Can over-education account for the positive association between education and within-groups wage inequality?

Santiago Budría[*] – University of Madeira and CEEAplA

December 2005

Abstract

International evidence shows that conditional wage dispersion increases as we move towards more educated individuals. This paper asks whether over-education accounts for this fact. The answer is a resounding 'no'.

Keywords: Returns to education, over-education, quantile regression.

JEL-Codes: C29, D31, I21

[*] I acknowledge financial support of the European Commission (EDWIN project HPSE-CT-2002-00108) and of the Portuguese Ministry of Education (FCT). I am most grateful to Corrado Andini for his valuable comments. Address correspondence to: Santiago Budría, Department of Economics, University of Madeira, Rua Penteada 9000-390, Funchal, Portugal. Phone: +351-291 705 055. Fax: +351-291 705 040. E-mail: sbudria@uma.pt.

0. Introduction

Recent international research has used the quantile regression technique to show that returns to education tend to be increasing when moving up along the wage distribution (Buchinsky, 1994, Machado and Mata, 2001, Pereira and Martins, 2002, 2004). This has been called 'the inequality increasing effect of education' (Machado and Mata, 2005, p. 457): if we give more education to workers who have the same observable characteristics but are located at different quantiles of the wage distribution, then their wages will become more dispersed.

This finding adds to the educational debate by warning policy makers that education may promote earnings differences. If conditional wage dispersion is higher among more educated individuals, then an educational expansion may raise overall wage inequality by raising the weight of the high-spread group. In this scenario, assessing the causes of the positive association between education and wage dispersion is of crucial importance, as countries where such causes are operating might wish to reverse the underlying causes.

This paper asks to what extent the over-education phenomenon accounts for the inequality increasing effect of education[1]. It has been documented that over-educated workers earn i) less than workers who have the same education but hold jobs for which they are adequately educated, and ii) more than workers in the same jobs who have less education (Alba, 1993, Sloane *et al.,* 1999, Dolton and Vignoles, 2000). Therefore, we expect that, relative to the adequately-educated, over-educated workers are located at lower quantiles of the earnings distribution and earn a lower return from their educational investment.

Then the question in this paper is: can over-education account for the tendency of education to be less rewarded at lower deciles of the wage distribution? This hypothesis

[1] In an international review, Groot and Van den Brink (2000) find that in most countries about one fourth of the working population is over-educated.

has not been empirically tested to date, even though many researchers have suggested it[2].

The next section describes the dataset and the definitions of over-education used in the paper. Section 2 calculates quantile returns to education and inspects whether over-education accounts for the differences in the returns across quantiles. Section 3 presents concluding remarks. Appendix A contains the detailed estimation results.

1. Data and definitions of over-education

I use the most recent wave (2001) of the European Community Household Panel dataset (ECHP). This survey contains useful information on personal and labour market characteristics, including maximum level of education completed, hours worked, and monthly net wage, for a variety of European countries. I take Portugal as case study, for in this country the inequality increasing effect of education has been found to be particularly acute (Pereira and Martins, 2002, 2004).

I focus on male wage earners aged between 18 and 60, who work normally between 15 and 80 hours a week, and are not employed in the agricultural sector. These restrictions leave us with a final sample of 2,042 individuals.

There are several approaches to measure the extent of over-education[3]. Following most other authors, I use the worker's self assessment regarding the match between the worker's skills and the firm's job requirements. In particular, I use two questions included in the ECHP,

- *Do you feel that you have skills or qualifications to do a more demanding job than the one you have now?*

[2] For instance, Machin (1996), Green *et al.* (1999), Pereira and Martins (2002) and Fersterer and Winter-Ebmer (2003). As Pereira and Martins (2002, p. 365) put it, "a situation where highly-schooled workers take jobs with a low skill requirement and consequent low pay would be consistent with this result".

[3] These approaches are basically three: job analysis, realized matches, and the worker's self-assessment. For further details, see Dolton and Vignoles (2000).

- *Have you had formal training or education that has given you skills needed for your present type of work?*

The previous questions provide us with different definitions of over-education (Alba and Blázquez, 2002). I define as

1) 'weakly over-educated' those workers that answer 'yes' to both the above questions
2) 'incorrectly qualified' those workers who answer 'no' to the first question and 'yes' to the second one[4]
3) 'strongly over-educated' those workers that answer 'yes' to the first question and 'no' to the second question, that is, those who are 'weakly over-educated' as well as 'incorrectly qualified'.

Table 1 reports the incidence of over-education by education levels. The most remarkable fact is that more educated workers are more likely to have excess education ('weakly over-educated') and, at the same time, less likely to work in jobs that are not related with their qualifications ('incorrectly qualified'). This finding warns researchers in the field against using a single measure of over-education, as different education groups are exposed to different types of educational mismatch.

