Returns to schooling, overeducation and wage inequality: new evidence[1]

Raul Ramos

AQR-IREA, Universitat de Barcelona

Dpt. Econometria, Estadística i Economia Espanyola,

Avda. Diagonal 690, 08034 Barcelona (España)

rramos@ub.edu

Abstract

The objective of the paper is to analyze the evolution of wage inequality in Spain using data from three waves of the Spanish *Encuesta de Estructura Salarial* and to test the potential role of overeducation to explain the observed trend. The obtained results show a decreasing trend in overall inequality although results from quantile regressions support the hypothesis on the "inequality increasing effect of education", although not caused by overeducation. This evidence is also confirmed using decomposition techniques showing that during the considered period the "price effect" has been higher than the "quantity effect".

Key words

Returns to schooling, overeducation, wage inequality, quantile regression, decomposition techniques

JEL Codes

J24, J31, I2

[1] Raul Ramos would like to express his gratitude for the support received from Spain's Ministerio de Ciencia e Innovación (VI Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008-2011) through the action ECO2010-16006.

Returns to schooling, overeducation and wage inequality: new evidence

1. Introduction and objectives

Recent international research has shown that wage inequality is higher among more educated individuals. This has been termed as the "inequality increasing effect of education". This result is, in fact, counterintuitive as policies aimed to increase the educational level of the population are expected to reduce earnings equality as they increases the proportion of high wage workers. A potential explanation of this result could be related to skill mismatches. If a fraction of educated workers do jobs that require lower qualification levels than the ones they have, they will earn less that they should and, as a consequence, wage differences between more educated people will increase. The only job, to my knowledge, that has tested formally the effect of overeducation on waqe inequality is Budría (2011). He has analysed the case of Portugal using data from two different surveys and considering different ways of measuring skill mismatches and his results, however, do not support the hypothesis.

Spain is the OECD country with the highest proportion of overeducated workers (OECD, 2007) and some previous works such as Izquierdo and Lacuesta (2007) and Felgueroso et al. (2010) have shown that the main factor explaining the decreasing trend of wage inequality in Spain between 1995 and 2002 is the drop in returns to schooling due to the high increase of university degree holders in the same period that have become overeducated workers. These two features make the analysis of the Spanish case particularly interesting to confirm Budría's results.

Taking into account this previous background, the objective of the paper is to analyze the evolution of wage inequality in Spain between 1995 and 2006 and to test the potential role of overeducation to explain the observed trend. The rest of the paper is structured as follows: first, section 2 presents the datasets and the definitions of educational mismatch used in the analysis; next, the results of empirical analysis are shown in section 3 and, last, the paper ends summarizing the main conclusions and suggesting some directions for further research.

2. Data sources, wage inequality and measurement of educational mismatch

The microdata used in the study are drawn from the Spanish *Encuesta de Estructura Salarial* for 1995, 2002 and 2006 (hereafter EES). This survey is regularly carried out by the Spanish National Statistics Office (*Instituto Nacional de Estadística*) and constitutes the Spanish sample of the European Structure of Earnings Survey, a 4-yearly survey conducted in member states of the European Union according to a standard methodology. The EES collects comprehensive information, provided by the management of the establishments, on the level and structure of remuneration of employees; workers' demographic and job characteristics (in particular, nationality, sex, age, level of education, tenure in the firm, occupation, type of contract, supervision tasks indicators and full-time/part-time indicators) along with detailed information for each respondent's establishment (industry, size, region and type of collective agreement). In order to have a homogenous firm coverage, I have excluded from the analysis sectors not included in the EES 1995 and firms with less than 10 workers.

The wage measure used in this paper is the logarithm of the gross hourly wage, computed as the annual gross earnings in euros divided by the number of hours worked in each year. Gross earnings cover remuneration in cash paid directly and regularly by the employer at the time of each monthly wage payment, before deductions for tax and employee social security contributions. Earnings cover all payments different from overtime pay, including commissions, travelling expenses, premium payments for shift, night and weekend work and all bonuses and allowances paid regularly in each pay period, and also annual bonuses not paid regularly. Outliers have also been excluded from the analysis. Table 1 shows the number of individuals included in the analysis for each of the considered years.

