



Munich Personal RePEc Archive

Are educational mismatches responsible for the ‘inequality increasing effect’ of education?

Budria, Santiago

University of Madeira ., CEEAplA

1 June 2010

Online at <https://mpra.ub.uni-muenchen.de/23420/>

MPRA Paper No. 23420, posted 22 Jun 2010 16:02 UTC

Are educational mismatches responsible for the 'inequality increasing effect' of education?

Santiago Budría* – University of Madeira and CEEApIA

March 2010

Abstract

This paper asks whether educational mismatches can account for the positive association between education and wage inequality found in the data. We use two different data sources, the European Community Household Panel and the Portuguese Labour Force Survey, and consider several types of mismatch, including overqualification, underqualification and skills mismatch. We test our hypothesis using two different measurement methods, the 'statistical' and the 'subjective' approach. The results are robust to the different choices and unambiguously show that the positive effect of education on wage inequality is not due to the prevalence of educational mismatches in the labour market.

Keywords: Overeducation, returns to education, educational mismatch, within-groups wage inequality

* I thank Ana Moro-Egido, Pedro Telhado-Pereira, Corrado Andini and Carlospaki Korre for their comments in the earlier stages of the paper. The financial support from the Spanish Ministry of Education through grant SEJ2006-11067 and the Junta de Andalucía through grant P07-SEJ-03261 is gratefully acknowledged. Address correspondence to: Santiago Budría, Department of Economics, University of Madeira, Rua Penteadá 9000-390, Funchal, Portugal. Phone: +351-291 705 055. Fax: +351-291 705 040. E-mail: sbudria@uma.pt.

“What matters, then, isn't what you do or where you live, but what you know. When two-thirds of all new jobs require a higher education or advanced training, knowledge is the most valuable skill you can sell. It's not only a pathway to opportunity, but it's a prerequisite for opportunity”

Extract from President Obama's education speech in Ohio, September 9, 2008

1. Introduction

Education plays a fundamental role in providing the basis for economic growth, social cohesion and personal development in modern societies. Better educated people tend to enjoy better health, exhibit pro-social behaviour, engage more in political and civic participation, raise more educated children, and are less likely to participate (actively or passively) in crimes. From an economic standpoint, it is widely recognized that knowledge and human capital play an increasingly central role in the economic success of nations and individuals. Better educated people are more productive and innovative, more likely to be economically active, earn higher wages and experience higher wage growth over their working lives¹. Consequently, the design of effective and efficient educational systems is currently a major policy issue in wealthy and poor societies around the globe.

There is, however, one aspect of education that rings some alarm bells among social scientists in general and economists in particular. 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. Still, recent international research has shown by means of quantile regression analysis that wage inequality is higher among more educated individuals (Buchinsky, 1994, Pereira and Martins, 2002, Fersterer and Winter-Ebmer, 2003, Martins and Pereira, 2004, Machado and Mata, 2005). To use an economist's term, this observation is *other things equal* (i.e., conditional on controlling for a wide range of labour market characteristics that may also affect the earnings distribution of the different education groups, such as professional experience, occupation, sector and gender) and, consequently, has

¹ For evidence on the social and economic benefits of education see, for example, Ashenfelter and Rouse (2000), Wolfe and Haveman (2001) and Dolton et al. (2009).

been termed the ‘inequality increasing effect’ of education: if (conditional) wage dispersion is higher for more educated individuals, then an educational expansion may add to overall wage inequality.

This finding raises important policy implications. First, it suggests that the inequality-reducing scope typically attributed to education is thornier than previously thought. Second, individuals consciously invest in themselves to improve their own, personal economic returns. A person may study architecture because she likes designing, but also because architects earn more than regular people. However, the higher inequality found among the educated warns that, in doing so, she exposes herself to greater wage uncertainty. Previous studies have shown that this uncertainty can be substantial (Cunha et al., 2005, Cunha and Heckman, 2007) and that it may exert a large influence on the decision on extended schooling (Carneiro et al., 2003, Hartog and Serrano, 2007, Hogan and Walker, 2007) and on the wage distribution in the society (Hartog et al., 2003, Bonin et al., 2007, Hartog and Vijverberg, 2007).

Despite these policy concerns, up to date the causes of the positive association between education and conditional wage dispersion remain mostly unknown. Among the plausible explanations, overeducation is the one that has been suggested more frequently. Machin (1996), Green *et al.* (1999), Fersterer and Winter-Ebmer (2003) and Martins and Pereira (2004), among others, put forward the notion that overeducation may be responsible for the higher wage dispersion among the educated or, alternatively, for the widening wage dispersion experienced by this group over the last years. As Martins and Pereira (2004, p. 365) put it, ‘situations where highly-schooled workers take jobs with a low skill requirement and consequent low pay would be consistent with these results’. The empirical evidence supports this intuition. First, it is well documented that a significant proportion of the labour force works in jobs that are not commensurate with their qualifications (Groot and Maassen van den Brink, 2000). Second, the survey of empirical studies shows that overeducated workers earn less than workers who have the same education but hold jobs for which they are adequately educated. The estimated wage gap can be as large as 8% in Kiker et al. (1997) for Portugal, 13% in Verdugo and Verdugo (1989) and 11% in Cohn and Kahn (1995) for the US, 26% in Groot (1996) for Holland and 35% in Dolton and Silles (2008) for the UK. When differentiating between the years required to match the educational requirement of the job and the years that exceed the educational level needed at the job, researchers usually find that excess education gives a 50% lower return than the return to required education (Hartog, 2000, McGuinness, 2006). Therefore, we expect that,

relative to the adequately educated, overeducated workers are located at lower deciles of the earnings distribution and earn a lower return from their educational investment. The incidence of overeducation would then act as a mechanism enhancing wage dispersion within similarly educated individuals.