2. Empirical models and results

I proceed by comparing the quantile returns to education of two wage equations. The first one corresponds to a standard specification,

$$ln\, w_i = \alpha_\theta + \delta_{\theta 1} exp_i + \delta_{\theta 2} exp_i^2 + \beta_{\theta 1} uppersec_i + \beta_{\theta 2} tertiary_i + e_{\theta i} \quad (1)$$

[4] Even though this situation does not correspond to over-education strictly speaking, it reflects a qualification mismatch that is worth exploring.

where *ln w* is the logarithm of hourly wages, *exp* is experience, while *uppersec* and *tertiary* are dummies that are activated if the maximum education level attained by the individual *i* is, respectively, upper secondary or tertiary education. The excluded education category is 'less than upper secondary education'[5].

The second specification extends the standard model to include over-education dummies,

$$ln\, w_i = \alpha_\theta + \delta_{\theta 1} exp_i + \delta_{\theta 2} exp_i^2 + \beta_{\theta 1} uppersec_i + \beta_{\theta 2} tertiary_i + \beta_{\theta 3} overuppersec_i + \beta_{\theta 4} overtertiary_i + u_{\theta i} \quad (2)$$

where *overuppersec* and *overtertiary* are, respectively, controls for over-education in the upper secondary and tertiary level[6]. I run this regression using the three alternative definitions of educational mismatch: weak over-education (Control 1), incorrect qualification (Control 2) and strong over-education (Control 3).

Figures 1 and 2 present the quantile returns to education arising from the different specifications[7]. The corresponding estimates are reported in Appendix A. The main finding is that returns to education are increasing over the wage distribution, and over-education can not account for this fact. In line with previous work, the coefficients of the standard model are increasing when moving from the lower to the upper quantiles. And, in the extended models, the quantile-return profile is as increasing over the wage distribution as in the standard model[8].

To provide a more detailed view, Tables 2 and 3 report return differentials between selected quantiles. Note how it is that differences across quantiles do not diminish after introducing controls for over-education. This result applies to both the secondary and the tertiary level, to any definition of over-education, and to any region of the wage distribution considered.

---

[5] In the ECHP the education variable is coded in three levels. These were constructed following the ISCED-97 classification (OECD, 2003).

[6] The use of over-education dummies in a wage regression is not new. See, for instance, Verdugo and Verdugo (1989), Dolton and Vignoles (2000) and Chevalier (2003).

[7] Note that these returns are measured in a comprehensive way, i.e., all indirect influences of education on wages – occupation, sector, etc. – are attributed to education itself.

[8] Though not reported, the over-education dummies are jointly significant in all regressions.

The last column of Tables 3 and 4 reports the F-test for the equality of coefficients at all quantiles. In the tertiary level, differences across quantiles remain statistically significant after controlling for over-education. In the secondary level, differences turn out to be non-significant in the standard model as well as in the extended models.

Overall, the results show that the magnitude and significance of the differences in the returns across quantiles remain practically unaffected after controlling for over-education. I argue, therefore, that the inequality increasing effect of education documented by previous works can not be attributed to the over-education phenomenon.

3. Conclusions

Returns to education are typically increasing when moving up along the wage distribution. While researchers have focused on the inequality implications of this finding, little attention has been paid to the analysis of its causes.

The results presented here warn policy makers that the positive association between education and within-groups earnings dispersion hinges on factors other than over-education. To the extent that these other factors are mostly unknown, further research needs to be done. A candidate explanation is ability. If ability interacts positively with education, then returns to education must be higher among workers at high-pay jobs, i.e., with more ability[9]. A second explanation has to do with qualifications. Differences in the returns within high-educated people may be due to differences in the type and quality of the qualifications provided by universities. According to this, those individuals located at the bottom part of the wage distribution, and thus earning lower returns, are those whose formal qualifications are less rewarded in the labour market.

[9] Arias *et al.* (2001) give partial support to this hypothesis, as they find that once ability is controlled for, the tendency of education to be more valued at high pay jobs, though still existent, becomes less acute.

In the future, the acquisition of new data containing detailed information on the individual's educational qualifications and ability tests would enormously help in the task of understanding the sources of wage dispersion within education groups in Portugal.

Appendix A

Table 1A. OLS and quantile returns to education – Standard model

	OLS	θ =.10	θ =.20	θ =.30	θ =.40	θ =.50	θ =.60	θ =.70	θ =.80	θ =.90
TERTIARY	**.949*****	.776***	.848***	.878***	.963***	1.003***	1.038***	1.070***	1.079***	1.100***
	(.047)	(.122)	(.067)	(.053)	(.049)	(.058)	(.043)	(.043)	(.045)	(.064)
UPPER SECONDARY	**.381*****	.282***	.308***	.334***	.387***	.398***	.413***	.412***	.419***	.389***
	(.031)	(.072)	(.047)	(.044)	(.034)	(.030)	(.029)	(.033)	(.034)	(.049)

Note: i) *** signals significant at the 1% level; ii) OLS estimation is heteroskedastic-robust; iii) quantile standard errors are obtained using 500 replications.