TABLE 1

Table 2 present different measures of wage inequality for the three survey years. According to figures in this table, overall inequality has clearly decreased between 1995 and 2006. Inequality between workers with higher wages has also decreased, although when looking at the lower tail of the distribution (P50/25, P50/10), the decrease in inequality was only present between 1995 and 2002, while afterwards it has remained fairly constant.

TABLE 2

The first row of table 3 shows the value of the average years of schooling for 1995, 2002 and 2006. According to these figures, the level of human capital in the Spanish economy has clearly increased during this period. Taking all this into account, the obtained preliminary evidence seems to contradict the "inequality increasing effect of education" as accumulation of human capital and inequality have shown clearly opposite trends. However and as previously mentioned, it is also important to take into account which has been the trend of education mismatch.

Three different methods have been suggested in the literature in order to measure educational mismatch[2]: the objective method[3], the subjective method (direct and indirect)[4] and the statistical method[5].

The objective method is based on the analysis carried out by experts of the jobs and its educational requirements. Once these educational requirements are established, they are compared with the educational levels of workers in order to determine the degree of mismatch. The limitations of this method are related to the need of simplification and the need to update the analysis regularly.

The subjective method is based on the answers provided by workers to questions related to their educational level and the needs of the job they occupy. It is based on the self-assessment of workers according to different questions that try to assess the degree of educational mismatch both in a direct and indirect way. The subjective method does not require as much

[2] Hartog (2000) provides an excellent survey of the different methods and their main advantages and weaknesses.
[3] Some studies applying this method are Rumberger (1987), Hartog and Oosterbeek (1988), Kiker and Santos (1991), Kiker *et al.* (1997) and García Montalvo (1995).
[4] Some examples are Duncan and Hoffman (1981) and Sicherman (1991) for the United States, Hartog and Oosteerbeck (1988) for the Netherlands and Alba-Ramirez (1993) and García Serrano and Malo (1996) for Spain.
[5] Verdugo and Verdugo (1989) and Kiker *et al.* (1997) defend this method: the first using the average while the second using the mode. Mendes de Olivera *et al.* (2000) also recommend this method but using a "corrected mode".

information as the previous one, but it is based only in the perception of the own individual with respect to its work, which is a clear limitation of the approach.

Last, the statistical method consists of comparing the years of study of an individual with the most usual schooling years of employed workers in the same occupation. A worker is considered as overeducated if his/her schooling years surpass in more than one standard deviation the average value in his/her occupation (Verdugo and Verdugo, 1989). Kiker et al. (1997) suggested using the mode instead of the average, arguing that the adequate schooling level to perform an occupation is the one that dominates among the workers of that occupation.

Taking into account that the EES does not contain subjective information about educational mismatch, but it provides detailed information of the schooling levels and the occupation (2 CNO digits), I have used the statistical method based on the mode to obtain evidence about the level of educational mismatch observed in the Spanish Economy for 1995, 2002 and 2006. The obtained results are shown in table 3. Between 1995 and 2006 the proportion of properly educated workers has clearly decreased, while both the proportion of overeducated and infraeducated workers has increased. The next section analyses the effects of this trend on individual wages and the time evolution of the returns to schooling and its effect on wage inequality.

TABLE 3

3. Empirical evidence

The analysis of the returns to schooling and the effect of educational mismatch on wages requires an estimation of ORU-type (*over-required under-education*) wage equations. This is a semi-logarithmic Mincerian earnings function, in which the years of study are broken down into three components: over-education (s^o), suitable education (s^r), and under-education (s^u):

$$\ln W_i = \gamma + \beta^o \cdot s_i^o + \beta^r \cdot s_i^r + \beta^u \cdot s_i^u + \alpha \cdot z_i + u_i \qquad (1)$$

where ln W_i is the natural logarithm of the hourly wage of individual i; s_i are the years of study of individual i, but broken down into years of over-education (s_i^o), years required (s_i^r) and years of under-education (s_i^u), and z_i includes other individual variables affecting earnings, such as work experience, gender, working part-time or other aspects of the particular job. Finally, it is assumed that u_i is a random disturbance term distributed as a normal variable with zero expectation and constant variance.

