variables in the entire sample, at the bottom of the table, and in the subset of observations selected for the analysis, in the top section of the table. In the entire sample the wage distribution shows a longer left tail, a larger number of women, of children and of people living in the South. The semi-logarithmic model is defined as

$$wage = \alpha\ education + \delta_1\ kids + \delta_2\ male + \delta_3\ married + \delta_4\ south + \delta_5\ age + \delta_6\ age^2. \tag{7}$$

Equation (7) is computed by OLS and by quantile regression. OLS yields an equation passing through the conditional mean, while the quantile regressions compute the line passing through different points of the conditional distribution. This allows to compute the impact of the regressors at different wage levels such as at highest, the lowest as well as the median wage.

Table 2 reports the OLS and the quantile regression estimates of (7). The estimated quantiles are the 10^{th}, the 25^{th}, the median, the 75^{th} and the 90^{th}. Returns to education increase with the quantiles ranging from 2.9% to 5.5%: for workers earning low wages, i.e. at the 10^{th} quantile, there is a stable 2.9% increase for each year of schooling, while at the higher wages, the 90^{th} quantile, each additional year of education grants a 5.5% premium. The impact of experience as mirrored by *age* and age^2 decreases across quantiles, while the southern regions penalty is less damaging at the higher quantiles and becomes non-significantly different from zero at the 90^{th} quantile. The gender wage gap shows an inverse u-shaped pattern reaching the peak at the first quartile.

The presence of changing coefficients across quantiles signals heterogeneity in the earning profiles (Blundell et al. 2001). The OLS estimates assume intermediate values, not far from the median regression results.

There is a wide literature focusing on the presence of an endogenous explanatory variable in equation (7). The focus is on the correlation between individual ability with both wages and education. Four additional variables are introduced to explain the endogenous right-hand side variable, and workers *education* becomes function of mother and father education of the household head, together with mother and father education of the household spouse. The idea behind is that parents' education is correlated to young people education but uncorrelated with offspring earnings.

Table 1: Summary statistics of the variables

	mean	s.d.	median	skewness	n
			in the selected sample		
wage	2.192	0.372	2.158	1.01	4267
age	44.261	9.799	46	−0.51	4267
education	12.159	3.686	13	0.21	4267
kids	0.1727	0.378	0	1.73	4267
married	0.6149	0.486	1	−0.47	4267
men	0.5507	0.497	1	−0.20	4267
south	0.2706	0.444	0	1.03	4267
			in the entire sample		
wage	2.079	0.552	2.13	−1.07	5160
age	48.292	23.27	51	−0.248	19400
education	9.138	4.89	8	0.021	19400
kids	0.279	0.448	0	0.983	19400
married	0.499	0.500	0	0.001	19400
men	0.479	0.499	0	0.082	19400
south	0.367	0.482	0	0.551	19400

Table 2: Simple OLS and quantile regression estimates

	0.10	0.25	0.50	0.75	0.90	OLS
	wage	*wage*	*wage*	*wage*	*wage*	*wage*
education	0.0299 (11)	0.0341 (21)	0.0385 (30)	0.0451 (25)	0.0549 (23)	0.039 (28)
kids	−0.075 (−2.15)	−0.106 (−4.51)	−0.077 (−4.74)	−0.068 (−3.87)	−0.0808 (−2.63)	−0.085 (−4.82)
male	0.0882 (6.02)	0.112 (11)	0.095 (9.94)	0.085 (7.24)	0.0723 (3.15)	0.077 (7.51)
married	0.055 (2.27)	0.0477 (2.77)	0.0534 (5.51)	0.0956 (5.79)	0.1034 (4.69)	0.075 (6.12)
south	−0.122 (−5.23)	−0.104 (−7.95)	−0.094 (−7.72)	−0.0541 (−3.36)	*0.0274* (1.04)	−0.050 (−4.50)
age	0.0180 (2.27)	*0.011* (1.63)	0.0119 (2.86)	*0.0064* (1.43)	*0.0084* (1.84)	0.011 (2.56)
age^2	*-0.00011* (−1.30)	*-0.00003* (−0.41)	*-0.00003* (−0.71)	*0.00003* (0.68)	*0.00002* (0.42)	*-0.00003* (−0.60)
constant	0.872 (5.16)	1.105 (7.23)	1.203 (12)	1.368 (15)	1.364 (12)	1.20 (12)

Note: Student-t values in parenthesis, in italics are the non-statistically significant estimates. Sample of size $n = 4267$.

Next, the Amemiya (1982) two-stage median regression approach is implemented. In the first stage workers education is computed as a function of parents' education in order to provide the variable $fitted\ education = \widehat{\widehat{education}}$. The second stage computes the equation

$$wage = \gamma\ \widehat{education} + \delta_1\ kids + \delta_2\ male + \delta_3\ married + \delta_4\ south + \delta_5\ age + \delta_6\ age^2. \tag{8}$$

Table 3 reports the results of equation (8) at the selected quantiles. Returns to education as estimated by standard IV and by IVQ are larger than respectively the OLS and the quantile regressions of Table 2. In IVQ returns sharply raise at the top quantiles. Purging endogeneity by means of family background variables sizably increases returns to education. The male premium is comparable with the previous results at the lower quantile, it is about half the estimates of Table 2 at the intermediate quantiles and becomes irrelevant at the top quantile, thus showing a decreasing in place of an inverse u-shaped pattern. The southern regions penalty is less severe and becomes a premium at the top quantiles. Experience as measured by *age* and age^2 is not statistically significant at most quantiles. In the standard IV estimates experience and southern penalty are not statistically significant, while the remaining estimated coefficients assume intermediate values, not far from the median IVQ results.

