Education, Over-education, and Wage Inequality: Evidence for Spain

Santiago Budría[1] – University of Madeira and CEEAplA

Ana I. Moro-Egido[1] – University of Granada

September 2005

Abstract

In this paper we use the European Community Household Panel to explore the connection between education, over-education, and wage inequality in Spain for the period 1994-2001. Drawing on quantile regression, we find that higher education is associated with higher wage dispersion. This indicates that an educational expansion towards tertiary education may have a positive impact on overall wage inequality. We find that over-education contributes to enlarge wage differentials within university graduates. Still, over-education itself can not account for the positive association between higher education and wage dispersion. Finally, we show that over the last years the wage distribution of over-educated workers with university education became more dispersed. This process, together with an increasing proportion of over-educated workers, contributed towards overall wage inequality through the within dimension.

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

JEL-Codes: C29, D31, I21

[1] We thank seminar participants at the University of Granada for helpful comments. Santiago Budría acknowledges the financial support of the European Commission, EDWIN project HPSE-CT-2002-00108. The data used in this paper is taken from the 1994-2001 waves of the European Community Household Panel (Eurostat). Author correspondence: 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

Conventional wisdom asserts that policies aimed to increase average schooling levels are expected to reduce earnings inequality by increasing the proportion of high-wage workers. A more balanced distribution of education, it is argued, will result in a more balanced distribution of earnings. Even though such policies may reduce *average* differences between otherwise differently educated individuals, their final impact on overall inequality is not clear cut. Recent empirical research has shown that education may promote earnings differences. Using international data, Pereira and Martins (2004) find that in most countries more educated individuals show higher wage dispersion. A similar result is reported by Buchinsky (1994) for US, Gosling *et al.* (2000) for UK, and Hartog *et al.* (2001) for Portugal. These findings warn that an educational expansion is expected to raise wage inequality *within* similarly educated individuals. Thus, by increasing the weight of the high-spread group, higher educational levels may have a positive impact on overall wage inequality.

In this paper, we use recent Spanish data to investigate the effects of education on wages. To that purpose, we calculate Ordinary Least Squares (OLS) and quantile returns to education. Estimation by OLS assumes that the marginal impact of education on wages is constant over the wage distribution. In this case, the effect of having one additional level of education can be represented by a shift (to the right) of the conditional wage distribution. Quantile returns, in turn, measure the wage effects of education at different points of the distribution, thus describing changes not only in the location but also in the shape of the distribution. By combining OLS with quantile regression, we can assess the impact of education on wage inequality between and within groups: while OLS returns measure the average differential between education groups, differences in quantile returns represent the wage differential between individuals that are in the same group but located at different quantiles.

We find that workers with tertiary and upper secondary education earn a wage premium of about 40% and 16%, respectively, relative to the base case of individuals with less than upper secondary education. Moreover, returns to education are not constant over the conditional wage distribution. Workers at high-pay jobs earn substantially higher returns from a university degree than workers at low-pay jobs. This is interpreted as a positive impact of higher education on wage dispersion.

While an increasing body of the literature has analyzed the impact of education on wage inequality, little is still known about the impact of over-education. Typically, over-educated workers earn less than workers who have the same education but hold jobs for which they are adequately educated (Alba, 1993, Sloane *et al.* 1999, Sloane, 2002, Dolton and Silles, 2001). However, this evidence is based on average differentials and, up to date, little is known about the effects of over-education at different points of the wage distribution. This misses important information regarding the effects of over-education on the wage structure. Thus, for example, if the pay penalty of over-education is higher for the earnings-poor, then an expansion in the proportion of over-educated workers is expected to deteriorate the labour market position of already disadvantaged individuals. Such expansion is expected, therefore, to increase wage inequality by enlarging the lower tail of the wage distribution.

In this paper we take a step towards filling this gap and calculate the wage effects of over-education at different points of the wage distribution. We find that over-education is associated with a substantial pay penalty, contributing to enlarge wage differentials within education groups. Furthermore, the impact of over-education on wages (and, thus, on within-groups inequality) is not constant over the wage distribution.

Potentially, the over-education phenomenon can account for the positive association between higher education and within-groups dispersion. A situation where a proportion of high-educated individuals take jobs with low skill requirement and low pay would be consistent with having increasing returns to higher education over the wage distribution.

We test this hypothesis, and find that, even though over-education contributes to enlarge wage differentials among university graduates, it can not satisfactorily account for the higher wage dispersion in the tertiary group.

Finally, we analyze changes in the returns to education for the period 1994-2001. In recent years, average schooling levels increased dramatically in Spain[2]. Alongside this process, a large proportion of high-educated workers entered jobs that required less schooling than they had obtained. As these workers are typically penalised in terms of wages, it is intriguing to speculate that a rise in the proportion of over-educated workers resulted into higher wage dispersion within the high-educated.

Our estimates show that average returns to education decreased over the period considered. This process was not homogenous across population groups. It was more pronounced for workers with a university degree and over-educated women. As regards the effects of over-education on wage inequality, we find increasing wage dispersion within over-educated workers with a tertiary level. This process, together with the rising proportion of over-educated workers in the Spanish labour market, had a positive impact on within-groups inequality.

The rest of the paper is organized as follows. Section 1 briefly presents the dataset, variables, and estimating sample used in the paper. Section 2 motivates the analysis by reporting some facts on wage inequality in Spain. Section 3 presents the quantile regression model. Section 4 explores the relation between education and wage inequality using cross-sectional data for the year 2001. Section 5 documents changes in the conditional wage structure that have taken place over the last years. The role that education and over-education have had in shaping the wage distribution is discussed. Section 6 presents the concluding remarks. The paper includes an Appendix describing

[2] This process was intense during the nineties. Among the 25-64 age group, the proportion of individuals with less than upper secondary education fell from 78% in 1991 to 58% in 2001, while the proportion of individuals with completed tertiary education rose from 10% in 1991 to 24% in 2001 (OECD, 2004).

the data source and variables used in the analysis. It also includes a set of additional tables.

1. Data and Variables

We use the Spanish waves (1994-2001) of the European Community Household Panel (ECHP, henceforth). This survey contains personal and labour market characteristics, including monthly wage, education level, hours worked, tenure, experience, sector, firm size, marital status and immigrant condition. Individuals are asked to report the maximum level of education that they have completed according to three categories: less than upper secondary, upper secondary and tertiary education.

We focus on wage earners in the private sector, aged between 18 and 60, who work normally between 15 and 80 hours a week, and are not employed in the agricultural sector. Thus, self-employed individuals, as well as those whose main activity status is paid apprenticeship, training, and unpaid family worker have been excluded from the sample.

Table 1 contains a set of descriptive statistics. Relative to men, women work less hours, earn lower wages, are more educated, have less experience and tenure, and are more prone to work in the service sector.

---------- insert Table 1 about here -------

2. Some facts on wage inequality

During the second half of the eighties and the first half of the nineties wage inequality increased in Spain. This phenomenon was partially accounted for by the evolution of wage differentials across education groups (Barceinas *et al.,* 2000, Cantó *et al.*, 2000).

Using more recent data, we find that from 1994 to 2001 wage inequality tended to decrease. Changes were small, though. As the first columns of Table 2 show, the Gini index, the ratio between wages at the 1st and the 5th deciles, and the ratio between wages at the 1st and the 9th deciles fell, respectively, from .31, 1.91, and 4.46 to .30, 1.90, and 4.08. Differentiating between education groups, we find that wage inequality decreased within workers with upper secondary education or less and increased within workers with tertiary education. At the end of the period, wage inequality is highest among the high-educated.