Although frequently suggested, this hypothesis has not been tested to date. In this paper we take a step towards filling this gap by asking: can overeducation account for the 'inequality increasing effect' of education? Answering this question is compelling, as educational mismatches are receiving a lot of attention as a potential source of the recent increase in total within-groups-inequality observed in developed countries. We take Portugal as case study, for in this country the inequality increasing effect of education has been found to be particularly acute (Martins and Pereira, 2004).

Overeducation describes the extent to which an individual possesses a level of education in excess of that which is required for his job. A first methodological point concerns the measurement of this excess education. Two approaches can be distinguished in the literature. The 'subjective' approach is based on the worker's self-assessment regarding the quality of the match between his education and the educational requirements of the job (e.g., Halaby, 1994). A variation of this method asks workers what the minimum educational requirements are for the job, and then compares this report with the actual education level of the worker (e.g., Duncan and Hoffman, 1981, Cohn and Khan, 1995, Daly et al., 2000, Dolton and Vignoles, 2000). The second approach consists on finding out the educational requirements externally. A worker is regarded to be overeducated (undereducated) if he has more (less) education than is required for the job. A number of studies establish the job level requirements by calculating the mean/mode value within occupations (e.g., Verdugo and Verdugo, 1989, Kiker et al., 1997, Bauer, 2002). This method is frequently termed the 'statistical' approach. Some other authors prefer to rest upon the formal evaluation of independent job analysts to determine the correct level of education required for each job title (e.g., Decker et al., 2002). All these methods present their advantages and limitations. For the purpose of the present paper, it suffices to recall that none of the above methods outperforms the others, and that the extent and effects of overeducation may differ across measures (Battu et al., 2000, Rubb, 2004, Verhaest and Omey, 2006, 2009). Similarly, certain personal characteristics appear to be more or less statistically related to overeducation depending on the measurement method.

On the basis of this ground, we test our central hypothesis using the two alternative measurement methods that have gained currency in the overeducation literature, the ‘subjective’ and the ‘statistical’ approach. This refinement is based on the utilization of two different datasets, the Portuguese Labour Force Survey and the European Community Household Panel, and allows us to better assess the robustness of our findings. A second methodological feature is that we consider different definitions of mismatch, not just overqualification. Specifically, we differentiate between ‘overqualification’, ‘underqualification’ and ‘skills mismatch’, and define yet another category, ‘strong mismatch’, to refer to those workers who are overqualified and, at the same time, lack necessary skills. Although most studies primarily focus on overqualification, there is no presumption that the effects of other forms of mismatch are less relevant. Moreover, a review of the existing literature suggests that differences in the amount, not just the incidence, of mismatch should be taken into consideration (Hartog, 2000). To that purpose, we use a statistical approach that explicitly differentiates between levels of over- and under-qualification.

Finally, most of the debate in the policy arena has gravitated around the question of to what extent the incidence of educational mismatches entail a productivity loss. On the one hand, mismatched workers may be in some way less able and lack some of the abilities and skills required to do a job commensurate with their education. In this case, the mismatch pay penalty would be a spurious statistical finding reflecting an omitted variables problem rather than a real economic problem. On the other hand, the mismatch phenomenon may reflect a real missadjustment between the worker’s potential and the job’s productivity ceiling. The available evidence is rather mixed (Green et al., 1999, Bauer, 2002, McGuinness, 2003). The results in this paper are based on quantile regression and provide valuable insights on this debate. Specifically, we investigate whether the wage effects of mismatch are homogeneous across segments (quantiles) of the earnings distribution. Assuming that earnings capacity is given by the individuals’ unobserved ability, the estimates at different quantiles provide snap-shots of how mismatched individuals within different ability groups are impacted relative to their well-matched counterparts. The major advantage of this approach is that it prevents us from comparing higher ability matched individuals with lower ability mismatched individuals, thus eliminating the potential bias arising from unobserved heterogeneity. We use the results to investigate whether educational mismatches entail a real productivity loss or, on the contrary, are the result of the low ability levels possessed by certain individuals.

The next section establishes the paper's research hypotheses. Section 3 presents the datasets and the definitions of mismatch used in the analysis. Section 4 outlines the quantile regression framework and introduces the Subjective and the Statistical model. Section 5 calculates quantile returns to education and inspects whether educational mismatches can account for the dispersion in the returns across quantiles. Besides, the wage effects of overqualification, underqualification, skills mismatch and strong mismatch at different points of the conditional wage distribution are documented. Section 6 discusses the results and presents concluding remarks. Appendix A contains the description of the variables used in the regressions. For a sensitivity analysis, Appendix B contains the estimates when a restricted rather than a full set of controls is used in the earnings equations.

2. Research hypotheses

We inherit the tradition of labour economists of estimating a set of wage equations in which individual earnings are explained in terms of a wide range of individual demographic and labour market characteristics. Among these characteristics, we include the crux of our analysis: education and mismatch status. The equations are used to calculate returns to schooling at different segments of the earnings distribution. Differences in these returns represent residual inequalities of pay that can be attributed to education (Buchinsky, 1994). Our first research question is whether the resulting inequalities are increasing as we move towards more educated groups. More specifically,

- *Question 1:* Is there an inequality increasing effect of education?

This hypothesis is the least original, and has been already covered in existing work. The following question is more innovative. Specifically, we hypothesize that discriminating between matched and mismatched workers should remove part of the observed dispersion within education levels. Specifically, our second research question is:

- *Question 2:* Is the inequality increasing effect of education due to the prevalence of educational mismatches in the labour market?

This question, which is central in the paper, can be potentially broken up into two separate questions if we consider the various definitions and measurements of mismatch outlined above.