Table 3: Instrumental variable results

	0.10	0.25	0.50	0.75	0.90	IV
	wage	*wage*	*wage*	*wage*	*wage*	*wage*
$\widehat{education}$	0.502	0.0465	0.0548	0.0786	0.1061	0.0623
	(7.01)	(6.79)	(8.91)	(11.4)	(7.64)	(11)
kids	−0.1273	−0.1585	−0.1071	−0.0967	−0.1398	−0.1297
	(−1.82)	(−3.89)	(−4.00)	(−3.64)	(−2.80)	(−4.04)
male	0.0860	0.0818	0.0484	0.0450	*0.0197*	0.0458
	(3.51)	(5.02)	(2.81)	(3.37)	(0.57)	(3.54)
married	0.1087	0.0611	0.0804	0.1051	0.1136	0.0842
	(1.89)	(2.04)	(3.87)	(4.56)	(2.28)	(3.11)
south	−0.0740	−0.0798	−0.0379	0.0494	0.1150	*0.0078*
	(−2.72)	(−5.15)	(−2.82)	(2.99)	(2.37)	(0.53)
age	*0.00449*	*0.0016*	*0.0009*	*-0.0041*	0.0264	*0.00069*
	(0.43)	(0.24)	(0.20)	(−0.63)	(−2.79)	(0.11)
age^2	*0.00004*	*0.00007*	0.00010	0.00021	0.00054	*0.00012*
	(0.41)	(0.92)	(1.91)	(2.71)	(4.63)	(1.77)
constant	0.8468	1.1844	1.234	1.1219	1.326	1.135
	(3.53)	(7.19)	(8.45)	(7.14)	(4.74)	(8.17)

Student-t values in parenthesis, in italics are the non-statistically significant estimates. Sample of size $n = 3182$.

In the alternative CH_IVQ method the first step is to compute the variable $\widehat{education}$ just as in the previous IVQ estimator. Next a grid search on α that is the coefficient of the endogenous explanatory variable in equation (7), allows to compute the variable

$diff = wage - \alpha\left(\widehat{\widehat{education}}\right)$. The variable *diff* is regressed on the exogenous variables and the instrument

$$diff = \gamma\ \widehat{\widehat{education}} + \delta_1\ kids + \delta_2\ male + \delta_3\ married + \delta_4\ south + \delta_5\ age + \delta_6\ age^2. \quad (9)$$

A Wald test is implemented to verify the null that the instrumental variable does not have any additional overlooked impact on log of wages. The failure to reject the null $H_0: \gamma = 0$ in equation (9) validates the CH_IVQ estimates.

The Wald test is here implemented through an auxiliary regression, where the residuals from the CH_IVQ estimated equation become the dependent variable and $\widehat{\widehat{education}}$ is the explanatory variable. The term nR^2 is compared to the critical value of a χ^2 with degrees of freedom given by the number of instrumental variables.

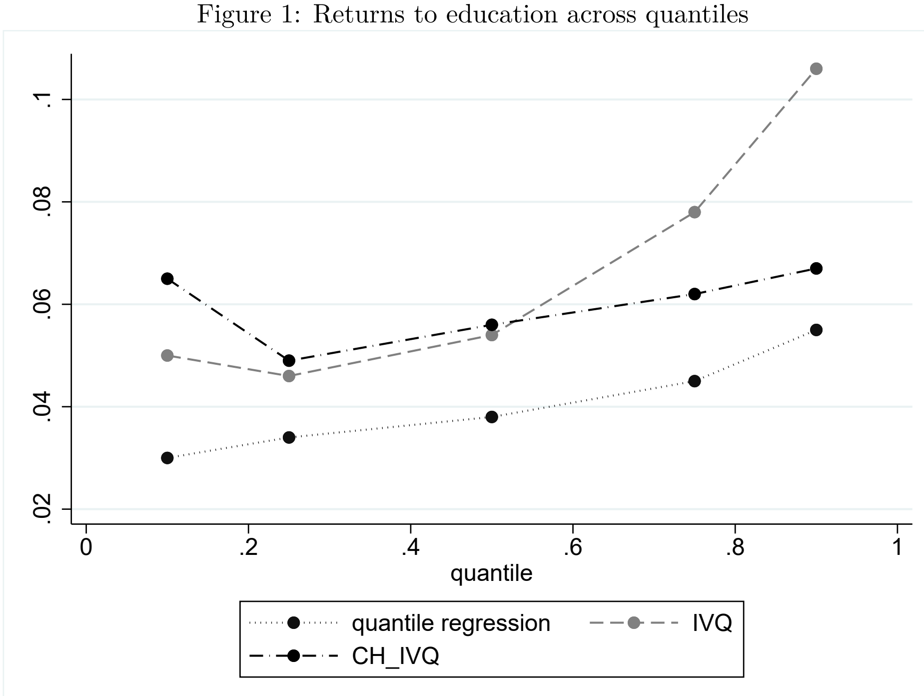
Table 4: Iterated instrumental variable results