Table 4 shows the returns per schooling year for 1995, 2002 and 2006 and also the returns to the years of required education, over-education and under-education. According to these results, the returns to schooling have dramatically fallen in the Spanish economy between 1995 and 2006. A similar evolution is observed for potential experience while the opposite trend is observed for tenure. This can be interpreted as evidence that firms are replacing generic human capital by on-the-job training. In fact, when the total number of schooling years is decomposed in required years of schooling, years of overeducation and years of undereducation, a similar picture is found although it is worth highlighting some interesting results: first, the returns to required years of schooling is more stable than returns to schooling. The returns to the years of overeducation are also positive although lower than the coefficient associated to the required ones, while the returns to the years of undereducation are negative (as expected). The differences between properly educated workers and those mismatched have been decreasing along time, a evolution that explains the decreasing trend in overall wage inequality.

TABLE 4

In order to confirm if this result holds across the entire wage distribution, quantile regressions have been estimated. The obtained results are shown in table 5. In line with previous works such as Budría (2011), I find that returns to education are increasing over the wage distribution even after controlling for the effect of educational mismatch, although conditional wage dispersion is slightly lower in this case. The differential effect between the .10 and the .90 quantile is reported in the last column of the table. According to the obtained results, differences across quantiles are statistically significant in all cases. Therefore, there is evidence to suggest that the tendency of education to be less rewarded in low pay jobs cannot be explained by the prevalence of educational mismatch.

TABLE 5

Last, in order to confirm the previous results regarding the role of educational mismatch to explain wage inequality, the Juhn et al. (1993) decomposition technique has been applied in order to decompose differences across the wage distributions between 1995 and 2002 and 2002 and 2006. This has involved to build hypothetical wages for each individual in each of the considered year and, next, to decompose the wage distribution differentials for each year according to the observed characteristics, the returns to the characteristics and the unobserved factors. The first term of the decomposition measures the portion of wage differentials which can be attributed to the differences in the two groups' observed characteristics. The second term captures the portion explained by the differences in the returns to the characteristics. Finally, the third term corresponds to the disparities in the effect of the unobserved factors. The obtained results are shown in table 6. The overall trend in wage inequality holds throughout the wage distribution, with the only exception of the lower tail between 2002 and 2006. According to the obtained results, schooling years have clearly played an important role to increase wage differentials along the whole distribution, although this "quantity" effect has been clearly offset by the "price effect", a result that is consistent with the previous evidence.

TABLE 6

4. Conclusions and further research

Using data from three waves of the Spanish Structure of Earnings Survey (SES), the objective of the paper was to analyze the evolution of wage inequality in Spain between 1995 and 2006 and to test the potential role of overeducation to explain the observed trend. With this aim, different specifications of Mincer equations considering the role of skill mismatches were estimated and the results from quantile regressions have shown that the returns to required schooling are increasing over the wage distribution and that differences between the 90%-10% percentiles are higher after controlling for educational mismatches. So, as in Budría (2011) the results support the hypothesis on the "inequality increasing effect of education", but not caused by overeducation. In fact, at the aggregate level, data show a continuous reduction on wage inequality between 1995 and 2006, which is not observed for other developed countries. The results obtained using the Juhn et al (2001) inequality decomposition technique have also shown that the reason why inequality has not increased in the considered period is that the fall in returns to schooling has offset the "inequality increasing effect of education" (the "price effect" has been higher than the "quantity effect").

5. References

Alba-Ramírez, A. (1993), "Mismatch in The Spanish Labor Market: Overeducation?", *Journal of Human Resources*, 28, pp. 259-278.

Budría, S. (2011), "Are Educational Mismatches Responsible for the 'Inequality Increasing Effect' of Education?", *Social Indicators Research*, forthcoming.

Duncan, G.; Hoffman, S. (1981), "The Economic Value of Surplus Education", *Economics of Education Review*, 1 (1), pp.75-86.