Specifically,

- *Question 2a:* Is the inequality increasing effect of education due to the prevalence of various types of educational mismatches in the labour market?
- *Question 2b:* Is the inequality increasing effect of education due to individual differences in the degree of educational mismatch?

Finally, our third research question is:

- *Question 3:* Is the pay penalty of educational mismatch homogenous across individuals in different segments of the earnings distribution?

3. Data and measurement of mismatch

We use information from two sources: the European Community Household Panel (ECHP) and the Portuguese Labour Force Survey (PLFS). The ECHP is a representative survey that covers 15 European countries. It contains personal and labour market characteristics, including wage, education, hours worked, tenure, experience, sector, firm size, marital status and immigrant condition, among other variables. For the present study, we use pooled data from 1994-2001 and the Portuguese subsample of the dataset. The PLFS is a quarterly survey of a representative sample of households in Portugal. Its sample size is about 45,000 individuals, and it has a rotating structure in which 1/6 of the sample is dropped randomly in each quarter. We use pooled data from 2000 to 2002. The variables included in the PLFS are very similar to those included in the ECHP. An advantage of the PLFS over the ECHP is that it describes more accurately the educational attainment of respondents. Specifically, the PLFS includes ten categories that range from 'No studies' to 'Doctoral degree', each of them associated with a certain number of (minimum) years of schooling. In the ECHP, in turn, the educational variable is coded in three broad categories (less than upper secondary, upper secondary and tertiary education) based on the ISCED-97 classification (OECD, 2003). An advantage of the ECHP over the PLFS is that it includes two self-assessed measures of the quality of the match between the worker's education and the requirements of the job.

In the two datasets, we restrict the sample to full-time, private sector, male wage earners aged

between 21 and 55, who work normally between 15 and 70 hours a week, and are not employed in the agricultural sector. The case of women is disregarded on account of the extra complication of potential selectivity bias. Workers with a monthly wage rate that is less than 10% or over 10 times the average wage have been also excluded. These restrictions leave us with a final sample of 8,319 individuals in the ECHP and 11,947 in the PLFS.

3.1 The subjective approach (ECHP)

The ECHP includes two questions with the worker's self-assessment regarding the quality of the match between acquired education and the requirements of the job. These questions have already been used by Alba-Ramírez and Blázquez (2002), Wasmer et al. (2007) and Budría and Moro-Egido (2008). The first self-evaluation question is

- **(S1)** *Do you feel that you have skills or qualifications to do a more demanding job than the one you have now?*

This information is used to identify those workers with excess education (S1: 'yes'). The second question is

- **(S2)** *Have you had formal training or education that has given you skills needed for your present type of work?*

This information allows us to identify those workers who did not acquire necessary skills through training and education (S2: 'no')². Using S1 and S2 we can construct the following categories

- 1) Workers with excess education but with appropriate skills (S1: 'yes', S2: 'yes'). We will term them as, simply, the 'overqualified'.

²Through the paper we abuse language somewhat and will refer to these workers as workers who 'lack necessary skills'. We are aware, however, that there might be individuals who have not had formal education and training for unskilled jobs but who have acquired the necessary background through other sources, including peer observation, learning by doing and general work experience. Although these channels are typically less relevant, they might be important for a small fraction of uneducated individuals working in low level jobs. As most other measures of mismatch, a limitation of our definition is that it focuses on formal education and training and disregards other sources of skills acquisition.

- 2) Non-overqualified workers who did not acquire necessary skills (S1: 'no', S2: 'no'). We will refer to them as the 'skills mismatched'.
- 3) Workers who, despite having excess education, lack necessary skills (S1: 'yes', S2: 'no'). We will term these workers as the 'strongly mismatched'.
- 4) Workers with appropriate education and skills (S1: 'no', S2: 'yes'). This is our reference group, composed by 'matched workers'.

As an illustration of the different types of mismatch, consider an individual with a bachelor degree in marketing employed as

- i) a salesman. In this situation, he may feel that his university degree allows him to do a more demanding job, even though it helps him to perform his current job. Thus, he would be considered an 'overqualified' worker.
- ii) a computer engineer. In this case, he may feel that his formal education does not allow him to perform a more demanding job nor has provided him with the skills needed to perform the job. Thus, he would be labelled as 'skills mismatched'.
- iii) a gardener. In this case, the individual presumably will report that his university degree allows him to do a more demanding job, yet it has not provided him with the skills needed to be a gardener. Thus, he will be considered as a 'strongly mismatched' worker.

Before advancing, two remarks are in order. First, most measures of mismatch used in the literature are exclusively based on the level of education attained by the individual. However, workers who declare not to be overqualified may have an inappropriate job match when the content, not the level, of their education is evaluated. As Sloane (2002, p.7) puts it 'in some measures of overeducation 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'. Exploring the effects of having inappropriate qualifications seems compelling, as there is no presumption that these are less important than the effects of having excess qualifications. This is why we complement the information reported in S1 with that reported in S2. Second, workers with excess education are typically regarded as 'overeducated' in the literature. However, as is clear from the previous examples, the term 'overeducation' may be seriously misleading. Workers who have excess education and, additionally, are mismatched in terms of skills can be hardly

labelled as ‘overeducated’, as their formal education did not provide them with the necessary background. This is why we split the group of workers with excess education (S1: ‘yes’) into those who are simply ‘overqualified’ (S2: ‘yes’) and those who are ‘strongly mismatched’ (S2: ‘no’).

In Table 1 we report the incidence of the different types of mismatch together with summary statistics for each group. The variables listed in the table are described in Appendix A. The proportion of overqualified, skills mismatched and strongly mismatched workers is, respectively, 12.6%, 44.6% and 32.4%. These figures indicate that most workers (77.0%) lack some skills that are required in their jobs. It is worth mentioning that approximately seven out of ten of the workers with excess qualifications (S1: ‘yes’) are strongly mismatched (S2: ‘no’).