	0.10	0.25	0.50	0.75	0.90
	wage	*wage*	*wage*	*wage*	*wage*
$\widehat{\widehat{education}}$	0.0658 (7.01)	0.0497 (6.88)	0.0561 (8.36)	0.0620 (8.39)	0.0669 (7.07)
kids	−0.1571 (−2.88)	−0.1534 (−3.66)	−0.1351 (−3.47)	−0.1246 (−2.91)	*-0.0539* (−0.98)
male	0.1229 (5.59)	0.1245 (7.35)	0.1209 (7.70)	0.1069 (6.19)	0.701 (3.17)
married	0.0953 (2.08)	*0.0375* (1.06)	*0.0417* (1.27)	*0.0520* (1.44)	0.1049 (2.27)
south	*-0.0813* (−0.25)	−0.1013 (−5.21)	−0.0855 (−4.74)	−0.0374 (1.88)	*0.0375* (1.48)
age	*-0.0058* (−0.57)	*0.0047* (0.61)	*0.0049* (0.68)	*-0.0071* (−0.89)	*0.0135* (1.31)
age^2	*0.0001* (1.41)	*0.00004* (0.49)	*0.00005* (0.64)	*0.00003* (0.37)	*-0.00002* (−0.24)
constant	0.8342 (3.54)	1.068 (5.89)	1.135 (6.74)	1.174 (6.34)	1.132 (4.77)
Wald test	1.273	0.318	0.000	0.954	5.727

The estimated z values are in parenthesis, in italics are the non-statistically significant estimates. Sample of size $n = 3182$. The Wald test considers the null $H_0: \gamma = 0$ in the auxiliary regression of residuals from equation (2) as a function of $\widehat{\widehat{education}}$. The critical value of a χ^2 with 1 degree of freedom is 3.84 at 5%, and 6.635 at 1% significance level.

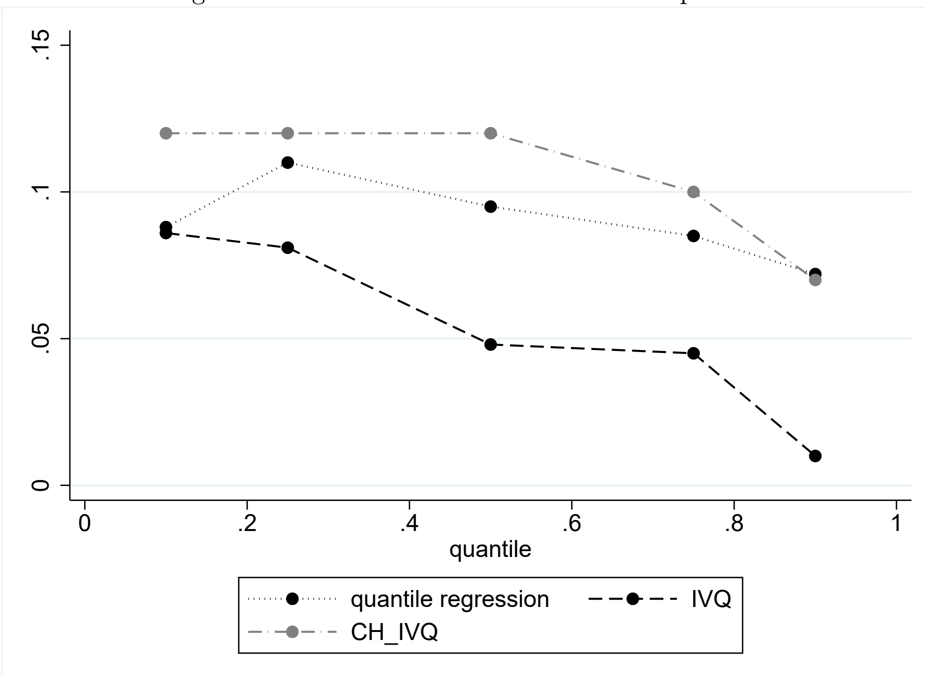
Table 4 provides the estimates of the CH_IVQ approach, and the last row of this table reports the values of the estimated Wald tests (the routine to implement in Stata this estimator, written by Kwak (2010), can be downloaded at: `www.stata.com/statalist/archive/2013-04/msg00023.html.`). This test fails to reject the null at all quantiles thus validating the CH_IVQ estimates. With respect to Table 3, returns to education are slightly larger at and below the median while become smaller at the top quantiles; the male markup is larger and significant at all quantiles and does not disappear; the southern workers penalty disappears at the 90^{th} quantile while only experience is not statistically significant. Figure 1 to 4 summarize the results of the two IVQ estimators together with the simple quantile regressions estimates of Table 2. These figures show the pattern of some coefficients, respectively returns to education, gender wage gap and regional penalty, across quantiles. Figure 2 and 3 report, respectively, the male premium, positive but decreasing across quantiles, and the woman penalty, that becomes less severe at the higher quantiles. Although the pattern across quantiles is similar for the three estimators – quantile regression, IVQ and CH_IVQ – the IVQ results are more extreme and turn out to be the largest or the smallest of all, particularly at the higher quantiles.

Figure 1: Returns to education across quantiles



Summarizing, with both IVQ and CH_IVQ estimators returns to education increase across quantiles and are generally larger than in the simple quantile regression estimates of Table 2. Therefore, the general finding that IV estimates of returns to schooling are larger than the simple OLS estimates (Brunello et al. 2001) carries on to the quantile regressions as well. With respect to the gender wage gap, IVQ signals a decreasing pattern of the estimated male coefficient, disappearing at the 90^{th} quantile. Vice versa, CH_IVQ yields larger male premium coefficients, declining only at and above the 75^{th} quantile but always statistically significant. The southern penalty diminishes across quantiles everywhere, in tables 2 to 4, and becomes a premium at the top quantiles in IVQ.