---------- insert Table 2 about here -------

In computations not reported here, we found that inequality between education groups tended to decrease over the period considered. In 2001 tertiary and secondary educated workers earned, respectively, 47.0% and 13.2% more than workers in the lowest education category. In 1994, these differentials were 72.4% and 26.5%, respectively.

Overall, the observed decrease in overall wage inequality can be attributed to decreases in within-groups inequality and, more primarily, between-groups inequality. This evidence is taken from raw statistics, which do not control for the groups' characteristics. In what follows, we investigate what has been the marginal contribution of education and over-education to the observed patterns.

3. The model

The quantile regression model can be written as

$$ln\, w_i = X_i\beta_\theta + e_{\theta i} \qquad \text{with } Quant_\theta(\, ln\, w_i|X_i) = X_i\beta_\theta \tag{1}$$

where X_i is the vector of exogenous variables and β_θ is the vector of parameters. $Quant_\theta(ln\ w_i|\ X_i)$ denotes the θth conditional quantile of $ln\ w$ given X. The θth regression quantile, $0<\theta<1$, is defined as a solution to the problem

$$\underset{\beta \in R^k}{Min}\left\{ \sum_{i:y_i \geq x_i\beta} \theta |ln\, w_i - X_i\beta_\theta| + \sum_{i:y_i < x_i\beta} (1-\theta)|ln\, w_i - X_i\beta_\theta| \right\} \tag{2}$$

which, after defining the check function $\rho_\theta(z)=\theta z$ if $z \geq 0$ or $\rho_\theta(z)=(\theta-1)z$ if $z < 0$, can be written as

$$\underset{\beta \in R^k}{Min}\left\{ \sum_i \rho_\theta(\, ln\, w_i - X_i\beta_\theta) \right\} \tag{3}$$

This problem is solved using linear programming methods. Standard errors for the vector of coefficients are obtainable by using the bootstrap method described in Buchinsky (1998).

Our wage equation is

$$ln\, w_i = \alpha_\theta + \delta_{\theta 1} X_i + \beta_{\theta 1} uppersec_i + \beta_{\theta 2} tertiary_i + e_{\theta i} \tag{4}$$

where $ln\ w_i$ is the logarithm of the gross hourly wage and X_i is a vector of explanatory variables, including experience (and squared), tenure, marital status, immigrant condition, sector (industry or service), and firm size[3]. The construction of these variables is described in the Appendix. The dummies *uppersec* and *tertiary* are activated only when the individual's maximum level of education is, respectively, upper secondary or tertiary education. Thus, *less than upper secondary* is the excluded education category. The use of dummies rather than years of schooling is motivated by two reasons. First, the use of education groups highlights the non-linearities of the response of wages to additional education. As we show, dispersion across quantiles

[3] We do not include controls for occupation. As the acquisition of education allows individuals to access certain occupations that are better rewarded, we prefer to interpret these wage gains as a return to education rather than a return to occupation.

increases non-monotonically as we move towards higher levels of education. Second, we believe that the labour market reward to formal qualifications is better captured by levels rather than by years of schooling.

4. Empirical results

In this section we calculate OLS returns and conditional returns to education at five representative quantiles: .10, .25, .50, .75, and .90. This is done separately for men and women. To simplify the analysis, we do not control for female self-selection into the labour market[4].

In Table 3 we report the coefficients on education. The full sets of estimates are reported in Tables 1B and 2B in the Additional Tables section. As expected, more educated individuals earn significantly higher wages. The OLS returns to a tertiary and upper secondary level are, respectively, 38.8% and 14.7% for men and 43.0% and 17.4% for women.

---------- Insert Table 3 about here -------

However, returns to education are not constant over the wage distribution. The average return to tertiary education for men masks a return of 29.6% in the lowest quantile and 53.0% in the top quantile. To facilitate the analysis, in Figures 1 and 2 we plot the quantile-return profiles. For men and women, the coefficient of tertiary education is clearly increasing as we move towards higher quantiles. In other words, workers at

[4] This is also the perspective used in Gardeazabal and Ugidos (2004), who use quantile regression and Spanish data to analyze the gender wage gap over the wage distribution. In a similar work, De la Rica, Dolado and Llorens (2005) control for female selectivity and find, using the ECHP, that the inverse of the Mill's ratio is not significant in the wage equation. Overall, the impact of the correction for sample selection on the return to schooling is found to be minor in most Spanish studies. Thus, for example, Barceinas *et al.* (2000) find that controlling for selectivity reduces the return to an additional year of schooling from 8.3% to 7.4%.

high-pay jobs earn substantially higher returns from university education than workers at low-pay jobs. This finding implies that if we give tertiary education to workers that are apparently equal but located at different quantiles, then their wages will become more dispersed. Thus, by raising the weight of the high-spread group, an educational expansion towards tertiary education may increase overall wage inequality. In contrast, the coefficient of secondary education exhibits low dispersion across quantiles. An educational expansion towards secondary education is expected, therefore, to have a more limited impact on within-groups dispersion.

---------- Insert Figures 1 and 2 about here -------

4.1. Over-education

In this section, we differentiate between adequately-educated and over-educated workers, depending on whether or not they are in jobs commensurate with their qualifications. We define the pay penalty of over-education as the differential in the return to education earned by these two groups. By calculating this penalty at different quantiles we can asses the impact of over-education on within-groups dispersion at different points of the wage distribution.

Most researchers have defined an individual as being over-educated if he has education in excess of that required to do his job. However, there are several approaches to measure the degree of over-education, each of one having its own limitations[5]. Following most other authors, we use the worker's self assessment regarding the match between the worker's skills and the firm's job requirements. In particular, we use two questions included in the ECHP,

[5] These approaches are basically three: job analysis, the statistical approach, and the worker's self-assessment. For a description of these methods, see Hartog (2000) and Sloane (2002).

- *Do you feel that you have skills or qualifications to do a more demanding job than the one you have now?*
- *Have you had formal training or education that has given you skills needed for your present type of work?*

We consider as over-educated workers those who answer "yes" to the first question and "no" to the second question[6]. According to this definition, in 2001 the proportion of over-educated workers was 27.1%. This figure is in line with those reported for other countries[7].

To explore the impact of over-education on wages, many researchers introduce an over-education dummy in the wage regression (Verdugo and Verdugo, 1989, Dolton and Vignoles, 2000, Chevalier, 2003). However, the effects of over-education may differ importantly across education levels. To take this into account, we use the following specification

$$ln\, w_i = \alpha_\theta + \delta_{\theta 1} X_i + \beta_{\theta 1} upper\, sec_i + \beta_{\theta 2} tertiary_i + \beta_{\theta 3} overuppersec_i + \beta_{\theta 4} overtertiary_i + e_{\theta i} \qquad (5)$$

where *uppersec* and *tertiary* are activated if the worker is not over-educated and has, respectively, upper secondary or tertiary education, and *overuppersec* and *overtertiary* are activated if the worker is over-educated and has upper secondary or tertiary education[8].

[6] Sloane (2002) warns that in some measures of over-education "reference is made to the level of education rather than the type of education. Thus a worker may still be mismatched if the level of education is appropriate, but its type inappropriate, such as an English graduate being hired as a statistician" (p. 7). By considering simultaneously the above questions, we take into account both the level and type of education. These questions have been also used by Alba and Blázquez (2002), who exploit the panel structure of the ECHP to investigate the relation between job promotions and over-education.

[7] See Groot and Van den Brink (2000) for a review of international studies analyzing the extent of over-education in several countries.