The large proportion of mismatched workers in our data should not come as a surprise, as subjective measures of mismatch tend to render large estimates. Indeed, our figures can be directly compared to those reported in Wasmer et al. (2007), who use the same dataset and taxonomy of mismatch to provide a European perspective on the topic. They report that in Europe as a whole, the incidence of overqualification, skills mismatch and strong mismatch is 33.0%, 24.7% and 21.1%, respectively. In Portugal, therefore, the extent of mere overqualification is low and the extent of skills mismatch is high by European standards. This outcome is consistent with the low educational attainment and training participation of the Portuguese labour force³.

Some interesting differences across groups emerge in Table 1. The overqualified earn the highest wages (1.25), are more likely to have tertiary education (22%), employer-financed training (32%), less experience (12.2 years), work in the services sector (58%), in larger firms (28% work in a firm with 100 workers or more), and are less likely to report bad health (1%). As opposite, workers who are skills mismatched or strongly mismatched earn lower wages (0.89 and 0.91, respectively), are less likely to have university education (1%), training (3% and 5%, respectively), work in larger firms (less than 15% work in a firm with 100 workers or more), and are more likely to report bad health (6% and 3%). The overqualified tend to work in white-collar occupations (‘Professionals’ and ‘Technicians and associate professionals’), while the

³ In Portugal, only 27.6% of the adult population (25-64 years old) has completed upper secondary education, while in Europe as a whole (EU-25) this proportion rises to 69.7%. Similarly, training participation in Portugal is 3.8%, against 10.1% in EU-25 (Eurostat, 2007).

skills and the strongly mismatched are more likely to be blue-collar workers ('Craft and related trades workers', 'Plant and machine operators and assemblers' and 'Elementary occupations'). Finally, differences across groups in terms of tenure, unemployment experience and marital status are relatively small.

In Table 2 we examine more closely the connection between mismatch status and education level. One might expect that the high educated are more likely to be overqualified and less likely to lack necessary skills, and this is what is observed. The proportion of overqualified workers is increasing in the education level, from 6.9% in the group with less than upper secondary education to 58.5% in the group with a tertiary education. As opposite, the incidence of skills mismatch is higher among the less educated, ranging from 2.8% in the tertiary-level group up to 51.1% in the less educated group. Finally, the proportion of matched workers is increasing in education, ranging from 8.1% (less than upper secondary) to 32.2% (tertiary). In other words, the self-reported variables seem to be behaving reasonably.

3.2 The statistical approach (PLFS)

Verdugo and Verdugo (1989) defined required schooling as a one standard-deviation range around the mean level of schooling within an occupation. Workers are considered to be adequately educated, overqualified or underqualified depending on whether their actual education falls within, above or below this range, respectively. For the present study, however, we use the modal value rather than the mean value. As Kiker et al. (1997) point out, this choice reduces the sensitivity to outliers and changes in workplace organization⁴.

With this method, actual years of schooling of individual i working in occupation j , (S_{ij}^a), can be decomposed into required years of schooling in occupation j , (S_j^r), years of overschooling, (S_{ij}^o), and years of underschooling (S_{ij}^u), where S_j^r is the modal value within the occupation and

$$S_{ij}^o = \begin{cases} S_{ij}^a - S_j^r & \text{if } S_{ij}^a - S_j^r > 0 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad S_{ij}^u = \begin{cases} S_j^r - S_{ij}^a & \text{if } S_j^r - S_{ij}^a > 0 \\ 0 & \text{otherwise} \end{cases}$$

⁴ Using the mean rather than the modal value produced only small changes in the estimates. The results are available upon request.

Two data features are desirable when implementing the statistical approach. First, the educational attainment of individuals should be sufficiently detailed. Otherwise, the modal level of education within an occupation may result into a broad education category that pools together workers whose education level is not comparable. Second, the occupation variable should be sufficiently disaggregated. Otherwise, we may be pooling together jobs with very different educational requirements. The PLFS exhibits these two ingredients. The educational classification is coded in ten categories that result into nine different levels of years schooling, and occupations are disaggregated on a 2-digit level based on the National Classification of Occupations (CNP). For each occupation, we compute the modal value of years of schooling and then calculate the corresponding years of over-, under- and required schooling for each individual within the occupation. Occupations with less than 10 observations were excluded from the analysis⁵.

A concern with our data is that as education levels are very low in Portugal, the modal level of education might be similarly low in almost all occupations. If this is the case, our measure of required education could be criticized for not being subtle enough to capture the expected variations across occupations. To allay this concern, in Figure 1 we report the frequency, as measured by the number of occupations, of each modal level of schooling. As expected, the modal value is low in most occupations, with 4 years of schooling being the most frequent outcome (12). Still, we detect some variation in required schooling across occupations. Specifically, almost 42% of the occupations (11 out of 26) are associated with 9 or more years of schooling, and in five occupations the modal value of schooling amounts to 16 years. Such dispersion seems substantial, as the sample mean of actual years of schooling is as low as 6.22 in the data.

In Table 3 we report the summary statistics. With the statistical method the proportion of overqualified and underqualified workers is 36.8% and 18.7%, respectively. Overqualified workers have on average 7.70 years of schooling, 3.35 of which are in excess to those required in the job. Underqualified workers, in turn, are less educated than average (4.98 years) and have 4.35 years of schooling less than what is required in their occupations. All groups work a similar number of hours, but the overqualified earn 11% less than the underqualified (1.01 against 1.12). The overqualified have less experience than average (14.7 against 24.3 years), less tenure (50% are below five years), tend to be single (43%) and are likely to be 'Skilled

⁵ This exclusion restriction affected 3 occupations and 0.1% of the workers in the initial sample.

workers, craftsmen and similar' (48%). The underqualified in turn, are more concentrated in the 'Middle management and technicians' and 'Administrative and related workers' occupations (20% and 32%, respectively).

