Economics Bulletin

Volume 31, Issue 3

Education and wage inequality in Europe

Marco Biagetti
Ministry of Economic Development, Department of Economic and Social Cohesion

Sergio Scicchitano
"La Sapienza" University and Italian Ministry of Economic Development

Abstract

In this paper we apply a quantile regression (QR) approach to the EU-SILC data set, in order to explore the connection between education and wage inequality in eight European countries. Our results corroborate the positive relation between wage increase and education and it holds true across the whole distribution. This effect is generally stronger at the highest quantiles of the distribution than at the lowest, implying that schooling increases wage dispersion. This evidence is found to be rather robust like the applied tests of linear hypothesis show.

We would like to thank the the Associate Editor Dan Anderberg, Marco Marini and an anonimous referee for valuable suggestions and comments. The views expressed in this article are those of the authors and, in particular, do not necessarily reflect those of the Ministry of Economic Development. The usual disclaimer applies.

Citation: Marco Biagetti and Sergio Scicchitano, (2011) ''Education and wage inequality in Europe'', *Economics Bulletin*, Vol. 31 No. 3 pp. 2620-2628.

Contact: Marco Biagetti - marco.biagetti@tesoro.it, Sergio Scicchitano - sergio.scicchitano@uniroma1.it.

Submitted: July 11, 2011. **Published:** September 13, 2011.

I. Introduction

In this article we investigate the effect of education on within-groups inequality in eight European countries, using the last wave of the European Union Statistics on Income and Living Conditions inquiry (EU-SILC). It is the new European homogenized panel survey that, to our knowledge, has never been used in such a comparative analysis so far.

Over the last two decades the increase in wage inequality between high-skilled and low-skilled workers became a relevant topic, particularly for the European countries. This issue started stimulating several possible explanations from the economic literature. The Skill-Biased Technological Change (SBTC) paradigm seems to be the chief justification, but also other causes, such as the role and characteristics of labor market Institutions (for instance, corporatism, centralization of the wage bargaining) and the international trade have been called[1]. In particular, there is a wide consensus that schooling has a relevant impact on wage inequality across countries. Returns to education tend to be increasing over the wage distribution and this evidence is interpreted as a direct effect of schooling on within-groups inequality.

As a matter of fact, some empirical analysis explored the gain effects of education over the entire wage distribution. In particular, since the seminal paper by Koenker and Basset (1978), quantile regression (QR) was frequently adopted. This method has been used in many articles to study single countries, but not as much to make scrutiny on comparable cross-countries differences. This in-depth examination has been mainly constrained by the availability of diverging data sets, as argued by Budria and Pereira (2005). In particular, on the one hand, these two authors evaluated the impact of education on wage inequality in nine European countries, by distinguishing for educational qualification. On the other hand, Martins and Pereira (2004, MP henceforth), showed that in sixteen European countries higher education is associated with higher wage dispersion. Both these papers were completed under a framework of research projects where each country team analyzed their own country different data set. In this article we apply a similar approach to eight European countries, using EU-SILC, widely considered an attractive source of information, as it adopts the same "community" questionnaire, thus obviously making comparisons across nations much easier. It has succeeded the European Community Household Panel (ECHP) since 2005, including the new European Member States. In particular the use of EU-SILC allows us to also include Poland, thus enabling – for the first time to our knowledge – the implementation of a comparison survey on returns to education between a new European Member State and many old members.

To address these issues we apply a QR semi-parametric approach which seems more interesting, as well as more suitable, for it allows us to get

[1] See among others Acemoglu (2002) for what concerns the SBTC, Di Nardo et al. (1996) for the impact of labor market Insitutions and Wood (1995) for the role of the international trade.

a more precise picture of the dynamics of the dependent variable at different points of the distribution, rather than at the conditional mean. Through it, the article points out a high cross-country heterogeneity in returns to education at different points of the wage distribution, which OLS modeling of conditional average of a dependent variable completely fails to account for.

This empirical paper is organized as follows. The next section describes the data. Section 3 illustrates our econometric specification. Section 4 reports the results, both in terms of OLS and QR. In Section 5 the robustness check is presented, while in the final section main conclusions are drawn.

II. Data selection

The analysis has been carried out on eight countries, the only ones having available data for our interest variables in EU-SILC. More specifically, for Austria, Spain, Ireland, Italy, Poland and Portugal the latest 2007 EU-SILC wave has been used. For Belgium and Greece the latest disposable wave for those variables was that of 2005.