[8] An alternative specification is the ORU model, in which years of schooling are decomposed into required, surplus and deficit years of schooling in relation to those necessary to do the job. Relative to our specification, the ORU model has one advantage: it controls for the amount of over-education. However, it presents two shortcomings. First, it assumes that the impact of over-education on wages is constant

The results in Table 4 match a priori expectations. Over-educated workers earn less than similarly educated workers whose skills are fully utilized. The pay penalty of over-education, measured by the return differential between adequately-educated and over-educated workers, is substantially large: 17.6% for men and 26.7% for women in the tertiary level, and 14.1% for men and 12.7% for women in the upper secondary level[9]. It is worth noting that upper secondary education fails to attract a significant wage premium if the worker is over-educated. Moreover, over-educated workers with university education earn wages that are similar to those earned by adequately-educated workers with upper secondary education.

---------- Insert Table 4 about here -------

Our estimates show that, by driving a wedge between matched and mismatched workers, over-education has a positive impact on wage inequality within educational levels. The quantile analysis reveals that, moreover, this impact is not constant over the wage distribution. This can be better seen in Figures 3 and 4. The pay penalty of over-education differs across quantiles. For men in the tertiary group, it is much lower at the top quantile than at the other quantiles. This implies that among university graduates, over-education contributes to enlarge the middle and bottom part of the wage distribution. For men in the secondary group, the effects of over-education are larger at the top quantile than at the other quantiles. Within this group, thus, over-education contributes to enlarge the top tail of the wage distribution. For women, the pay penalty of over-education fluctuates across quantiles without a clear tendency.

---

across education levels. Second, and more important, in the quantile regression framework the use of years of schooling rather than levels of education would assume that the marginal impact of education (and over-education) on within-groups dispersion is the same for all education levels. Clearly this is not the case, since, as we show, tertiary education has a much larger impact on within-groups dispersion than secondary and primary education.

[9] Our estimates for Spain are in line with previous estimates for UK. Battu *et al.* (1999) and Dolton and Vignoles (2000) use over-education dummies and find a pay penalty that ranges from 15% to 20%. Chevalier (2003) differentiates between types of over-education, and reports estimates that range from 5% to 29%.

---------- Insert Figures 3 and 4 about here ------

There are two additional findings arising from the quantile analysis. First, the differential between the tertiary and secondary level is increasing over the wage distribution. This pattern holds for men and women and for all possible comparisons between adequately-educated and over-educated workers. Thus, for example, over-educated men with higher education earn on average 4.6% and 18.7% more, respectively, than adequately-educated men with a secondary level and over-educated men with a secondary level. However, these differentials are -.2% and 13.1% in the bottom quantile and as high as 22.9% and 48.9% in the top quantile. According to these results, an educational expansion from secondary to tertiary education is expected, regardless of the educational match attained by university students, to increase overall within-groups dispersion.

Second, we can test whether over-education is responsible for the higher wage dispersion in the tertiary group. If over-educated workers contribute to enlarge wage differentials among university graduates, then we should expect lower wage inequality among the group of university graduates who are not over-educated. As Table 4 shows, the return to a tertiary level earned by the adequately-educated is clearly increasing over the wage distribution, going from 32.9% to 52.2% among men and from 48.2% to 58.8% among women. This pattern indicates that over-education itself can not explain why conditional wage dispersion is substantially higher among university graduates than among workers with less education. Notwithstanding this, over-education contributes to enlarge wage differentials among university graduates. Using equation (4), the .90-.10 differential for the tertiary group was found to be 23.4% for men and 19.3% for women. However, using equation (5) this differential falls to 19.3% for adequately-educated men and 10.6% for adequately-educated women. It seems, therefore, that in a world without educational mismatches the relation between higher education and wage dispersion would be still existent, though less acute.

5. Changes over time

In the following we examine how the impact of education and over-education on wages evolved from 1994 to 2001. We draw on equation (5) to perform the analysis.

In Table 5 we summarize changes in OLS returns to education. The general pattern is decreasing earnings differentials across education levels. As regards the impact of over-education on within-groups dispersion, we find different patterns. For men, the pay penalty of over-education decreased from 28.4% to 17.6% in the tertiary group and increased from 3.8% to 14.1% in the upper secondary group. For women, the pay penalty of over-education increased from 13.6% to 27.6% in the tertiary group and decreased from 15.7% to 12.7% in the secondary group. In other words, the role of over-education in generating within-groups dispersion became more acute among men with upper secondary education and women with higher education and less acute among men with tertiary education and women with upper secondary education.

The previous analysis is confined to changes in average wage differentials. Thus, it disregards changes in the conditional wage distribution of different population groups. To address this issue, in Figures 5-8 we analyze changes in quantile returns. The estimates for 1994 are reported in Table 3B of the Additional Tables section.

---------- Insert Figures 5-8 about here -------

We start by analyzing the group of adequately-educated workers. As Figures 5 and 6 show, education premia decreased, particularly for university graduates. Among men with university education and women with secondary education, this process was more pronounced at the middle and upper quantiles than at the bottom quantile. As a consequence, wage dispersion fell in these two groups. Among women with university education the decline of returns was more pronounced at the lower quantiles, contributing towards wage dispersion. Overall, these results indicate that over the

sample period the tendency of higher education to be more valued at high-pay jobs became less acute for men and more acute for women.

In Figures 7 and 8 we focus on the sample of over-educated workers. The most prominent result is the sharp increase in wage inequality within university graduates. This process is driven by a sharp increase in the return at the top quantile among men and a drammatical decrease in the returns at the bottom quantiles among women. Even though assessing the causes of this result is beyond the scope of this paper, we may advance some candidate explanations. It is likely that the educational expansion occurred over the last decades has been parallel to an increasing proportion of low ability individuals accessing higher education, resulting in a more dispersed distribution of ability and skills. At the same time, there have been important changes in the types of qualifications provided by universities (Subirats, 2001). Education, social, and media studies have become increasingly demanded. Whilst these courses have become popular, students in these areas may have faced difficulties in finding adequate jobs and, consequently, have taken up mismatched work. Moreover, there has been an expansion of the institutions providing higher education. This has probably resulted into greater variation in the type, quality, and skills of over-educated workers with a university degree, insofar as an important fraction of job entrants have taken mismatched work.

From the previous analysis, we can draw important conclusions regarding the role that tertiary education has had in shaping wage inequality. The wage premium earned by tertiary educated workers decreased over the period considered, which contributed to reduce wage inequality through the between- dimension. As regards within-groups inequality, different patterns across population groups emerge. On the one hand, changes in conditional dispersion were small among adequately-educated workers. Still, for university graduates in this group, wage inequality decreased among men and increased among women. On the other hand, changes were substantial among over-educated workers with a university degree. For this group, the wage distribution became increasingly dispersed, contributing towards overall wage inequality. In the case of

women, this process was reinforced by an increase in the wage gap between the adequately-educated and the over-educated.

Furthermore, changes in the structure of pay were accompanied by changes in the composition of the workforce. On the one hand, the number of university graduates rose. Given the positive association between tertiary education and wage dispersion, this educational expansion had a positive impact on overall within-groups inequality. Alongside this process, the proportion of over-educated workers increased. From 1994 to 2001, it went from 16.7% to 26.6% among men and from 19.1% to 27.8% among women. Given that these workers are penalised in terms of wages, and the increasing wage dispersion among university graduates in this group, the expansion of the over-education phenomenon contributed to raise overall wage inequality through the within-dimension.

6. Conclusions

In this paper we explored the connection between education, over-education, and wage inequality in Spain. We used the quantile regression technique to calculate returns to education at different points of the wage distribution. Our findings were several. First, the returns to tertiary education are highly increasing as we move towards higher quantiles of the conditional wage distribution. Since higher education is a better investment for the earnings-rich, an educational expansion towards tertiary education is expected, ceteris paribus, to boost overall within-groups dispersion.