3.3 Are the subjective and statistical indicators comparable?

Comparisons between methods must be undertaken very carefully, insofar as in most cases they measure different things. Thus, for example, 36.8% of the workers in the PLFS sample have excess education, while in the ECHP only 12.6% is regarded as being 'overqualified'. The difference is potentially intriguing. It may also be intriguing the fact that the 'overqualified' in the ECHP earn higher wages than their well-matched counterparts (1.25 versus 1.19), while the opposite occurs in the PLFS (1.01 versus 1.06). However, it should not be so if we recall that the group of 'overqualified' individuals in the ECHP was restricted to include only those who have the necessary skills for their jobs. Individuals with excess education but with insufficient skills are the 'strongly mismatched', who represent 32.5% of the total population and earn very low wages (0.91). Aggregating, we find that in the ECHP the total fraction of workers with excess education amounts to 45.0% and earns an average wage of 1.00. These figures come closer to the corresponding estimates in the PLFS (36.8% and 1.01, respectively).

Similarly, 44.6% of the ECHP sample individuals report that they do not have excess qualifications and that they lack necessary skills. Presumably, an important fraction of these workers has less education than required. However, the incidence of underqualification is as low as 18.7% in the PLFS. Again, the difference is less intriguing if we consider that the ECHP 'skills mismatch' group may include two types of workers. On the one hand, the underqualified, i.e., those who did not acquire necessary skills because they did not acquire a sufficiently high level of education. On the other hand, those with 'wrong' qualifications, i.e., those who completed a high level of education but work in jobs that are not related to the content of their education. Therefore, the ECHP measure provides a more global indicator of the different skills shortages that take place in the labour market.

4. Estimating models

The quantile regression model can be written as

$$\ln w_i = X_i \beta_\theta + e_{\theta i} \quad \text{with } \text{Quant}_\theta(\ln w_i | X_i) = X_i \beta_\theta \quad (1)$$

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 θ th conditional quantile of $\ln w$ given X . The θ th regression quantile, $0 < \theta < 1$, is defined as a solution to the problem

$$\text{Min}_{\beta \in R^k} \left\{ \sum_{i: \ln w_i \geq X_i \beta_\theta} \theta |\ln w_i - X_i \beta_\theta| + \sum_{i: \ln w_i < X_i \beta_\theta} (1 - \theta) |\ln w_i - X_i \beta_\theta| \right\} \quad (2)$$

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

$$\text{Min}_{\beta \in R^k} \left\{ \sum_i \rho_\theta(\ln w_i - X_i \beta_\theta) \right\} \quad (3)$$

This problem is solved using linear programming methods, while standard errors for the vector of coefficients are obtained using bootstrap techniques (Buchinsky, 1998).

4.1 The Subjective model

We use the following earnings equation for the ECHP data,

$$\ln w_i = \alpha_\theta + \delta_\theta X_i + \beta_\theta E_i + \gamma_\theta M_i + e_{\theta i} \quad (4)$$

where $\ln w_i$ is the logarithm of the net hourly wage, X_i is a vector of controls, E_i includes two education dummies, one for upper secondary education and one for tertiary education, and M_i includes three dummy variables controlling for, respectively, overqualification, skills mismatch and strong mismatch. This specification is similar to that used in Verdugo and Verdugo (1989), who were the first to use a categorical variable to measure the effects of overeducation on wages. Later on, this equation was adopted by Dolton and Vignoles (2000) and Bauer (2002), among others.

4.2 The Statistical model

For the PLFS, we use the following earnings equation

$$\ln w_i = \alpha_\theta + \delta_\theta X_i + \beta_{u\theta} S_i^u + \beta_{r\theta} S_i^r + \beta_{o\theta} S_i^o + e_{\theta i} \quad (5)$$

where β_{u0} is the return to a year of schooling below the schooling requirement, β_{r0} is the return to years of required education, and β_{e0} is the return to an additional year of schooling beyond those required. This specification, termed the ORU model in the literature, was first introduced by Duncan and Hoffman (1981) and since then has inspired many other studies (e.g., Cohn and Kahn, 1995, Kiker et al., 1997). In words of Hartog (2000, p. 144), ‘the ORU earnings function has proven itself as an extension of the canonical Mincerian earnings function, by passing statistical testing in several countries for several datasets and several years’. An advantage of the statistical model over the subjective is that it controls for the amount, not just for the incidence, of mismatch.

5. Results

In this section we calculate returns to education and examine how their dispersion across quantiles changes when explicit controls for mismatch are included in the earnings equation. In a second stage, we document how the effect of mismatch on wages differs across segments of the wage distribution. Through the discussion we concentrate on the education and mismatch variables and disregard the effects of other covariates on earnings⁶.

5.1 The Subjective model

In Table 4 we report the average (OLS) and quantile returns to schooling obtained with the ECHP data under different specifications. All the estimates are controlling for labour market experience, training provided by the employer, job tenure, unemployment experience, marital status, immigrant condition, health status, establishment size, industry and occupation.

5.1.1 OLS estimates

In Panel 1 we report the results of a simplified version of Eq. (4) which does not include the mismatch variables (‘Standard model with education levels’). We find that the average wage premium of tertiary and secondary education is, respectively, 51.9% and 16.7%⁷. These

⁶ The estimates for the full set of controls are available upon request.