We have focused our analysis on the personal file of EU-SILC. We chose to concentrate the survey on males aged between 25 and 65 working full-time: women were disregarded on account of potential selectivity biases. Younger males were dropped because they are still in the "almost exclusively" educational period of their life, i.e. they are very likely enrolled in a secondary or tertiary course and at the same time do not perform any work activity. People who had missing or NA data on the educational variable were also dropped. Our dependent variable is the hourly (logarithmic) gross wage, available for 24,118 full-time working males aged between 25 and 65 for the six countries of the 2007 wave and 3,621 Belgian and Greek men from the 2005 wave. Thus our analysis focuses on 27,739 individuals overall.

EU-SILC does have data for the highest educational attainment from which we built up our first independent variable (schooling years) following the usual framework, i.e. by taking the highest ISCED level of education attained by a male worker, and assigning the legal minimum number of years typically required to achieve it for each level[2]. Our second and third regressors are respectively the number of years spent in paid work and its squared: the former is regarded as being a proxy for individual experience while the latter takes account of possible non linearities. Summary statistics for these variables are shown in Table 1.

[2] Those who reached only an ISCED 1 grade have been given five years of schooling; eight years have been assigned to those with an ISCED 2 grade; 13 years to those with an ISCED 3; 14 to people who attained an ISCED 4 grade and 18 years to those who reached an ISCED 5.

Tab. 1. Summary statistics

Country	Var.	Obs	Mean	S. D.
Austria	Wage	2774	2.72	0.44
	Schooling	3280	13.90	2.72
	Experience	3259	23.54	10.14
Belgium	Wage	1754	2.78	0.37
	Schooling	2234	13.91	4.10
	Experience	2232	20.23	10.62
Greece	Wage	1867	2.03	0.41
	Schooling	3038	11.25	4.80
	Experience	3038	19.69	11.26
Ireland	Wage	1574	3.06	0.54
	Schooling	2172	12.49	4.79
	Experience	2100	25.76	11.87
Italy	Wage	7324	2.45	0.38
	Schooling	10512	11.58	3.83
	Experience	10512	19.59	10.48
Poland	Wage	5273	1.13	0.52
	Schooling	6862	13.12	3.21
	Experience	6819	19.70	10.93
Portugal	Wage	1735	1.68	0.55
	Schooling	2204	8.08	4.37
	Experience	2197	24.39	12.33
Spain	Wage	5438	2.37	0.45
	Schooling	7181	11.56	5.00
	Experience	7076	22.63	11.75

Note. Data for Austria, Spain, Ireland, Italy, Poland and Portugal are from cross sectional UDB SILC 2007 – version 1 of March 2009; Belgium and Greece from EU-SILC 2005.

III. Econometric specification

The first equation has the following simple form:

$$\ln w_i = \alpha_i + \beta_i \mathrm{S} + \gamma_i E + \delta_i E^2 + \varepsilon_i \tag{1}$$

where S represents the years of education (schooling) and E is the work experience.

Equation (1) is solved through a classic OLS method, based on the mean of the conditional distribution of the dependent variable. As it is well known, it implicitly assumes that the impact of the regressors along that conditional distribution are irrelevant. This fact is referred to as a pure location shift. In other words, the x's are unable to cause a scale effect or any other consequence on the distributional shape. But as covariates may influence the conditional distribution of the response in many other ways, an estimate of the whole distribution of conditional quantiles of the dependent variable seems more appropriate to study the influence of the regressors on its shape. We do this performing a QR approach, which has the following functional form (Koenker & Basset, 1978): min

$$\sum_{i:\ln w_i \geq x_i\beta} \theta \left| \ln w_i - \left(\alpha_\theta + \beta_\theta S_i + \gamma_\theta E_i + \delta_\theta E_i^2 \right) \right| + \\ + \sum_{i:\ln w_i < x_i\beta} (1-\theta) \left| \ln w_i - \left(\alpha_\theta + \beta_\theta S_i + \gamma_\theta E_i + \delta_\theta E_i^2 \right) \right| \tag{2}$$

Equation (2) is normally written as:

$$\min_{B \in R^k} \sum_i \rho_\theta \left(\ln w_i - \alpha_\theta - \beta_\theta S_i - \gamma_\theta E_i - \delta_\theta E_i^2 \right) \tag{3}$$

where $\varrho_\theta(z) = \theta z$ if $z \geq 0$ or $\varrho_\theta(z) = (\theta - 1)z$ if z<0 and B=α, β, γ, δ: of course, k=4 in this case

This problem is solved using linear programming methods. To estimate standard errors and confidence intervals, a bootstrap technique is used to replicate the above procedure. In this study 300 replications are carried out.