Table 2A. OLS and quantile returns to education – Model with Control 1

	OLS	θ =.10	θ =.20	θ =.30	θ =.40	θ =.50	θ =.60	θ =.70	θ =.80	θ =.90
TERTIARY	**.965*****	.898***	.846***	.867***	1.018***	1.090***	1.080***	1.111***	1.102***	1.224***
	(.083)	(.214)	(.114)	(.118)	(.106)	(.079)	(.075)	(.064)	(.099)	(.098)
UPPER SECONDARY	**.302*****	.221***	.251***	.240***	.262***	.320***	.316***	.320***	.337***	.324***
	(.043)	(.100)	(.046)	(.050)	(.053)	(.045)	(.047)	(.041)	(.062)	(.089)

Note: i) *** signals significant at the 1% level; ii) OLS estimation is heteroskedastic-robust; iii) quantile standard errors are obtained using 500 replications.

Table 3A. OLS and quantile returns to education – Model with Control 2

	OLS	θ =.10	θ =.20	θ =.30	θ =.40	θ =.50	θ =.60	θ =.70	θ =.80	θ =.90
TERTIARY	**.969*****	.845***	.847***	.913***	.986***	1.036***	1.070***	1.080***	1.102***	1.120***
	(.051)	(.113)	(.067)	(.059)	(.048)	(.059)	(.047)	(.043)	(.049)	(.069)
UPPER SECONDARY	**.494*****	.398***	.426***	.471***	.480***	.527***	.517***	.514***	.487***	.495***
	(.037)	(.080)	(.061)	(.049)	(.037)	(.038)	(.039)	(.030)	(.045)	(.045)

Note: i) *** signals significant at the 1% level; ii) OLS estimation is heteroskedastic-robust; iii) quantile standard errors are obtained using 500 replications.

Table 4A. OLS and quantile returns to education – Model with Control 3

	OLS	θ =.10	θ =.20	θ =.30	θ =.40	θ =.50	θ =.60	θ =.70	θ =.80	θ =.90
TERTIARY	**.963*****	.845***	.848***	.906***	.987***	1.032***	1.042***	1.080***	1.098***	1.118***
	(.049)	(.122)	(.063)	(.062)	(.049)	(.058)	(.048)	(.046)	(.049)	(.067)
UPPER SECONDARY	**.408*****	.325***	.324***	.355***	.405***	.411***	.437***	.447***	.449***	.454***
	(.034)	(.075)	(.051)	(.050)	(.038)	(.034)	(.035)	(.038)	(.039)	(.050)

Note: i) *** signals significant at the 1% level; ii) OLS estimation is heteroskedastic-robust; iii) quantile standard errors are obtained using 500 replications.

References

Alba, A. (1993), Mismatch in the Spanish Labour Market. Overeducation?, *Journal of Human Resources* 28, 259–278.

Alba, A. and M. Blázquez (2002), Types of Job Match, Overeducation, and Labour Mobility in Spain, in Büchel, F., A. de Grip and A. Meitens (eds), *Overeducation in Europe: Current Issues in Theory and Policy.* Edward Elgar Publishing, Cheltenham, UK.

Arias, O., K. Hallock and W. Sosa-Escudero (2001), Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data, *Empirical Economics*, 26(1), 7-40.

Buchinsky, M. (1994), Changes in the US Wage Structure 1963-1987: Application of Quantile Regression, *Econometrica* 62, 405-458.

Chevalier, A. (2003), Measuring over-education, *Economica* 70, 509-531.

Dolton, P. and A. Vignoles (2000), The incidence and effects of over-education in the UK graduate labour market, *Economics of Education Review* 19, 179-98.

Fersterer, J. and R. Winter-Ebmer (2003), Are Austrian Returns to Education Falling Over Time?, *Labour Economics* 10(1), 73-89.

Green, F., S. McIntosh and A. Vignoles (1999), *Overeducation' and Skills – Clarifying the Concepts*, Centre for Economic Performance, London School of Economics, 1-54.

Groot, W. and H. Van den Brink (2000), Overeducation in the Labor Market: A Meta-Analysis, *Economics of Education Review* 19, 149-158.

Machado, J. and J. Mata (2001), Earning functions in Portugal 1982-1994: evidence from quantile regressions, *Empirical Economics* 26, 115–134.

Machado, J. and J. Mata (2005), Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression, *Journal of Applied Econometrics* 20(4), 445-465.