⁷ Through the paper we refer to the coefficients reported in the tables as ‘wage effects’ or ‘wage differentials’. To be precise, however, the percentage wage difference is given by $e^{\beta}-1$, rather than by β itself, especially when the

coefficients change by little (to 49.4% and 15.1%, respectively) when we add, in Panel 2, the vector M_i to control for educational mismatches ('Subjective model'). This new specification shows that skills mismatched and strongly mismatched workers earn on average 7.3% and 4.1% less, respectively, than their well matched counterparts, while the overqualified are not exposed to a significant pay penalty. Two things are worth noting. First, there is evidence to suggest that strongly mismatched workers should be considered as a distinct group: they have excess qualifications, but earn 4.1% less than the overqualified, and they lack necessary skills, but earn 3.2% more than the skills mismatched. In other words, this group of workers manages to partially compensate the lack of necessary skills with excess education. Second, the null impact of overqualification on wages seems to be at odds with earlier research reporting negative and significant effects. However, we must note that our group of overqualified workers is restricted to include only those who have the necessary skills for their jobs. Those others with insufficient skills (the strongly mismatched) do earn lower wages. Moreover, similar studies focus on university graduates, among which the effects of mismatch are particularly large. This becomes transparent in our next specification.

In Panel 3, we include a full set of interactions between the education groups and the different types of mismatch ('Subjective model with interactions'). As is apparent, the pay penalty of mismatch is closely related to the educational attainment of the individual. Specifically, we find that workers with a university degree are exposed to large wage decreases if they enter in jobs for which they are skills mismatched (-26.8%) or strongly mismatched (-16.6%). These effects are almost four times larger than among individuals with the lowest education level (-7.3% and -4.2%, respectively) and turn to non-significant in the upper secondary group. Finally, we note that overqualification *per se* does not bring a significant wage decrease among any of the education groups.

5.1.2 Quantile estimates

Question 1: The inequality increasing effect of education. In line with previous works, we find that the 'Standard model with education levels' yields returns to education that are highly increasing over the wage distribution (Panel 1). The return to tertiary education goes from 38.3% in the first quantile up to 61.2% in the top quantile, while the return to secondary education rises from 10.4% to 21.8%. The differential effect between the .10 and the .90

estimates are large. We do not perform this transformation to facilitate the correspondence between the tables and the text.

quantile is reported in the second last column of Table 4. This statistic summarizes the excess conditional wage inequality within the education group relative to the reference group (less than upper secondary education). Thus, for example, the 23 percentage points (pp) differential in the tertiary level indicates that the wage gap between two workers who are seemingly equal but located at the two extreme quantiles will be 23% higher if these workers have tertiary and not less than upper secondary education. In the last column, we list the p-values of F-tests for the joint equality of coefficients at all quantiles. According to the results, differences across quantiles are statistically significant. This is, in sum, the essence of the ‘inequality increasing effect’ of education documented in Buchinsky (1994), Pereira and Martins (2002), Fersterer and Winter-Ebmer (2003), Martins and Pereira (2004) and Machado and Mata (2005).

Questions 2 and 2a: Educational mismatches and the inequality increasing effect of education.

In Panel 2 we include explicit controls for mismatch. We find that conditional wage dispersion within the educated is slightly lower in the resulting model. Specifically, when we switch from the standard model to the ‘Subjective model’, the .90-.10 spread of tertiary education decreases by 26.1%, from 23.0 to 17.0 pp, while the corresponding spread for secondary education decreases by 32.5%, from 11.4 to 7.7 pp. Even though these reductions are sizable, most of the dispersion remains, and the corresponding tests reject the equality of coefficients across quantiles at conventional significance levels. It is worth noting that the estimating equation controls for three different types of educational mismatch simultaneously (overqualification, skills mismatch and strong mismatch). Therefore, the results indicate that the inequality increasing effect of education cannot be attributed to the incidence of various types of mismatch.

This is also the case in Panel 3, where we allow the mismatch effect to differ across education groups. Indeed, the results are now more conclusive. Far from a reduction, the ‘Subjective model with interactions’ exhibits larger conditional wage dispersion than the standard model. Specifically, the within-group conditional dispersion rises by 9.5% in the tertiary level and by 51.1% in the upper secondary level when we switch from Panel 1 to Panel 3. This can be seen by comparing the .90-.10 spreads of the tertiary and secondary group in the ‘Subjective model with interactions’ (25.3 and 25.7 pp, respectively) with the corresponding spreads in the ‘Standard model’ (23.0 and 11.4 pp, respectively). 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 (various types of) mismatch. This is particularly evident in this more elaborated

model, according to which returns to secondary and tertiary education among well-matched workers are highly increasing over the earnings distribution.

For illustrative purposes, in Figures 2 and 3 we plot the quantile-return profile of education under the different specifications. As is apparent, the returns in the extended models are as increasing over the wage distribution as in the standard model. To provide a sensitivity analysis, in Table B1 of Appendix B we have re-estimated the coefficients using a restricted rather than a full set of controls in the earnings equations⁸. The first lesson is that returns to education are roughly doubled when we switch from the full to the restricted specification, a result that is consistent with the sensitivity analyses reported in the literature (Card, 2001). The same applies to the mismatch coefficients: about 50% of these effects can be attributed to observables that are correlated with mismatch. Thus, for example, the average pay penalty of skills mismatch and strong mismatch in the ‘Subjective model’ rises from 7.3% and 4.1%, respectively, to 14.3% and 9.9% when we switch from the full to the restricted specification. The second lesson is that in the restricted models the conditional wage dispersion among the educated largely persists after controlling for mismatch, as in the full models. Thus, for example, in Table B1 the .90-.10 spread in the tertiary group is 41.5 pp in the ‘Standard model’, 37.0 pp in the ‘Subjective model’ and 45.1 pp in the ‘Subjective model with interactions’. In all cases, the spread is statistically significant at the 1% level.