Felgueroso, F.; Jiménez-Martín, S.; Hidalgo, M. (2010), "Explaining the fall of the skill wage premium in Spain", Documento de Trabajo FEDEA 2010-19.

García Montalvo, J. (1995), *Empleo y Sobrecualificación: El Caso Español*, Documento de Trabajo FEDEA 95-20.

García-Serrano, C.; Malo, M. A. (1996), "Desajuste Educativo y Movilidad Laboral en España", *Revista de Economía Aplicada,*. 4 (11), pp. 105-131.

Hartog, J. (2000) "Over-education and earnings: Where are we, where should we go?", *Economics of Education Review*, 19, pp. 131-147.

Hartog, J.; Oosterbeek, H. (1988), "Education, Allocation and Earnings in the Netherlands: Overschooling?", *Economics of Education Review*, 7 (2), pp.185-194.

Izquierdo, M.; Lacuesta, A. (2007), "Wage Inequality in Spain: Recent development", ECB Working Paper 78.

Juhn, C., Murphy, K.; Pierce, B. (1993): "Wage inequality and the rise in returns to skill", *Journal of Political Economy*, 101 (31), pp. 410-442.

Kiker, B.; Santos, M. (1991), "Human Capital and Earnings in Portugal", *Economics of Education Review*, 10 (3), pp. 187-203.

Kiker, B., Santos, M.; Mendes de Oliveira, M. (1997), "Overeducation and Undereducation: Evidence for Portugal", *Economics of Education Review*, 16 (2), pp. 111-125.

Mendes de Olivera, M., Santos, M.; Kiker, B. (2000), "The Role of Human Capital and Technological Change in Overeducation", *Economics of Education Review*, 19, pp. 199-206.

OECD (2007), *International Migration Outlook*, Paris.

Rumberger, R. (1987). "The Impact of Surplus Schooling on Productivity and Earnings", *Journal of Human Resources*,.XXII (1), pp. 24-50.

Sicherman, N. (1991), "Overeducation in the Labour Market", *Journal of Labor Economics*, 9, pp. 101-122.

Verdugo, R.; Verdugo, N. (1989), "The Impact of Surplus Schooling on Earnings", *Journal of Human Resources*, 24 (4), pp. 629-643.

6. Tables

Table 1. Description of the EES sample

	1995	2002	2006
Total individuals	132,828	129,364	176,275
Individuals in the sample	124,883	105,853	120,107
Sample without outliers	121,557	103,737	117,705

Source: Own calculation from EES 1995, 2002 and 2006.

Table 2. Measures of wage inequality

	1995	2002	2006
Coefficient of variation of hourly wages	0.579	0.546	0.503
Gini index of hourly wages	0.289	0.268	0.255
P50/P10 of hourly wages	1.630	1.522	1.524
P50/P25 of hourly wages	1.339	1.285	1.286
P75/P50 of hourly wages	1.482	1.438	1.421
P90/P75 of hourly wages	1.457	1.422	1.388
P75/P25 of hourly wages	1.984	1.848	1.827
P90/P10 of hourly wages	3.519	3.114	3.006
Variance of the logarithm of hourly wages	0.241	0.202	0.186

Source: Own calculation from EES 1995, 2002 and 2006.

Table 3. Schooling levels and educational mismatch

	1995	2002	2006
Average years of schooling	7.916	8.634	10.423
Proportion of properly educated workers	0.569	0.691	0.368
Proportion of overeducated workers	0.241	0.222	0.329
Proportion of infraeducated workers	0.190	0.087	0.303

Source: Own calculation from EES 1995, 2002 and 2006.

Table 4. Mincer and ORU equations OLS robust estimates

Log of hourly wages	1995	2002	2006	1995	2002	2006
Schooling years	0.063	0.052	0.040			
Required years of schooling				0.0745	0.0672	0.0734
Years of overeducation				0.0459	0.0333	0.0385
Years of undereducation				-0.0341	-0.0263	-0.0166
Experience	0.027	0.020	0.013	0.0253	0.0182	0.0152
Experience sq	0.000	0.000	0.000	-0.0003	-0.0002	-0.0002
Tenure	0.007	0.016	0.018	0.0080	0.0156	0.0161
Tenure sq	0.000	0.000	0.000	0.0000	-0.0002	-0.0002
R-squared	0.465	0.468	0.427	0.498	0.507	0.481
Number of observations	121,557	103,737	117,705	121,557	103,737	117,705

Source: Own calculation from EES 1995, 2002 and 2006.