Machin, S. (1996), Wage Inequality in the UK, *Oxford Review of Economic Policy*, 12(1), 47-64.

OECD (2003), *OECD Education at a Glance,* Annex 3: Sources methods and technical notes 2003 edition. OECD: Paris.

Pereira, P. and P. Martins (2002), Is there a Return-Risk Link in Education, *Economics Letters* 75, 31–37.

Pereira, P. and P. Martins (2004), Does Education Reduce Wage Inequality? Quantile Regressions Evidence from Fifteen European Countries, *Labour Economics*, 11(3), 355-371.

Sloane, P.J., H. Battu and P.T. Seaman (1999), Over-education, Undereducation and the British Labour Force, *Applied Economics* 31(11), 1437–1453.

Verdugo, R. and N.T. Verdugo (1989), The impact of surplus schooling on earnings: Some additional findings, *Journal of Human Resources* 24, 629-695.

Tables

Table 1. The incidence of over-education by education levels (%)

	WEAKLY OVER-EDUCATED	INCORRECTLY QUALIFIED	STRONGLY OVER-EDUCATED
TERTIARY	56.71	8.50	5.34
UPPER SECONDARY	54.18	51.44	26.69
LESS THAN UPPER SECONDARY	37.55	81.65	29.52

Table 2. Dispersion across quantiles – Tertiary education

	.90–.10	.90–.50	.50–.10	.70–.30	.70–.50	.50–.30	F-test[(a)]
STANDARD MODEL	0.324	0.097	0.227	0.192	0.067	0.125	F(4, 2060) = 3.30**
MODEL WITH CONTROL 1	0.326	0.134	0.192	0.244	0.021	0.223	F(4, 2058) = 2.37**
MODEL WITH CONTROL 2	0.275	0.084	0.191	0.167	0.044	0.123	F(4, 2058) = 2.89**
MODEL WITH CONTROL 3	0.273	0.086	0.187	0.174	0.048	0.126	F(4, 2058) = 2.76**

Note: (a) The F-test tests whether the coefficients at all quantiles (.10, .20,…, .90) are statistically different.

Table 3. Dispersion across quantiles – Secondary education

	.90–.10	.90–.50	.50–.10	.70–.30	.70–.50	.50–.30	F-test[(a)]
STANDARD MODEL	0.137	0.021	0.116	0.078	0.014	0.064	F(4, 2060) = 1.36
MODEL WITH CONTROL 1	0.116	0.017	0.099	0.080	0.000	0.080	F(4, 2058) = 1.25
MODEL WITH CONTROL 2	0.089	-0.040	0.129	0.043	-0.013	0.056	F(4, 2058) = 1.09
MODEL WITH CONTROL 3	0.124	0.038	0.086	0.092	0.036	0.056	F(4, 2058) = 1.51

Note: (a) The F-test tests whether the coefficients at all quantiles (.10, .20,…, .90) are statistically different.

Figures

Figure 1. OLS and quantile returns to tertiary education

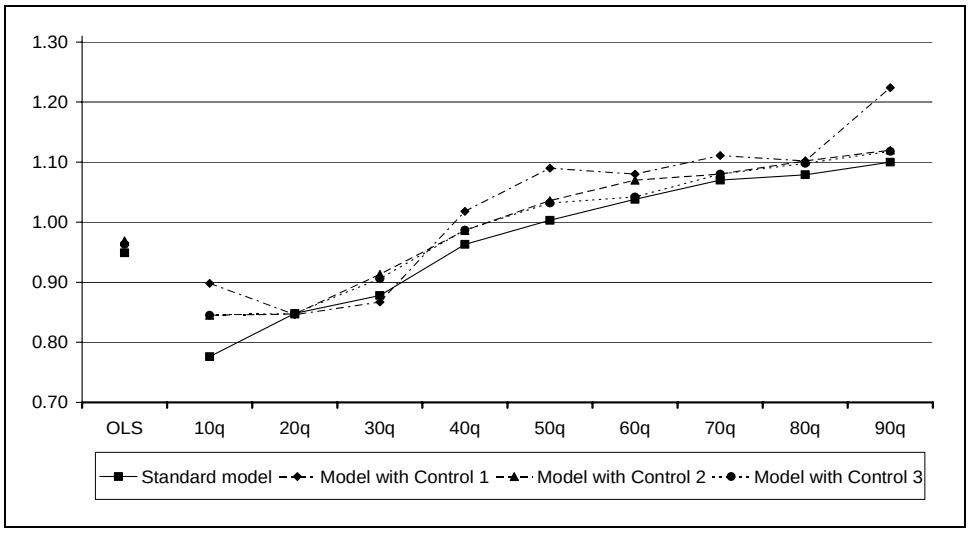

Figure 2. OLS and quantile returns to secondary education