Figure 2: Male estimated coefficient across quantiles

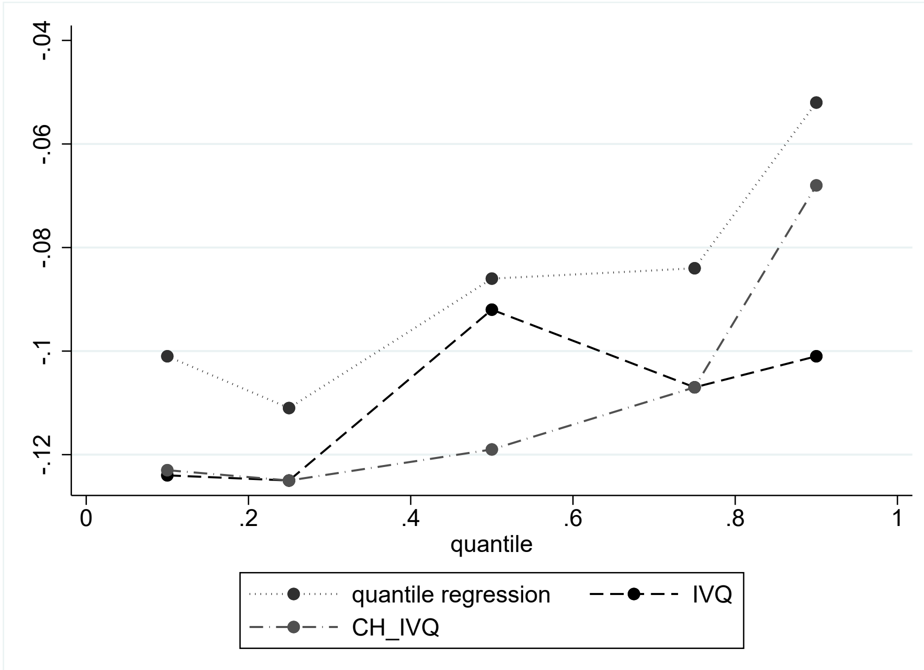


5 Decomposition results

Focusing on returns to education, this variable can be transformed into a dichotomous one by gathering in one group workers with at most a vocational degree, up to 11 school years, while the other group collects workers with higher degrees – high school and university. The corresponding dummy variable assumes respectively value $d_i = 0$ for workers having at most up to 11 years of schooling, $n_l = 1750$, and $d_i = 1$ for $n_h = 2517$ workers with higher degrees. This allows to focus on returns of high school, university and doctorate compared to vocational, junior high and elementary school premium.

Consider the vector of regression coefficients in equation (7) as estimated in the subset of workers with higher education, $\beta_1' = [\delta_1, \delta_2, ..., \delta_6]_1$, $wage_1$ their earnings and $x_1 = [kids\ male\ married\ south\ age\ age^2\ constant]_1$ the $(n, 7)$ matrix

Figure 3: Women estimated coefficient across quantiles

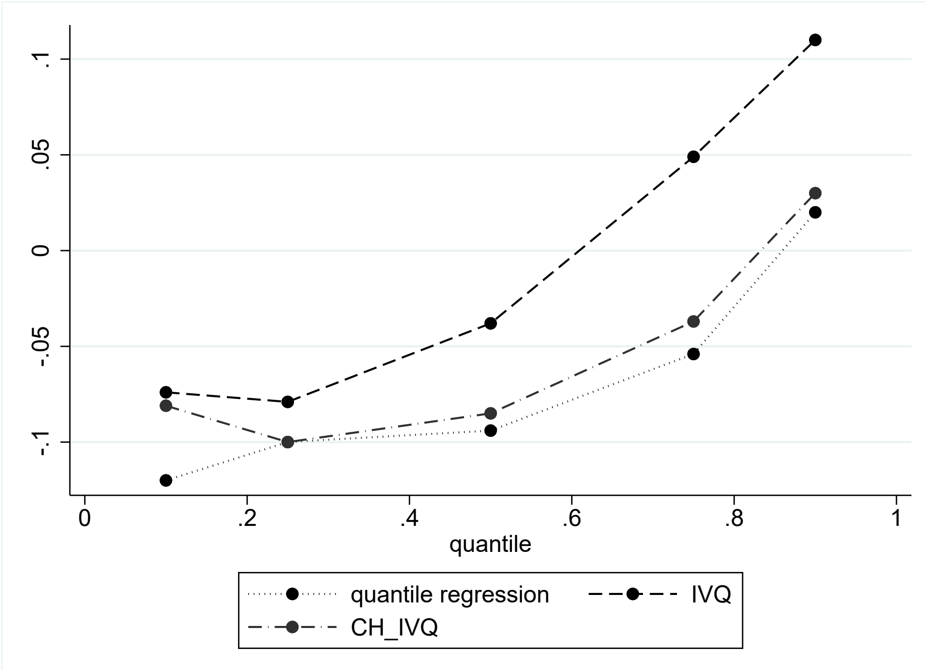


of the explanatory variables for group 1. Analogously, $\beta_0' = [\delta_1, \delta_2, ..., \delta_6]_0$, is the vector of regression coefficients estimated in group 0, $wage_0$ their wages and $x_0 = [kids\ male\ married\ south\ age\ age^2\ constant]_0$ the explanatory variables for the group with lower education. The Oaxaca (1973) and Blinder (1973) decomposition allows to write the difference between the two subsets as follows