Second, over-educated workers earn substantially less than their adequately educated and matched peers. This market outcome results into higher wage dispersion within educational groups. Moreover, the impact of over-education on within-groups inequality is not constant over the wage distribution. Among men with university education, this impact is larger at the bottom tail of the distribution, while the opposite occurs among men with upper secondary education.

Third, we asked whether or not over-education is responsible for the positive association between higher education and within-groups dispersion. We found that, even though over-education contributes to enlarge wage differentials among university graduates, wage inequality is also large among adequately-educated workers in the tertiary level. We concluded, therefore, that over-education itself can not explain why the conditional wage distribution of workers with university education is more disperse than the conditional wage distribution of workers with less educational attainment. Arguably, there are a number of other potential causes. Higher dispersion in skill and ability requirements among individuals with higher education, and differences in the types and qualities of qualifications awarded by universities could account for some of the observed variation. A complementary view is that higher education does not function as a screening device and, consequently, the group of university graduates is rather heterogeneous in terms of ability. If ability interacts with schooling, then returns to education must be higher among workers at high-pay jobs, i.e., with more ability. The acquisition of new data containing detailed information on the type of qualifications, ability scores, school quality, and occupational categories may help to test these hypotheses.

Fourth, we investigated to what extent the over-education phenomenon has contributed to increase wage inequality in Spain. The proportion of over-educated workers rose markedly over the years considered. Alongside this process, important changes in the structure of pay took place. Among university graduates, wage dispersion rose substantially within over-educated workers and, in the case of women, the pay penalty of over-education increased. These changes contributed towards wage dispersion.

A clear implication of our analysis regards the demand for education. To the extent that prospective students are not aware of the characteristics which will place them at some point of the wage distribution, the returns to their educational investment are largely unpredictable. This uncertainty is reinforced by the fact that they may end up in jobs for which they are over-educated and, thus, earn lower wages than their matched peers with

similar education. It seems, therefore, that from an individual perspective investing in education is subject to a substantial amount of wage risk.

There are a number of reasons why policy makers should be concerned about over-education. Like unemployment, over-education could be seen as an indicator of the extent to which the labour market underutilises available human capital. Dolado *et al.* (2000) report that over the last years less educated workers have been crowded out from their traditional jobs towards jobs that require no educational qualifications, while a large proportion of high-educated workers has entered jobs that require less schooling than they have obtained. Considering education as an individual investment, this process has resulted in a huge waste of resources. Our results indicate that, in addition, it has had a positive impact on within-groups inequality.

We are aware that, from an individual perspective, over-education may be a temporary phenomenon. High-educated individuals may take up mismatched work to acquire other forms of human capital and move into matched work as their work experience increase. This "stepping stone" hypothesis has been suggested by Sicherman (1991). What is not in doubt, however, is that for some workers over-education is a long-run phenomenon[10]. Our findings indicate that if the incidence of over-education continues to increase, further and persistent wage inequality whithin high-educated workers can be expected.

Finally, there is evidence that in Europe the incidence of over-education has increased over time (Hartog, 2000). Unfortunately, existing knowledge on the connection between over-education and wage inequality is still too limited. Our analysis for Spain can be easily extended to other European countries that integrate the ECHP. The data harmonization provided by this dataset would allow for a straight comparison between different countries. To our eyes, assessing the impact that the over-education

[10] Thus, for example, Dolton and Vignoles (2000) and Sloane *et al.* (1999) have shown for the UK that a substantial fraction of workers remain in jobs for which they are over-qualified during long periods.

phenomenon is having in the European wage structure is a compelling task for future research.

Appendix. Description of data source and estimating samples

The European Community Household Panel (ECHP) is available from 1994 to 2001 for Spain. It has a sample size of about 5,000 households and 14,000 individuals, who are interviewed over time. They report personal and family characteristics, including marital and educational status, as well as gross monthly wages and worked hours. We have dropped workers with a monthly wage rate that is less than 10% or over 10 times the average wage. This correction for outliers affects only 1.9% of the total sample. In the following, we describe the variables used in the analysis,

Gross hourly wage. Defined as monthly gross salary in the main job divided by four times the weekly hours worked in the main job.

Level of education. Individuals are asked to report the maximum level of completed schooling, according to three categories: less than upper secondary, upper secondary, and tertiary education. These education categories are constructed following the ISCED-97 classification.

Experience. Defined as age minus age of first job.

Tenure. Defined as the difference between the year of the survey and the year of the start of the current job. We have constructed three categories: from 1 to 4 years, from 5 to 14 years, and 15 years or more.

Married. It is a dummy that takes the value 1 if the individual is married, zero otherwise.

Immigrant. It is a dummy activated if the individual was born in a foreign country.

Industry. It is a dummy that takes the value 1 if the individual works in the industry sector, zero if he works in the service sector. The agricultural sector, which accounted for 6% of the working population in 2001, was dropped on the account of the particularities of this sector.

Firm size. Individuals are asked to report the number of employees that actually work in their firm. We have constructed four categories, from 1 to 19 employees, from 20 to 99 employees, from 100 to 499 employees, and 500 employees or more.

References

Abadie, A. (1997), Changes in Spanish Labor Income Structure during the 1980's: A Quantile Regression Approach, *Investigaciones Económicas* 21(2), 253-272.

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.

Asplund, R. and E. Barth (2005), *Education and Wage Inequality in Europe: A Literature Review.* The Research Institute of the Finnish Economy, Helsinki 2005.

Barceinas, F., J. Oliver, J.L. Raymond and J.L. Roig (2000), *Private rates of return to human capital in Spain: new evidence.* Fundación de las cajas de ahorro confederadas para la investigación económica y social, Working Paper 162/2000.

Battu H., C. Belfield and P. Sloane (1999), "Over-education among graduates: a cohort view", Education Economics, 7, 21-38.

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

Buchinsky, M. (1998), Recent advances in quantile regression models: a practical guideline for empirical research, *Journal of Human Resources* 33, 88-126.

Cantó, O., A.R. Cardoso, C. Farinha, M. Izquierdo and J.F. Jimeno (2000*), Integration and Inequality: Lesson from the Accessions of Portugal and Spain to the EU*. FEDEA Working Paper 2000-10.

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

De la Rica, S., J.J. Dolado and V. Llorens (2005), *Ceiling and Floors: Gender Wage Gaps by Education in Spain*, IZA Discussion Paper 1483.

Dolado, J.J., F. Felgueroso and J.F. Jimeno (2000), *Explaining Youth Labor Market Problems in Spain: Crowding-out, Institutions, or Technology Shifts?,* IZA Discussion Paper 142.

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

Dolton, P. and M. Silles (2001), *Over-education in the Graduate Labour Market: Some Evidence from Alumni Data.* Centre for the Economics of Education, Discussion Paper Series, 9.

Febrer, A., and J. Mora (2005), *Wage Distribution in Spain, 1994-1999: An Application of a Flexible Estimator of Conditional Distributions,* IVIE Working Papers, EC-2005-04.

Gardeazábal, J. and A. Ugidos (2004), Measuring the gender gap at different quantiles of the wage distribution, *Journal of Population Economics* 44, 224–236.

Gosling, A., S. Machin and C. Meghir (2000), The changing distribution of Male Wages in the UK, *Review of Economic Studies* 67, 635-666.

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

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

Hartog, J., P. Pereira and J.A. Vieira (2001), Changing Returns to Education in Portugal during the 1980s and Early 1990s: OLS and Quantile Regression Estimators, *Applied Economics* 33, 1021-2037.

OECD (2004), *Education at a Glance*. Paris.

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.