Question 3: Is the pay penalty of educational mismatch homogenous across individuals in different segments of the earnings distribution? Next, we turn to the quantile estimates of overqualification, skills mismatch and strong mismatch. The results are interesting on their own, as they document substantial differences across earnings quantiles. One finding stands out prominently from Table 4: the effects of educational mismatches are increasing as we move up

⁸ Following a human capital interpretation of the mismatch phenomenon, some authors have suggested that workers may accept mismatched work in exchange of training, to compensate for low tenure and experience, or to access higher level occupations (Sicherman, 1991, Groot, 1996, Sloane et al., 1999). The full set of controls used in our earnings equations is aimed to remove the impact of these and other variables from the mismatch effect. It may be argued, however, that most of these covariates are endogenous and that a more parsimonious specification would capture the ‘true’ penalty of mismatch more appropriately. Thus, for example, we may not include controls for training. As the acceptance of mismatched work may allow some individuals to participate in training activities that later on are rewarded in the labour market, we could interpret these wage gains as a return to mismatch rather than a return to training. A similar argument applies to other variables, such as occupation and tenure. Following this reasoning, the estimates reported in Appendix B are obtained dropping from vector X_i all the controls except experience.

the wage distribution. The results are illustrative. In Panel 2, the coefficient of overqualification switches from non-significant to a (mildly) significant -5.5% when moving from the .10 to the .90 quantile. The coefficient of skills mismatch more than doubles, ranging from -5.1% in the first quantile up to -14.7% in the top quantile, and the pay penalty of strong mismatch more than triples, going from a non-significant -2.7% to a significant -9.1%. A glance to the second last column of Table 4 shows that the .90-.10 spreads for overqualification, skills mismatch and strong mismatch amount to, respectively, -9.3, -9.7 and -6.3 pp.

These spreads are statistically significant, and rise further when we switch to Panel 3. In the 'Subjective model with interactions' the corresponding pay penalties are remarkably large in the upper segments of the distribution, particularly among the educated. At the top quantile, the overqualification effect is as large as -18.9% for workers with a tertiary education and -6.6% for those with an upper secondary education. The .90-.10 spreads of the overqualification variable amount to -16.4 and -23.3 pp in these two groups. Similarly, the wage effects of skills mismatch at the top quantile are as large as -66.0% in the tertiary level and -18.5% in the upper secondary level, and the corresponding spreads rocket to -71.8 and -29.6 pp, respectively. Finally, we find that the .90-.10 spread of the strong mismatch effect is not significant among university graduates, but it amounts to -21.1 pp among workers with upper secondary education⁹.

In a related study, McGuinness and Bennet (2007) obtain a pay penalty of overqualification that is highly decreasing, not increasing, along the earnings distribution. Such divergence with our results may be due to the fact that theirs are based on a sample of recent graduates, among which the overqualified status is more likely to have a transient nature. It is likely that among this group, the overeducated are either high-ability individuals who accept mismatched work to access high-level occupations (and high wages) or low-ability individuals who, immediately after graduation, enter in low-level jobs while they search for suitable jobs. This would be consistent with having decreasing effects of overqualification over the wage distribution. Another difference is that McGuinness and Bennet base their results on the overqualification/non-overqualification distinction, and do not control for other types of

⁹ The more differentiated view provided in Panel 3 comes at the cost of reduced cell size in some groups. Thus, for example, the number of overqualified workers drops from 1,047 in the total sample to 231 when we consider only the group of workers with university education, and a more reduced group results when we consider skills mismatches and strong mismatch among university graduates. Even though the interaction coefficients reported in Panel 3 exhibit moderate standard errors and are not erratic across quantiles, the reduced cell size of specific groups in a quantile regression framework recommend us to interpret the results with some caution.

mismatch that can be equally relevant for wages, such as skills mismatches. As a result, their estimates across quantiles may be obscuring subtle differences across heterogeneous workers.

Educational mismatches and unobserved ability. All in all, the results uncover across-quantiles variation at large scale. Somewhat surprising, this variation has been typically overlooked in the literature, even though it can provide useful hints to better understand the mismatch phenomenon. In the quantile regression framework, the estimates at different quantiles represent the effects of a given covariate for individuals that have the same observable characteristics but, due to unobservable earnings capacity, are located in different segments of the conditional distribution. Therefore, those workers who end up in high-paid jobs are those who have more productive abilities, where by abilities we refer to those marketable skills, academic credentials and motivations that allow a worker to earn a higher wage given a vector of observable characteristics. Having the labour market segmented by ability deciles, with individual ability indexed by the individual's position in the conditional wage distribution, the estimates at different quantiles provide snap-shots of how mismatched individuals within the different ability groups are impacted. Turning to the results, we find that educational mismatches are events that reduce wages amongst all ability groups. As is apparent in Table 4, in almost every quantile and every education group is the pay penalty of skills mismatch found to significantly decrease wages. And the same holds, though to a lesser extent, for the strong mismatch effect. Taken together, these observations can be hardly reconciled with the frequent interpretation that mismatched workers earn less because they are less able. At the top of that, we find that workers in the high-ability segments of the distribution are precisely those who are exposed to larger wage decreases if they end up in mismatched work. This observation gives further support to the view that educational mismatches represent a complex phenomenon in which workers with very different backgrounds and abilities are involved. Recently, De Grip et al. (2008) have shown that matched and mismatched workers show similar levels of ability, including verbal memory, cognitive flexibility, verbal fluency and information processing speed. Our results are consistent with this view, as they suggest that describing the mismatch phenomenon as a result of lower ability levels is an oversimplification.