$$\begin{aligned} wage_1 - wage_0 &= x_1\beta_1 - x_0\beta_0 = \\ &= x_1(\beta_1 - \beta_0) + (x_1 - x_0)\beta_0 = \\ &= wage_1 - wage_{0/1} + wage_{0/1} - wage_0, \end{aligned} \tag{10}$$

where in the second line $x_1\beta_0$ has been added and subtracted. The term $wage_1 = x_1\beta_1$ and $wage_0 = x_0\beta_0$ are the realizations of the dependent variable within each subset, and the term $wage_{0/1} = x_1\beta_0$ is defined by the covariates in subset 1 evaluated at the coefficients of subset 0, the unobserved counterfactual. Counterfactual distributions are the result of a change in the covariates, or a change in their relationship with the dependent variable, the regression coefficients. In terms of the wage distribution, $wage_1$ and $wage_0$ are respectively the observed wages in groups 1 and 0, while $wage_{0/1}$ is the group 1 covariate multiplied by the coefficient of group 0 and tells us what group 1 would earn if rewarded with group 0 coefficients. The first term in the third line of

Figure 4: Southern penalty estimated coefficient across quantiles



the decomposition measures the difference in wages due to changes in the regression coefficients, $wage_1 - wage_{0/1} = x_1(\beta_1 - \beta_0)$. The second term looks at the difference in wages due to differences in the covariates, $wage_{0/1} - wage_0 = (x_1 - x_0)\beta_0$, for instance differences in workers characteristics like gender, age or region, and provides a measure of the composition effect. Often the interaction term $(\beta_1 - \beta_0)(\overline{x}_1 - \overline{x}_0)$ is added to pick up any further effect. These terms are generally computed at their average values, yielding the Oaxaca-Blinder decomposition

$$\overline{wage}_1 - \overline{wage}_0 = \overline{x}_1(\beta_1 - \beta_0) + (\overline{x}_1 - \overline{x}_0)\beta_0, \tag{11}$$

where $\overline{wage}$ and $\overline{x}$ are the sample averages and the vectors of parameters β_1, β_0 are replaced by their least squares estimates. The result is an average measure of the wage differences between the two subsets. Table 5 gathers these estimates and shows a statistically significant difference between the two overall averages, $\overline{wage}_0 - \overline{wage}_1 = -0.1776$. This premium is linked to a significant average coefficients effect of $\overline{x}_1(\beta_0 - \beta_1) = -0.1983$, which is slightly offset by a small covariates effect. However, the terms in a decomposition can take different values according to the selected quantile of the log of wages distribution, center, lower and upper tails. Therefore, the decomposition can be estimated not only on average but also in the tails, by means of the quantile regression-based decomposition approach discussed in

Table 5: Oaxaca-Blinder average decomposition, impact of higher versus lower degrees

	average	std. err
lower education	2.109	0.0095
higher education	2.286	0.0088
$\overline{wage}_0 - \overline{wage}_1$	−0.1776	0.0130
due to covariates	0.0447	0.0060
due to coefficients	−0.1983	0.0131
interaction term	−0.0241	0.0064

Machado and Mata (2005). In a quantile regression decomposition it is possible to verify whether any discrepancy is statistically significant at each quantile and whether such a discrepancy is stable or changes across quantiles.
The Machado and Mata (2005) approach computes $m = 100$ different quantile regressions within each group defined as

$$wage_i = \delta_1\ kids_i + \delta_2\ male_i + \delta_3\ married_i + \delta_4\ south_i + \delta_5\ age_i + \delta_6\ age_i^2 \qquad i = 0, 1 \tag{12}$$

the estimated coefficients and the explanatory variables are separately bootstrapped within each group, yielding $\hat{\hat{\beta}}_0(\theta)$, $\hat{\hat{\beta}}_1(\theta)$, $\widetilde{X}_0$ and $\widetilde{X}_1$. The quantile regression decomposition can be computed considering the terms $\widehat{\widehat{wage}}_0 = \widetilde{X}_0\hat{\hat{\beta}}_0(\theta)$, $\widehat{\widehat{wage}}_1 = \widetilde{X}_1\hat{\hat{\beta}}_1(\theta)$, $\widehat{\widehat{wage}}_{0/1} = \widetilde{X}_1\hat{\hat{\beta}}_0(\theta)$ to replace the average values in the Oaxaca-Blinder decomposition

$$\widehat{\widehat{wage}}_1 - \widehat{\widehat{wage}}_0 = \widetilde{X}_1\left(\hat{\hat{\beta}}_1(\theta) - \hat{\hat{\beta}}_0(\theta)\right) + (\widetilde{X}_1 - \widetilde{X}_0)\hat{\hat{\beta}}_0(\theta). \tag{13}$$

This allows to implement the decomposition at various quantiles θ and yields results on the behavior in the tails that cannot be detected by the standard Oaxaca-Blinder average decomposition approach. The routine to implement this estimator in Stata, written by S. Souabni (2013), can be downloaded at `http://fmwww.bc.edu/repec/bocode/m/mmsel.ado`. Table 6 reports the results of the decomposition at the deciles. They show a positive impact of higher education that is statistically significant throughout, increases across quantiles and is due to a pure coefficient effect. The latter represents the sheer premium to higher education, purged by any possible difference in the covariates. Indeed, the covariates effect is very small in the table, although statistically significant. Figure 5 depicts the pattern of the total difference between groups, $\widehat{\widehat{wage}}_1 - \widehat{\widehat{wage}}_0$, as decomposed by differences due to the covariates and to the coefficients. The covariates effect is small and almost constant while the coefficients effect is wider and steadily increasing across deciles.