Sloane, P. J. (2002), Much Ado About Nothing? What does the Over-education Literature really Tell Us? Keynote Address, International Conference on Over-education in Europe: What Do We Know? 22-23 November, Berlin.

Sicherman, N. (1991), Over-education in the Labor Market, *Journal of Labor Economics* 9(2), 101-122.

Subirats, J. (2001), Universidad en España: ¿época de cambios o cambio de época?, *Educar 28,*11-39.

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. Descriptive statistics (2001)

	Men		Women	
	Mean	St. dev	Mean	St. dev
No. of observations	1,749		937	
Age	37.00	10.54	34.13	9.58
Married	0.64	0.48	0.50	0.50
Immigrant	0.01	0.78	0.02	0.70
Weekly hours	43.32	7.27	38.70	6.19
Ln (monthly wage)	14.48	11.63	14.07	11.16
Ln (hourly wage)	6.78	8.26	6.51	7.95
Experience	19.14	12.25	14.07	10.91
Education				
Tertiary	25.27		37.25	
Secondary	22.18		26.25	
Primary	52.54		36.50	
Tenure				
0-4 years	56.83		64.57	
5-14 years	21.78		22.84	
≥ 15 years	9.26		5.12	
Sector				
Industry	56.95		25.72	
Services	43.05		74.28	
Firm size				
1-19 employees	46.77		53.26	
20-99 employees	29.90		25.29	
100-499 employees	14.87		14.73	
≥ 500 employees	8.46		6.72	

Table 2. The evolution of inequality by education groups (from 1994 to 2001)

	Total Sample			Tertiary			Upper Secondary			Less than upper secondary		
	Gini	**W1/W5**	**W1/W9**	Gini	W1/W5	W1/W9	Gini	W1/W5	W1/W9	Gini	W1/W5	W1/W9
2001	**0.30**	**1.90**	**4.08**	0.30	1.80	4.21	0.29	1.80	4.07	0.24	1.58	3.23
1994	**0.31**	**1.91**	**4.46**	0.28	1.78	3.56	0.29	1.86	4.14	0.26	1.64	3.93

Table 3. Conditional returns to education (2001)

MEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY	**0.388*****	0.296***	0.307***	0.336***	0.464***	0.530***
	(0.030)	(0.069)	(0.035)	(0.034)	(0.038)	(0.042)
UPPER SECONDARY	**0.147*****	0.105***	0.093***	0.152***	0.182***	0.162***
	(0.028)	(0.057)	(0.038)	(0.028)	(0.032)	(0.041)
WOMEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY	**0.430*****	0.375***	0.430***	0.414***	0.479***	0.568***
	(0.043)	(0.112)	(0.064)	(0.052)	(0.045)	(0.060)
UPPER SECONDARY	**0.174*****	0.239***	0.172***	0.154***	0.150***	0.198***
	(0.041)	(0.106)	(0.058)	(0.046)	(0.044)	(0.061)

Note: i) * signals significant at the 10% level, ** signals significant at the 5% level, and *** signals significant at the 1% level; ii) standard errors in parenthesis; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors are obtained using 500 replications. The reference individual is a worker with less than upper secondary education, less than 5 years of tenure, single, not immigrant, working in the service sector in a firm with less than 20 employees.

Table 4. Conditional returns to education and over-education by education levels (2001)

MEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY ADEQUATELY-EDUCATED	**0.415*****	0.329***	0.335***	0.378***	0.495***	0.522***
	(0.031)	(0.075)	(0.037)	(0.039)	(0.040)	(0.045)
TERTIARY OVER-EDUCATED	**0.239*****	0.185	0.184**	0.222***	0.242***	0.480***
	(0.071)	(0.161)	(0.082)	(0.056)	(0.085)	(0.140)
UPPER SECONDARY ADEQUATELY-EDUCATED	**0.193*****	0.187***	0.121***	0.173***	0.204***	0.251***
	(0.032)	(0.062)	(0.038)	(0.032)	(0.041)	(0.063)
UPPER SECONDARY OVER-EDUCATED	**0.052**	0.054	-0.049	0.077	0.067**	-0.009
	(0.040)	(0.071)	(0.068)	(0.060)	(0.035)	(0.057)
WOMEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY ADEQUATELY-EDUCATED	**0.472*****	0.482***	0.470***	0.449***	0.541***	0.588***
	(0.045)	(0.115)	(0.060)	(0.057)	(0.050)	(0.060)
TERTIARY OVER-EDUCATED	**0.205****	0.281**	0.070	0.218**	0.252***	0.245***
	(0.081)	(0.170)	(0.168)	(0.086)	(0.090)	(0.123)
UPPER SECONDARY ADEQUATELY-EDUCATED	**0.210*****	0.303***	0.229***	0.162***	0.167***	0.233***
	(0.043)	(0.108)	(0.057)	(0.044)	(0.055)	(0.076)
UPPER SECONDARY OVER-EDUCATED	**0.083**	0.204	0.021	0.025	0.114**	0.052
	(0.069)	(0.182)	(0.108)	(0.093)	(0.056)	(0.098)

Note: i) * signals significant at the 10% level, ** signals significant at the 5% level, and *** signals significant at the 1% level; ii) standard errors in parenthesis; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors are obtained using 500 replications. The reference individual is a worker with less than upper secondary education, less than 5 years of tenure, single, not immigrant, working in the service sector in a firm with less than 20 employees

Table 5. OLS returns to education and pay penalty of over-education (1994-2001)

MEN WITH TERTIARY EDUCATION	1994	2001	2001–1994
ADEQUATELY-EDUCATED	50.4	41.5	-8.9
OVER-EDUCATED	22.0	23.9	1.9
PAY PENALTY OF OVER-EDUCATION	28.4	17.6	-10.8
MEN WITH UPPER SECONDARY EDUCATION	1994	2001	2001–1994
ADEQUATELY-EDUCATED	20.3	19.3	-1.0
OVER-EDUCATED	16.5	5.2	-11.3
PAY PENALTY OF OVER-EDUCATION	3.8	14.1	10.3
WOMEN WITH TERTIARY EDUCATION	1994	2001	2001–1994
ADEQUATELY-EDUCATED	56.6	47.2	-9.4
OVER-EDUCATED	43.0	20.5	-22.5
PAY PENALTY OF OVER-EDUCATION	13.6	27.6	14.0
WOMEN WITH UPPER SECONDARY EDUCATION	1994	2001	2001–1994
ADEQUATELY-EDUCATED	24.0	21.0	-3.0
OVER-EDUCATED	8.3	8.3	0.0
PAY PENALTY OF OVER-EDUCATION	15.7	12.7	-3.0

Figures

Figure 1 – Conditional returns to education – Men (2001)