IV. Results

In Table 2 we show OLS returns as well as conditional returns at seven representative quantiles. Both OLS and QR estimated coefficients are positive and significant at the 1% level in every country. Differences between percentiles of the wage distribution computed for six different extremes taken by twos (θ95-θ5, θ90-θ10 and θ75-θ25) are also reported. In terms of OLS returns of education, Portugal shows the highest coefficient (8%). At the bottom of the wage distribution Greece displays the lowest value (3.5%), while Italy shows a slightly higher coefficient (4.3%). Moreover, results for Poland point out that the estimate of the rate of return in that country is among the highest in Europe.

The differences between percentiles computed at the six extremes considered decrease with the distance between percentiles in every country: in other words, θ95-θ5 (θ75-θ25) is always higher (lower) than θ90-θ10. Our estimates show that the country having the greatest OLS coefficient and the largest spreads between percentiles in all of the three comparisons, is still

Portugal. The same situation occurs for Greece at the bottom of the distribution, which is also the only country with no difference between the 75th and 25th quantile. Poland shows both relatively high returns on education and high inter-quantile differences. Austria has a particular evidence: it displays a high OLS returns on education but quite low differences between the percentiles of the wage distribution.

These results signalize that OLS technique really misleads relevant information about cross-counties differences in the impact of education on within group inequality at different points of the wage distribution. There is a clear evidence that wages increase with education and this is true across the whole distribution. Furthermore, this effect is generally more important at the highest quantiles than at the lowest, implying that higher schooling increases wage dispersion. Also Greece, which was found the only exception by MP, follows the same pattern[3]. Despite this common pattern across countries, different paths across countries from the bottom to the top of the distribution arise. Focusing only on Poland, a changing track over the wage distribution is clear: returns to education of Polish adult male workers trace a curve which is concave in the lower half and then it becomes convex from that point on, with a couple of jumps (around 1%) from the 5th to the 10th quantile and from the 90th to the 95th.

Table 2. Conditional returns to schooling – OLS, QR and inter-quantile differences.

Quantile	Austria	Belgium	Spain	Greece	Ireland	Italy	Poland	Portugal
OLS	0.067	0.049	0.045	0.035	0.055	0.043	0.063	0.079
	(22.34)	(26.33)	(40.93)	(19.54)	(20.44)	(40.34)	(29.19)	(31.26)
θ=.05	0.055	0.035	0.031	0.023	0.036	0.030	0.033	0.036
	(9.30)	(8.49)	(11.50)	(6.22)	(2.75)	(10.20)	(3.63)	(5.90)
θ=.10	0.050)	0.033	0.030	0.028	0.042	0.027	0.042	0.044
	(11.74	(10.12)	(15.16)	(7.87)	(7.20)	(14.35)	(9.27)	(11.06)
θ=.25	0.052)	0.037	0.037	0.034	0.049	0.030	0.054	0.069
	(20.44)	(22.73)	(18.62)	(15.82)	(15.32)	(22.51)	(24.52)	(25.75)
θ=.50	0.063	0.049	0.044	0.034	0.054	0.035	0.065	0.082
	(18.96)	(21.21)	(30.35)	(13.08)	(19.25)	(30.82)	(24.95)	(64.62)
θ=.75	0.071	0.054	0.051	0.034	0.058	0.047	0.064	0.092
	(16.44)	(30.3)	(25.35)	(17.00)	(13.41)	(25.36)	(26.99)	(37.79)
θ=.90	0.078	0.057	0.055	0.041	0.066	0.059	0.068	0.088
	(13.26)	(19.97)	(31.01)	(9.60)	(20.53)	(29.08)	(16.57)	(29.32)
θ=.95	0.085	0.061	0.061	0.044	0.075	0.067	0.076	0.088
	(17.8)	(11.52)	(39.97)	(7.95)	(12.54)	24.86)	(19.31)	(12.13)
θ95-θ5	0.030**	0.026***	0.030***	0.021***	0.039***	0.037***	0.043***	0.052***
θ90-θ10	0.028***	0.024***	0.025***	0.013***	0.024***	0.032***	0.026***	0.044***
θ75-θ25	0.019***	0.017***	0.014***	0.000	0.009**	0.017***	0.010**	0.023***
Obs.	2,756	1,744	5,350	1,867	1,522	7,324	5,236	1,702

Note. Data for Austria, Spain, Ireland, Italy, Poland and Portugal are from cross sectional UDB SILC 2007 – version 1 of March 2009; Belgium and Greece from EU-SILC 2005. All coefficients of the quantile regression significant at $p<0.001$, with *t-statistics* in parentheses. Significance of coefficients of the inter-quantile differences : * $p<0.10$. ** $p<0.05$. *** $p<0.01$.