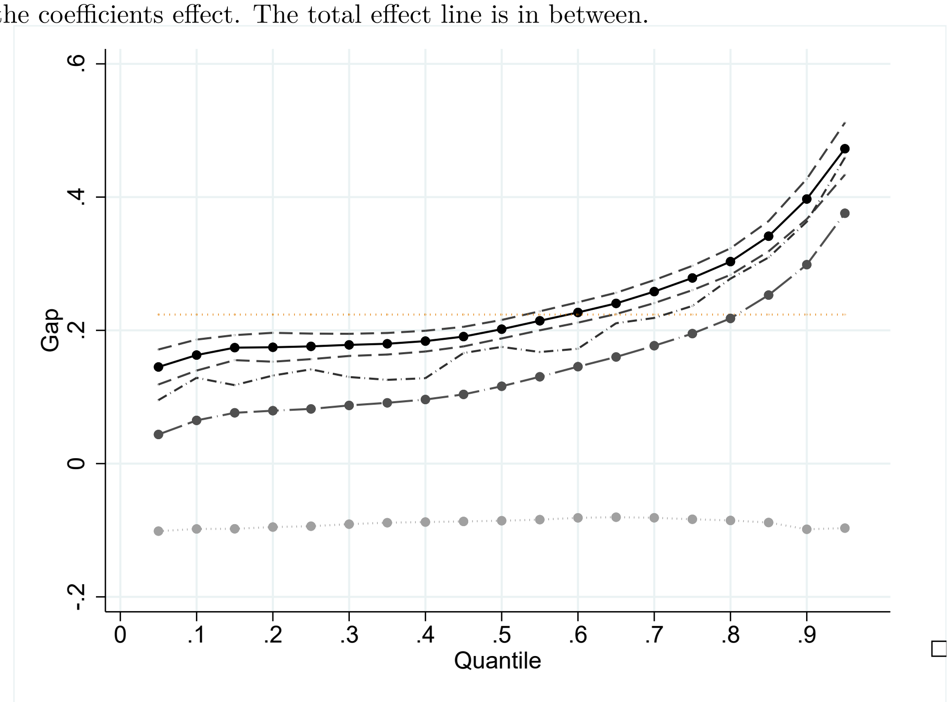
These results can be compared with a different decomposition approach which allows to consider the endogeneity of education. Using the quantile treatment effect estimator (QTE) discussed by Frölich and Melly (2008, 2010), the coefficients effect

Table 6: Decomposition of the impact of higher education across deciles, $\widehat{wage}_1 - \widehat{wage}_0$

quantile	total effect	std. err.	effect of characteristics	std. err.	effect of coefficients	std. err.
0.1	0.0649	0.011	−0.0980	0.012	0.1629	0.011
0.2	0.0794	0.010	−0.0952	0.010	0.1746	0.011
0.3	0.0873	0.008	−0.0908	0.009	0.1782	0.008
0.4	0.0961	0.008	−0.0876	0.008	0.1838	0.007
0.5	0.1161	0.007	−0.0856	0.008	0.2017	0.007
0.6	0.1455	0.008	−0.0813	0.008	0.2269	0.007
0.7	0.1769	0.008	−0.0812	0.009	0.2582	0.008
0.8	0.2179	0.009	−0.0852	0.011	0.3032	0.010
0.9	0.2986	0.015	−0.0985	0.018	0.3971	0.015

Number of workers in group 1 with higher education, $n_h = 1899$; number of workers in group 0 having at most a vocational degree, $n_l = 1283$.

Figure 5: Covariates, coefficients and total effects across quantiles in the Machado and Mata decomposition. The bottom line shows the covariates effect, the top one the coefficients effect. The total effect line is in between.



□

can be computed at many quantiles. This estimator considers a weighted linear quantile regression of the dependent variable on a constant term and on the dummy variable assuming unit value for higher education d_i

$$\begin{aligned} \sum \rho(e_i)w_i &= \sum [\rho(wage_i - a_0 - a_1 d_i)w_i] = \min, \\ w_i &= (2d_i - 1)\left[\frac{Z_i - p_i(x)}{p_i(x)(1 - p_i(x))}\right], \end{aligned} \tag{14}$$

where $\sum \rho(e_i)$ is the usual quantile regression objective function. The weights w_i are function of the dummy d_i, the instrumental variable Z_i, $\widehat{\widehat{education}}$ in our case, and the propensity score $p_i(x)$. The latter is the probability of each worker to belong to the group of higher/lower education, it is function of all the explanatory variables of the model and is estimated by a logit model. The weighting system allows to even out the covariates of the two groups, covariates of workers with higher versus lower education, net of endogeneity. Any difference between the two groups is exclusively due to the coefficients, and this estimator has the advantage of measuring their individual impact. Table 7 reports the results of the comparison $\widehat{\widehat{\widehat{wage}}}_1 - \widehat{\widehat{\widehat{wage}}}_0$ as computed by the IPW approach in (14). The routine to implement in Stata this estimator can be found at: st0203 from `http://www.stata-journal.com/software/sj10-3.` Higher education has a positive and significant coefficient effect at all quantiles, that increases across quantiles.