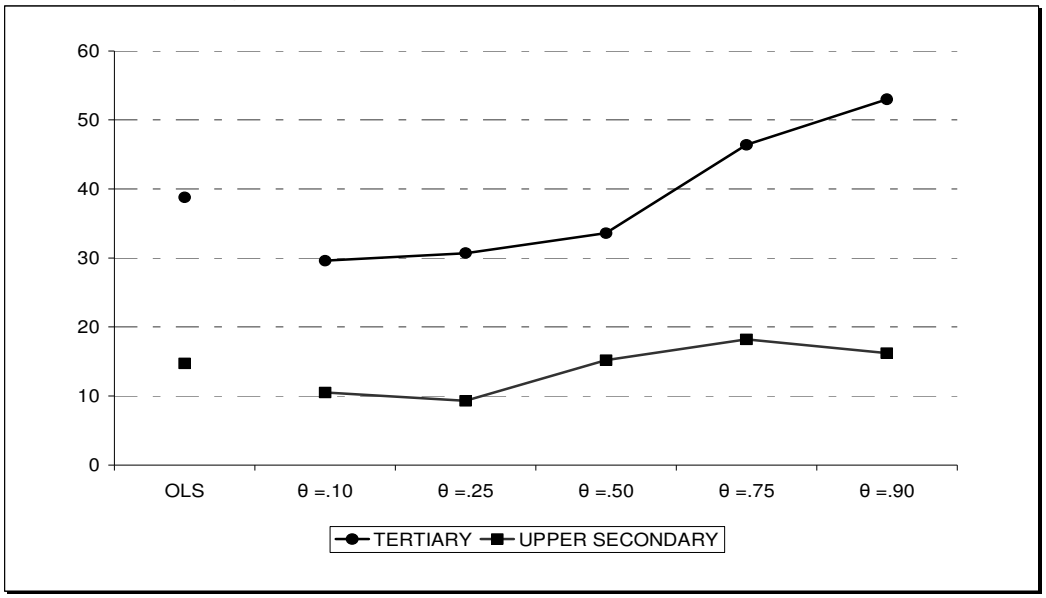


Figure 2 – Conditional returns to education – Women (2001)

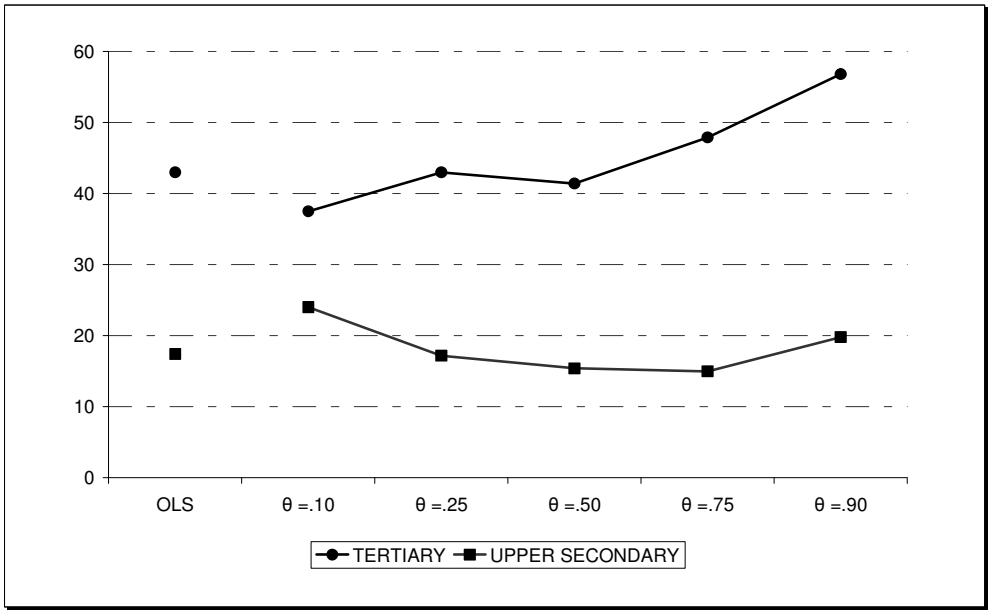


Figure 3 – Conditional returns to education and over-education – Men (2001)

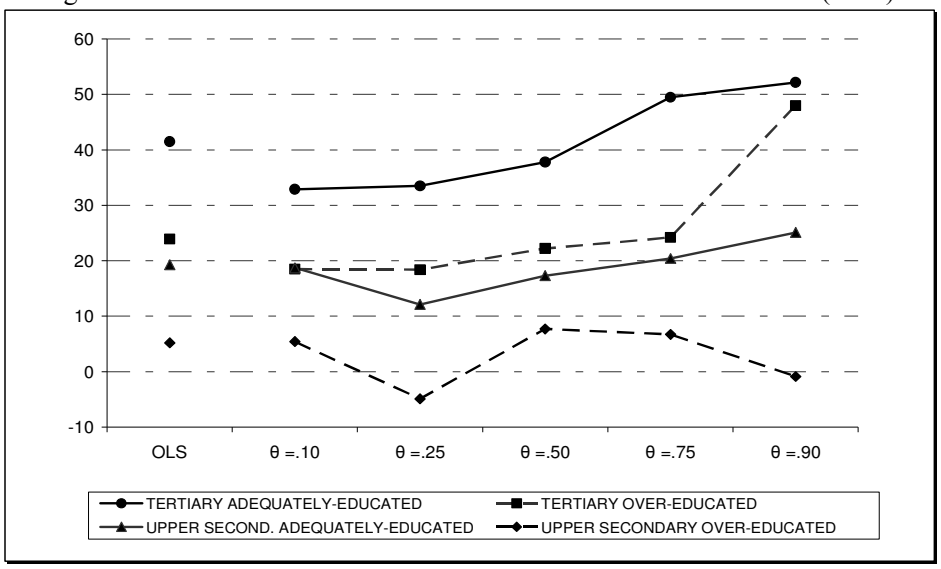


Figure 4 – Conditional returns to education and over-education – Women (2001)

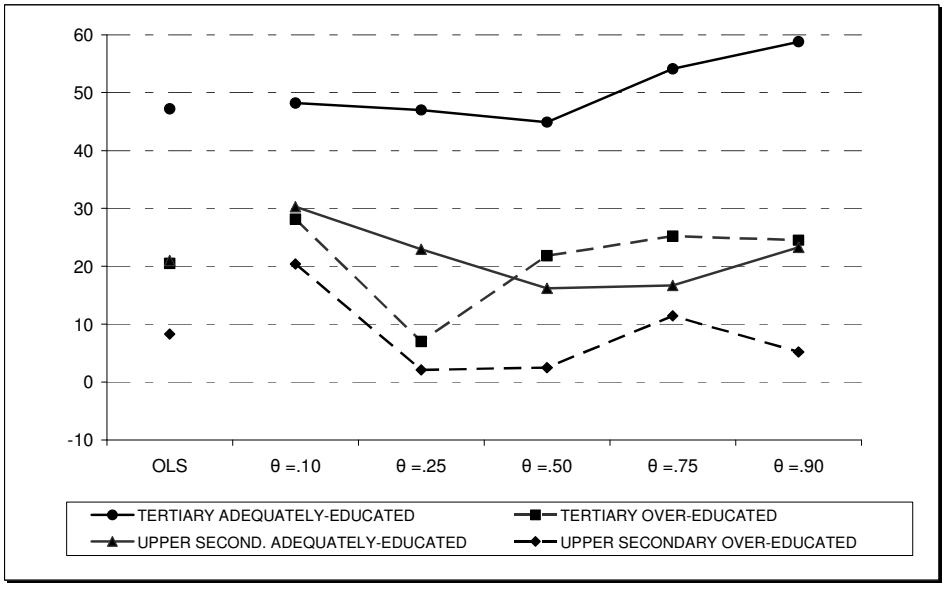


Figure 5 – Changes in conditional returns to education – Men (1994-2001)