Gender, contract, activity sector, region, firm size, main market and type of collective agreement calso included.

Full time workers between 16 and 65 years old.

All coefficients are statistically different from zero at the usual significance levels.

Table 5. Mincer and ORU quantile regression estimates

Estimated returns to schooling

	OLS	Quantile regression			
		0.1	0.5	0.75	0.9	0.9-0.1
No control for skill mismatch						
1995	0.063	0.041	0.063	0.072	0.075	0.034
2002	0.052	0.030	0.050	0.060	0.067	0.037
2006	0.040	0.023	0.037	0.046	0.051	0.029
Control for mismatch						
1995	0.070	0.046	0.070	0.079	0.082	0.035
2002	0.066	0.044	0.065	0.075	0.081	0.036
2006	0.064	0.038	0.062	0.071	0.076	0.038

Estimates of the returns to required years of schooling

	OLS	Quantile regression			
		0.1	0.5	0.75	0.9	0.9-0.1
1995	0.075	0.050	0.075	0.085	0.088	0.038
2002	0.067	0.046	0.066	0.076	0.082	0.036
2006	0.073	0.048	0.071	0.079	0.085	0.037

Estimates of the returns to years of over-schooling

	OLS	Quantile regression			
		0.1	0.5	0.75	0.9	0.9-0.1
1995	0.046	0.032	0.045	0.052	0.056	0.024
2002	0.033	0.019	0.031	0.038	0.044	0.025
2006	0.039	0.022	0.037	0.044	0.051	0.029

Estimates of the returns to years of under-schooling

	OLS	Quantile regression			
		0.1	0.5	0.75	0.9	0.9-0.1
1995	-0.034	-0.028	-0.035	-0.034	-0.034	-0.006
2002	-0.026	-0.025	-0.027	-0.024	-0.023	0.002
2006	-0.017	-0.013	-0.016	-0.018	-0.019	-0.006

Source: Own calculation from EES 1995, 2002 and 2006.

Gender, contract, activity sector, region, firm size, main market and type of collective agreement calso included.

Full time workers between 16 and 65 years old.

Differences between the 90% and the 10% percentile are statistically significant at the usual significance levels.

Table 6. Juhn, Murphy and Pierce (1993) decomposition of wage inequality

2002-1995	d9010	d7525	d9050	d5010
Total difference (%)	-12.2%	-7.1%	-5.4%	-6.8%
Quantity effect (Q)	-1.1%	-1.0%	2.8%	-3.9%
Price effect (P)	-5.9%	-3.8%	-5.0%	-1.0%
Unobservables (%)	-5.2%	-2.4%	-3.2%	-2.0%
Breakdown of Q	d9010	d7525	d9050	d5010
Schooling years	3.4%	1.9%	2.5%	0.9%
Other human capital	-1.6%	-1.3%	0.7%	-2.3%
Other factors	-3.0%	-1.6%	-0.5%	-2.5%
2006-2002	d9010	d7525	d9050	d5010
Total difference (%)	-3.5%	-1.1%	-3.7%	0.2%
Quantity effect (Q)	0.6%	0.7%	-0.7%	1.3%
Price effect (P)	-4.5%	-2.5%	-3.0%	-1.5%
Unobservables (%)	0.4%	0.6%	0.0%	0.4%
Breakdown of Q	d9010	d7525	d9050	d5010
Schooling years	1.2%	1.4%	-0.5%	1.7%
Other human capital	-0.2%	-0.1%	0.4%	-0.6%
Other factors	-0.4%	-0.6%	-0.6%	0.2%

Source: Own calculation from EES 1995, 2002 and 2006.

Gender, contract, activity sector, region, firm size, main market and type of collective agreement calso included.

Full time workers between 16 and 65 years old.

All coefficients are statistically different from zero at the usual significance levels.