Table 7: Coefficients effect at the quantiles

	0.10	0.25	0.50	0.75	0.90
	w_i $wage_i$	w_i $wage_i$	w_i $wage_i$	w_i $wage_i$	w_i $wage_i$
w_i *higher educ.*	0.1073	0.1500	0.1786	0.1927	0.2971
w_i *kids*	−0.1148	−0.1380	−0.0652	−0.0403	−0.0383
w_i *male*	0.0741	0.0987	0.0697	0.0469	0.0722
w_i *married*	0.0788	0.0505	0.0679	0.1059	0.1950
w_i *south*	−0.1341	−0.1169	−0.1187	−0.0109	0.0224
w_i *age*	−0.0057	−0.0035	0.0058	−0.0012	−0.0227
w_i age^2	0.0001	0.0001	0.00007	0.0002	0.0004
w_i *constant*	1.628	1.7575	1.674	1.913	2.383

The weights w_i are function of the propensity scores $p_i(x)$ and of the instrumental variable $z_i = \widehat{\widehat{education}}$.

Returns to higher education are quite large at all quantiles in both Table 6 and 7, much larger than in Table 4 or 5. The latter measure the overall premium to education whereas tables 6 and 7 compute the premium of higher with respect to lower education. The wide discrepancy in the comparison of the results in Table 4, 5 with those in Table 6, 7 can be related to non-linearities in returns, that are much

larger than the overall premium for the higher school degrees at all quantiles. The last two tables show an increasing impact of higher education across quantiles. However, returns in Table 7 are smaller and present a smoother pattern than in Table 6. Thus, accounting for endogeneity causes a smoother coefficients effect of higher education across deciles.
This second approach has the advantage of pointing out the impact of each variable within the decomposition. This tells us something about the individual pattern of gender wage gap and regional penalty, and Table 7 shows a decreasing but non-vanishing male premium across quantile, confirming the CH_IVQ results, and a reduction in the southern workers penalty that becomes a premium at the top quantile.
In sum returns to education increase with the completed degree and grow across quantiles. The decomposition shows that such increase is a sheer premium to higher education which does not depend upon any other difference in the characteristics of the two groups. It even offsets the southern penalty at the top quantile. However, OECD (2019) statistics show that in Italy returns to higher education are below the OECD average.

6 Conclusions

The analysis of a wage equation using Italian data has uncovered some characteristics of the model which deeply affect the results:

1. heterogeneity, that causes changing coefficients across quantiles;

2. endogeneity, that once purged signals greater returns to education at all quantiles, a decreasing pattern of gender wage gap and a milder regional penalty across quantiles;

3. non-linearity, that reveals greater returns to higher with respect to lower degrees, increasingly so along the wage distribution.

The data set has been analyzed at various quantiles with several quantile regressions estimators. The changes in the estimated coefficients across quantiles show returns to education increasing with quantiles, with a 3% raise at the lower wages that grows above 5% at the higher earnings. This implies that education increases earnings later more than earlier in a lifetime, since it takes time to progress in a career. The presence of differing estimates across quantiles shows the limitations of OLS, which computes only an average rate of return along the entire wage distribution.
The endogeneity is related to the correlation between error term and education, and it is mostly driven by the non-observable individual talent/ability. The latter is unobservable, but it has an impact on both returns to education and gender wage gap. Ignoring endogeneity yields biased and inconsistent estimates, and the bias can be negative or positive depending on the individual discount rate on future

earnings. Two different quantile regression instrumental variable approaches are implemented to solve endogeneity. The instrument is provided by grand-parents education, since parents education is correlated with offspring education and is uncorrelated with descendants' earnings. In both the quantile regression instrumental variables estimators implemented, returns to education are larger than in the simple quantile regressions approach. This is a known result at the mean, but in our model it holds also at the quantiles: returns to education estimated by instrumental variables improve upon the simple regression results at all quantiles. Thus, endogeneity causes the under-estimation of returns to education in both the least squares and the quantile regression approaches. The bias in the simple regressions estimates shows a disproportionate relevance in these methods of workers with high discount rates on future earnings. The decreasing pattern of the male coefficient across quantiles, measuring the gender wage gap, is confirmed by both quantile regressions instrumental variable estimators.
Finally, the non-linearity can be analyzed by splitting the various school degrees in two groups, lower and higher education. The comparison of the two different sub-samples shows that returns increase as a function of the higher completed degree. Indeed, by comparing the earning profiles of workers with lower versus higher education, returns to education to higher degrees are larger and the gap rises across quantiles. The results of a quantile regression-based decomposition show that the difference in returns between higher and lower degrees is a sheer premium to higher education, independent of any other difference in the characteristics of the two groups. The increment in returns to higher versus lower degrees is much wider than the premium to education computed by the previous quantile regressions approaches, thus signaling non-linearities. The inclusion of the instrumental variable in the quantile decomposition approach yields a smoother pattern of returns to higher education across quantiles, but still endorses the completed degree as a source of heterogeneity. In terms of policy implications, the results show the need of actions leading to a wider inclusion of women in the regular job market in terms of both sheer number of workers and their remuneration. The southern workers penalty is mostly due to a lagging southern economy and calls for relevant and effective intervention, although for highly educated workers at the top wages the southern economy provides better opportunities, wider returns. Finally, returns to education with respect to higher degrees is still below the OECD average value and in need of improvements. This can partly explain the lower number of workers with university degree in Italy.

References

[1] Amemiya T., (1982), Two stage least absolute deviations estimators, *Econometrica* 50, 689–711.

[2] Bettio F., Villa P., (1999), To what extent does it pay to be better educated?