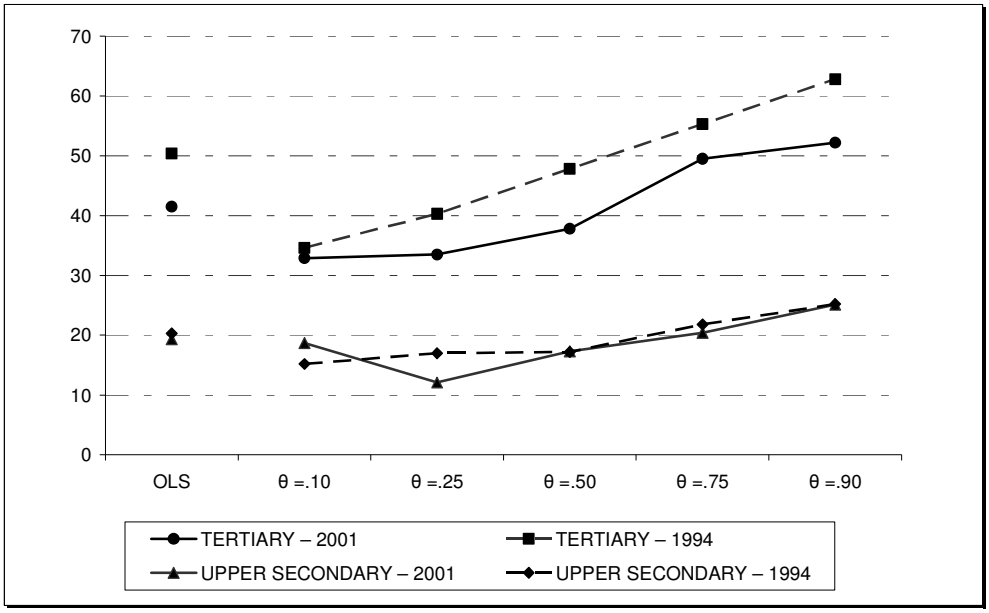


Figure 6 – Changes in conditional returns to education – Women (1994-2001)

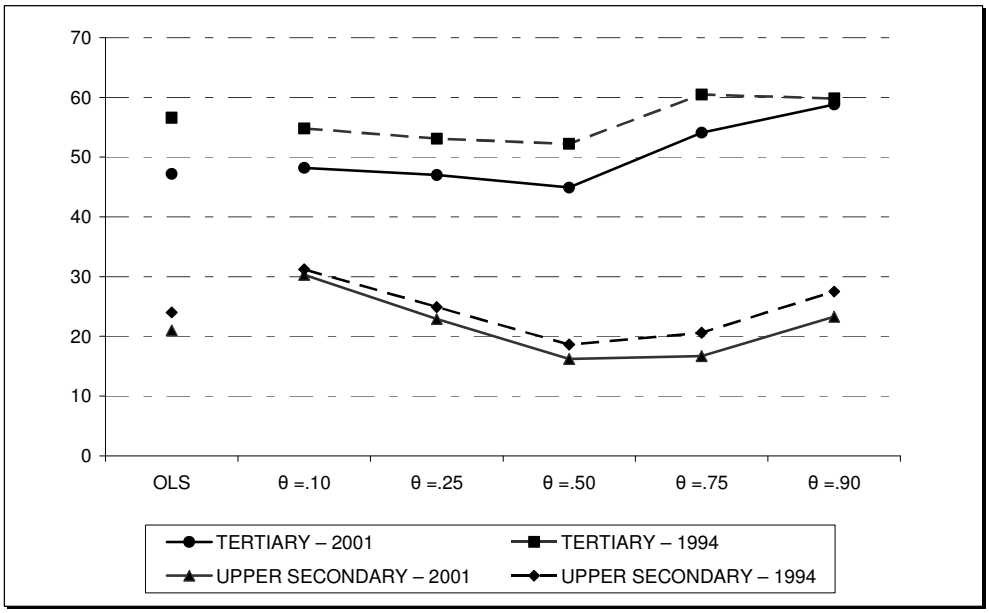


Figure 7 – Changes in conditional returns to over-education – Men (1994-2001)

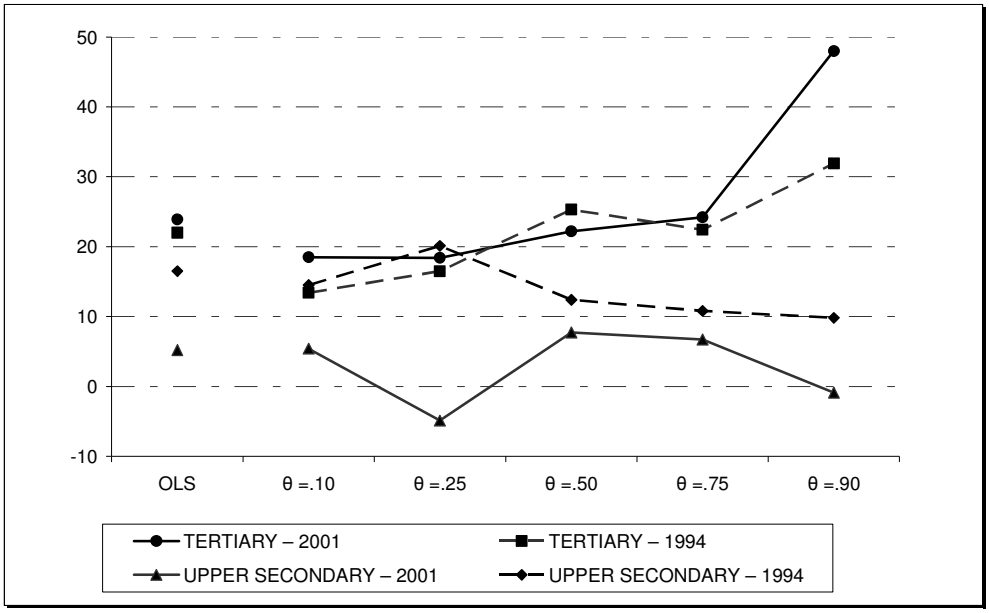
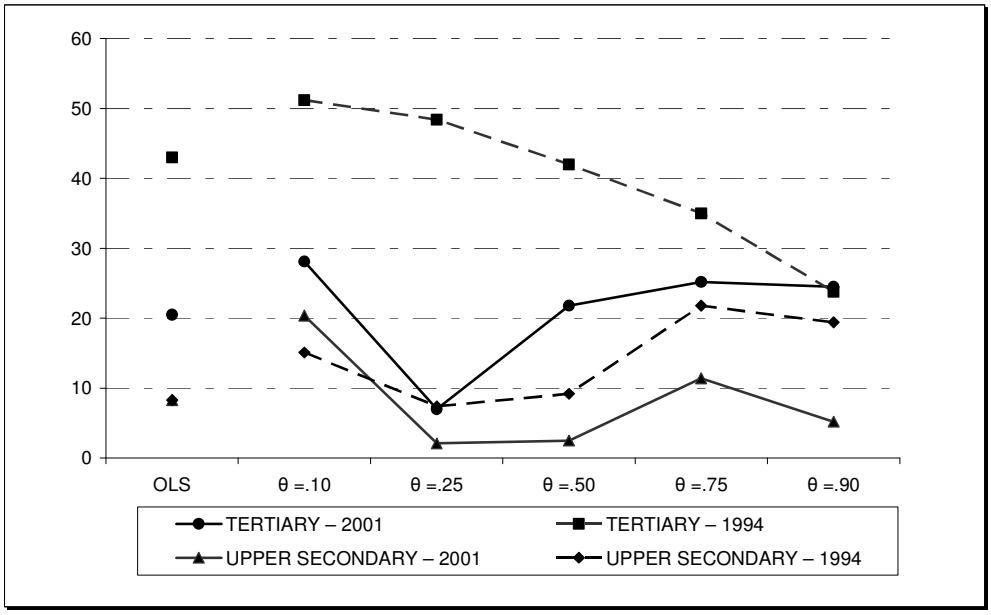


Figure 8 – Changes in conditional returns to over-education – Women (1994-2001)



Additional Tables

Table 1B. OLS and Quantile Regression (2001) – Dependent variable: Ln. Gross hourly wage