For the sake of precision, below here a comparison with the MP's estimates is presented:

[3] It can be further noted that in that study data for Austria, Greece and Italy were based on net wages, "which troubles a full comparison with the remaining countries" (MP, p. 365).

Table 3: our estimates vs. MP's

	Austria		Spain		Greece		Ireland		Italy		Portugal	
	Our estim.	MP	Our estim	MP	Our estim	MP	Our estim.	MP	Our estim.	MP	Our estim.	MP
OLS	0.067	0.093	0.045	0.082	0.035	0.063	0.055	0.086	0.043	0.062	0.079	0.119
θ=.10	0.050	0.070	0.030	0.065	0.028	0.073	0.042	0.075	0.027	0.065	0.044	0.065
θ=.50	0.063	0.091	0.044	0.087	0.034	0.056	0.054	0.099	0.035	0.056	0.082	0.122
θ=.90	0.078	0.120	0.055	0.087	0.041	0.055	0.066	0.099	0.059	0.068	0.088	0.145

For the countries (and quantiles) whose returns had been evaluated by MP, it is absolutely evident that our estimates are sensibly lower. This may be due to the different dataset used. Nonetheless, futher explanations are given below in this paper.

V. Robustness check

Further, in the same Table 2 we test whether gaps between quantile coefficients estimated in our QR are statistically significant. The test has been carried out with respect to the three spreads considered in the paper (θ95-θ5=0, θ90-θ10=0 and θ75-θ25=0) and to all quantiles. More specifically, p-values are obtained through a bootstrapped variance-covariance matrix that includes between quantiles blocks. The results indicate that the first linear hypothesis (θ95-θ5=0) is found to be significant at all levels of confidence for almost each of the eight countries. Only Austria displays a weaker difference, as the associated p-value is not significant at the 1% confidence level. As to the second linear hypothesis (θ90-θ10=0), overall significance is found. As expected, significance decreases when the third linear hypothesis (θ75-θ25=0) is analyzed: in particular it is discovered to be not significant at the 1% confidence level for Ireland and Poland: further, it is statistically not significant for any confidence level in Greece. Finally, the joint equality of coefficients at all quantiles is rejected as well at the 1% confidence level.

VI. Conclusions

In this paper we have applied a QR technique to the last available data collected from the EU-SILC survey, in order to explore the connection between education and wage inequality in eight European countries. Our comparative study gives a contribution to the "little comparable evidence for Europe" (Budria and Pereira 2005, p.1). We found that, although with a relevant cross-countries heterogeneity, wages increase with education and this holds true across the whole distribution. Furthermore, this effect is generally more important at the highest quantiles of the distribution than at the lowest, implying that schooling increases wage dispersion. This evidence is found to be rather robust as showed through several specification tests.

The effects computed at different quantiles (0.1, 0.5 and 0.9) are shown to be lower than those calculated by MP. This may be due to the different dataset used, Nonetheless, alternative interpretations can be thought of: indeed, as MP's estimates were computed more than a decade ago, this result suggests that returns to education have decreased all the way throughout Europe due to a higher number of strongly educated people competing for a (high-skilled) job on one hand, and the mismatch between educational attainment and quality of work, very often resulting in the so-called overeducation phenomenon, on the other. Anyway, any insight on whether the returns have been really decreasing in the last 10-13 years is left to further research.

We have thus corroborated the idea that a semi-parametric QR approach is more interesting, as well as more appropriate, because it measures the wage effect of education at different quantiles, thus describing relevant cross-countries changes or bounces not only in the location, but also in the shape of the distribution.

Our results may be driven by several, arguably not alternative, explanations. In particular the increasing in the education premia over the wage distribution, which in this paper is confirmed to be a common characteristic across European labor markets, is fully compatible with the SBTC theory.

References

Acemoglu, D. (2002) Technical Change, Inequality, and the Labor Market, *Journal of Economic Literature*, **40**, 7-72.

Budria, S and Pereira, P. S. (2005) Educational Qualifications and Wage Inequality: Evidence for Europe, IZA Discussion Papers, **1763**.

Di Nardo, J., Fortin, N. M. and Lemieux, T. (1996) Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach, *Econometrica*, **65**, 1001-46.

Koenker, R. and Bassett G. (1978) Regression quantiles. *Econometrica*, **46**, 33–50.

Martins, P. T. and Pereira P.S. (2004) Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries, *Labour Economics*, **11(3)**, 355-71.

Wood, A. (1995) How Trade Hurt Unskilled Workers, *Journal of Economic Perspectives*, **9(3)**.