Education and the work market for women in Italy, *South European Society and Policy* 4, 150–170.

[3] Blinder A., (1973), Wage discrimination: reduced form and structural estimates, *Journal of Human Resources* 8, 436–455.

[4] Blundell R., Dearden L., Sianesi B., (2001), *Estimating the returns to education: models, methods and results*, Center for Economics of Education, London.

[5] Blundell R., Dearden L., Sianesi B., (2005), Evaluating the effect of education on earnings: models, methods and results from National Child Development Survey, *Journal of the Royal Statistical Society A* 168, 473–512.

[6] Brunello G., Comi S., Lucifora C., (2001), The Returns to Education in Italy: A New Look at the Evidence, [in:] *The Returns to Education in Europe*, [ed.:] Harmon C., Walker I. and Westergard-Nielsen N., Cheltenham (UK) Edward Elgar.

[7] Brunello G., Miniaci R., (1999), The Economics Returns to Schooling for Italian Men. An Evaluation Based on Instrumental Variables, *Labour Economics* 6, 509–519.

[8] Buchinsky M., (1998), The dynamics of changes in the female wage distribution in the USA: a quantile regression approach, *Journal of Applied Econometrics* 13, 1–30.

[9] Cannari L., D'Alessio G., (1998), Il rendimento dell'istruzione: alcuni problemi di stima,*Temi di Discussione* 253, Banca d'Italia.

[10] Card D., (2001), Estimating the return to schooling: progress on some persistent econometric Problems, *Econometrica* 69, 1127–60.

[11] Card D., DiNardo J., (2002), Skill biased technological change and rising wage inequality: some problems and puzzles, *Journal of Labor Economics* 20, 733–783.

[12] Chernozhukov V., Hansen C., (2005), An IV model of quantile treatment effect, *Econometrica* 73, 245–261.

[13] Chernozhukov V., Hansen C., (2006), Instrumental quantile regression inference for structural and treatment effect models, *Journal of Econometrics* 132, 491–525.

[14] Chernozhukov V., Hansen C., Jansson M., (2007), Inference approaches for instrumental variable quantile regression, *Economic Letters* 95, 272–277.

[15] Chernozhukov V., Hansen C., (2008), Instrumental quantile regression: A robust inference approach, *Journal of Econometrics* 142, 379–398.

[16] Chernozhukov V., Fernandez-Val I., Melly B., (2013), Inference on counterfactual distributions, *Econometrica* 81, 2205–2268.

[17] Ciccone A., Cignano F., Cipollone P., (2006), The private and social return to schooling in Italy, *Temi di Discussione* 569, Banca d'Italia.

[18] Colussi A., (1997), Una analisi cross section del tasso di rendimento dell'istruzione in Italia, *Politica Economica* 13.

[19] Dickinson M., Harmon C., (2011), Economic returns to education: what we know, what we don't know and where we are going-some brief pointer, *Economic of Education Review* 30, 1118–1122.

[20] European Comission (2019), The gender pay gap situation in the EU. https://ec.europa.eu/info/policies/justice-and-fundamental-rights/gender-equality/equal-pay/gender-pay-gap-situation-eu_en.

[21] Flabbi L., (1999), Returns to Schooling in Italy: OLS, IV and Gender Differences, *Universitá Bocconi Working Paper*.

[22] Frölich M., Melly B., (2008), *Unconditional quantile treatment effect under endogeneity*, Insitute for the Study of Labor (IZA) 3288 at http://ideas.repec.org/p/iza/izadps/dp3288.html.

[23] Frölich M., Melly B., (2010), Estimation of quantile treatment effects with Stata, *Stata Journal* 10, 423–457.

[24] Furno M., (2014), Returns to education and gender gap, *International Review of Applied Economics* 28, 628–649.

[25] Giustinelli P., (2004), Quantile Regression Evidence on Italian Education Returns, *Rivista di Politica Economica* XI-XII, 49–100.

[26] Harmon C., Oosterbeek H., Walker I., (2003), The returns to education: microeconomics, *Journal of Economic Surveys* 17, 115–155.

[27] Koenker R., (2005), *Quantile regression*, Cambridge University Press.

[28] Kwak D., (2010), Instrumental variable quantile regression method for endogenous treatment effect, *unpublished manuscript*.

[29] Meghir C., Palme M., (2000), Estimating the Effect of Schooling on Earnings Using a Social Experiment, *IFS* 99/12.

[30] Melly B., (2006), *Estimation of counterfactual distributions using quantile regression* w.p. University of St. Gallen.

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

[32] Machado J., Mata J., (2005), Counterfactual decomposition of changes in wage distributions using quantile regression, *Journal of Applied Econometrics* 20, 445–465.

[33] Oaxaca R., (1973), Male-female wage differentials in urban labor markets, *International Economic Review* 14, 693–709.

[34] OECD, (2019), Education at a Glance 2019 OECD Indicators. `https://www.oecd-ilibrary.org/docserver.`

[35] OECD, (2015), PISA 2015 country note for Italy. `https://www.oecd.org/pisa/PISA-2015-Italy.pdf.`

[36] Souabni S., (2013), MMSEL: Stata module to simulate (counterfactual) distributions from quantile regressions, Statistical Software Components S457441, Boston College Department of Economics.

[37] Sulis G., (2007), Gender wage differentials in Italy: a structural estimation approach, *Crenos working paper* 15.

[38] Zizza R., (2013), The gender wage gap in Italy, *Questioni di Economia e Finanza* 172, Banca d'Italia.