	Men					
	OLS	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
Tertiary	0.388***	0.296***	0.307***	0.336***	0.464***	0.530***
	(0.030)	(0.069)	(0.035)	(0.034)	(0.038)	(0.042)
Upper Secondary	0.147***	0.105***	0.093***	0.152***	0.182***	0.162***
	(0.028)	(0.057)	(0.038)	(0.028)	(0.032)	(0.041)
Experience*100	0.038***	0.048***	0.032***	0.026***	0.029***	0.033***
	(0.004)	(0.008)	(0.006)	(0.004)	(0.004)	(0.005)
(Experience*100)^2	-0.010***	-0.008***	-0.005***	-0.004***	-0.004***	-0.004***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Tenure: 5-14 years	0.100***	0.220***	0.088***	0.062***	0.043*	-0.006***
	(0.023)	(0.051)	(0.031)	(0.026)	(0.026)	(0.047)
Tenure: ≥ 15 years	0.189***	0.306***	0.131***	0.150***	0.122**	0.279
	(0.037)	(0.058)	(0.036)	(0.033)	(0.068)	(0.071)
Married	0.089***	0.170***	0.087***	0.055**	0.055**	0.069*
	(0.027)	(0.064)	(0.035)	(0.029)	(0.029)	(0.035)
Immigrant	0.021	-0.155	0.012	0.066	-0.060	0.179
	(0.110)	(0.200)	(0.201)	(0.092)	(0.147)	(0.241)
Industry	0.001	0.128**	0.079***	0.040	-0.060***	-0.124***
	(0.023)	(0.051)	(0.029)	(0.026)	(0.027)	(0.031)
Firm size: 20-99 employees	0.094***	0.061	0.087***	0.098***	0.108***	0.076***
	(0.025)	(0.055)	(0.032)	(0.024)	(0.030)	(0.031)
Firm size: 100-499 employees	0.221***	0.175**	0.217***	0.256***	0.242***	0.268***
	(0.033)	(0.086)	(0.043)	(0.033)	(0.038)	(0.050)
Firm size: ≥ 500 employees	0.309***	0.328***	0.278***	0.320***	0.330***	0.209***
	(0.041)	(0.069)	(0.048)	(0.043)	(0.053)	(0.067)
Constant	6.032***	5.279***	5.881***	6.181***	6.426***	6.661***
	(0.041)	(0.081)	(0.064)	(0.040)	(0.039)	(0.043)

Note: i) * signals significant at the 10% level, ** signals significant at the 5% level, and *** signals significant at the 1% level; ii) standard errors in parenthesis; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors are obtained using 500 replications. The reference individual is a worker with less than upper secondary education, less than 5 years of tenure, single, not immigrant, working in the service sector in a firm with less than 20 employees

Table 2B. OLS and Quantile Regression (2001) – Dependent variable: Ln. Gross hourly wage

	Women					
	OLS	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
Tertiary	0.430***	0.375***	0.430***	0.414***	0.479***	0.568***
	(0.043)	(0.112)	(0.064)	(0.052)	(0.045)	(0.060)
Upper Secondary	0.174***	0.239***	0.172***	0.154***	0.150***	0.198***
	(0.041)	(0.106)	(0.058)	(0.046)	(0.044)	(0.061)
Experience*100	0.045***	0.060***	0.052***	0.040***	0.034***	0.039***
	(0.006)	(0.011)	(0.008)	(0.008)	(0.006)	(0.008)
(Experience*100)^2	-0.008***	-0.011***	-0.010***	-0.007***	-0.006***	-0.007***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Tenure: 5-14 years	0.252***	0.469***	0.293***	0.216***	0.127***	0.072
	(0.035)	(0.085)	(0.047)	(0.045)	(0.040)	(0.058)
Tenure: ≥ 15 years	0.326***	0.335***	0.246***	0.292***	0.318***	0.164
	(0.072)	(0.133)	(0.085)	(0.141v	(0.086)	(0.144)
Married	0.023	0.138*	0.061	-0.020	-0.013	0.006
	(0.035)	(0.081)	(0.047)	(0.045)	(0.033)	(0.049)
Immigrant	-0.188*	-0.425***	-0.108	-0.195	-0.266	-0.011
	(0.127)	(0.231)	(0.212)	(0.156)	(0.189)	(0.205)
Industry	0.009	0.112	0.064	0.022	-0.049	-0.082
	(0.037)	(0.105)	(0.056)	(0.043)	(0.042)	(0.056)
Firm size: 20-99 employees	0.131***	0.125*	0.148***	0.117***	0.128***	0.058
	(0.041)	(0.082)	(0.052)	(0.047)	(0.048)	(0.060)
Firm size: 100-499 employees	0.207***	0.101	0.225***	0.217***	0.205***	0.155**
	(0.052)	(0.111)	(0.075)	(0.064)	(0.052)	(0.078)
Firm size: ≥ 500 employees	0.268***	0.320***	0.296***	0.277***	0.271***	0.212***
	(0.063)	(0.125)	(0.073)	(0.083)	(0.075)	(0.082)
Constant	5.767***	4.877***	5.420***	5.898***	6.228***	6.446***
	(0.056)	(0.123)	(0.109)	(0.071)	(0.054)	(0.059)

Note: i) * signals significant at the 10% level, ** signals significant at the 5% level, and *** signals significant at the 1% level; ii) standard errors in parenthesis; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors are obtained using 500 replications. The reference individual is a worker with less than upper secondary education, less than 5 years of tenure, single, not immigrant, working in the service sector in a firm with less than 20 employees

Table 3B. Conditional returns to over-education by education levels (1994)

MEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY ADEQUATELY-EDUCATED	**0.504*****	0.346***	0.403***	0.478***	0.553***	0.628***
	(0.032)	(0.068)	(0.041)	(0.034)	(0.039)	(0.054)
TERTIARY OVER-EDUCATED	**0.220*****	0.134	0.165	0.253***	0.224***	0.319***
	(0.070)	(0.163)	(0.112)	(0.084)	(0.076)	(0.129)
UPPER SECOND. ADEQUATELY EDUCATED	**0.203*****	0.152***	0.170***	0.172***	0.218***	0.252***
	(0.033)	(0.069)	(0.035)	(0.037)	(0.035)	(0.068)
UPPER SECONDARY OVER-EDUCATED	**0.165*****	0.145	0.201***	0.122**	0.108**	0.098
	(0.041)	(0.095)	(0.050)	(0.042)	(0.060)	(0.075)
WOMEN	**OLS**	θ =.10	θ =.25	θ =.50	θ =.75	θ =.90
TERTIARY ADEQUATELY-EDUCATED	**0.566*****	0.548***	0.531***	0.522***	0.605***	0.597***
	(0.046)	(0.127)	(0.054)	(0.051)	(0.052)	(0.060)
TERTIARY OVER-EDUCATED	**0.430*****	0.512***	0.484***	0.420***	0.350***	0.238**
	(0.082)	(0.164)	(0.153)	(0.100)	(0.076)	(0.122)
UPPER SECOND. ADEQUATELY EDUCATED	**0.240*****	0.312***	0.249***	0.186***	0.206***	0.275***
	(0.046)	(0.116)	(0.064)	(0.050)	(0.056)	(0.080)
UPPER SECONDARY OVER-EDUCATED	**0.083**	0.151	0.074	0.092	0.218**	0.194**
	(0.094)	(0.337)	(0.152)	(0.111)	(0.115)	(0.093)

Note: i) * signals significant at the 10% level, ** signals significant at the 5% level, and *** signals significant at the 1% level; ii) standard errors in parenthesis; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors are obtained using 500 replications. The reference individual is a worker with less than upper secondary education, less than 5 years of tenure, single, not immigrant, working in the service sector in a firm with less than 20 employees