5.2 The Statistical model

In Table 5 we present the results of the second model. These are calculated controlling for labour market experience, training participation, job tenure, marital status, type of contract,

immigrant condition, establishment size and occupation.

5.2.1 OLS estimates

In Panel 1 we report the estimates of a standard earnings equation with actual years of schooling in the right hand side ('Standard model with years of schooling'), and in Panel 2 we report the results of the ORU model. First we focus on the OLS estimates. We find that the returns to surplus schooling (3.9%) are lower than the returns to required schooling (6.6%), and that a year of deficit schooling carries a significant wage penalty (-4.8%). Consistent with this, the returns to required schooling are above the returns to actual years of schooling (4.3%). These regularities are in line with the survey of the evidence reported in Hartog (2000) and McGuinness (2006)¹⁰.

5.2.2 Quantile estimates

Next, we examine the dispersion across quantiles obtained with the two specifications. Our answers to *Question 2* and *Question 2a* are consistent with those given by the ECHP dataset: returns to schooling are highly increasing over the wage distribution and educational mismatches cannot account for this fact.

In Panel 1 the return to an additional year of schooling ranges from 2.9% in the bottom quantile to 4.9% in the top quantile, and the spread is statistically significant. Adding explicit controls for over- and under-qualification in Panel 2 does not reduce the within-education-groups dispersion. Far from a reduction, the .90-.10 spread of required schooling (5.3 pp) is found to be 2.8 times larger than the .90-.10 spread of actual years of schooling (1.9 pp). The results in Appendix B are similar, with the coefficient of required schooling ranging from 5.6% in the bottom quantile up to 9.6% in the top quantile. This indicates that the inequality increasing effect of education is indeed sharper among the well-matched than among the total working population (Figure 4). Or, in other words, conditional wage dispersion among the educated would be also substantial if there were not mismatched workers in the labour market.

Question 2b: Degree of mismatch and the inequality increasing effect of education. A particularity of the statistical model is that it controls not only for the type of educational mismatch (over- versus under-qualification) but also for the amount of mismatch (as measured

¹⁰ Across studies, the return to surplus, required and deficit schooling range from 3% to 5%, from 5% to 11%, and from -2% to -6%, respectively.

by years of over- or under-schooling). The estimates unambiguously show that, although potentially relevant, this ingredient does not alter our main conclusion.

Question 3: Is the pay penalty of educational mismatch homogenous across individuals in different segments of the earnings distribution? In an earlier work, Hartog et al. (2001) used an ORU specification to explore the evolution of returns to schooling in Portugal from 1982 to 1992. They found that during the sample period the return to a year of schooling above the job requirement was typically higher among workers with higher earnings. Our results are consistent with this view. Specifically, we detect substantial individual heterogeneity in the returns to over- and under- schooling. The returns to overschooling are low among low-ability workers and high among high-ability workers. The estimated coefficient more than doubles as we move from the very bottom to the top of the distribution, going from 2.3% to 4.8%. The pay penalty of underschooling shows a similar pattern, ranging from 3.7% to 4.9%. It is interesting to note that as we move from the bottom to the top quantile, the returns to required schooling increase by 5.3 pp, i.e., 2.2 times more than the returns to overschooling (2.4 pp). This implies that relative to required schooling, the wage loss of overschooling increases over the wage distribution, and the same applies to underschooling. These profiles confirm, again, the earlier findings from the ECHP: the wage effects of mismatch are higher precisely among those workers with higher unobserved earnings capacity.

6. Conclusions

Returns to education are increasing over the wage distribution. While researchers have focused on the inequality implications of this finding, little attention has been paid to its potential causes. This is the first paper that formally tests whether educational mismatches can account for the tendency of education to be less rewarded in low-paid jobs. Answering this question was compelling, as in the political arena educational mismatches are being put forward as one of the major sources behind the high earnings dispersion of educated workers in developed countries. We have used two alternative datasets (the ECHP and the PLFS), two different measurement methods (the ‘subjective’ and the ‘statistical’ approach), different definitions of mismatch (‘overqualification’, ‘underqualification’, ‘skills mismatch’ and ‘strong mismatch’) and alternative earnings equations.

The paper has produced two main findings. Both of them are robust to changes in the

measurement and definition of mismatch and changes in the estimating model. The first result is that within-education-groups wage dispersion is at least as large among well-matched workers as among the total working population. Therefore, we must reject the hypothesis that the higher wage dispersion among the educated is due to the prevalence of (different types and degrees of) educational mismatches in the labour market.

The second result is that the wage effects of mismatch are by no means constant over the conditional wage distribution. Rather, they are found to be remarkably larger at the upper segments of the distribution. This result contributes to a better perception of the causes and consequences of the phenomenon. First, it highlights the substantial individual heterogeneity that surrounds the mismatch effect. This heterogeneity stems not only from differences between education groups (university graduates are exposed to larger wage losses if they end up in mismatched work) but within groups as well. Researchers and policy makers should take this heterogeneity into account when attempting to ascertain the impact of educational mismatches on different population groups and on the total earnings distribution. To that purpose, focusing on averages may be seriously misleading. Second, most of the debate in the policy arena has gravitated around the question of to what extent the incidence of mismatch entails a productivity loss. It is very difficult to know whether the lower earnings observed for mismatched workers are caused by their mismatch, or whether individuals with lower earnings capacity end up in mismatched work. Several papers have explored this issue using panel data (Bauer, 2002), proxies of skills (McGuinness, 2003) and treatment effects models (Dolton and Silles, 2008). In this paper we have provided an alternative view using quantile regression. The major advantage of this approach is that it documents how workers who are mismatched within homogenous ability groups are impacted relative to their well-matched counterparts. We have found evidence to suggest that interpreting the mismatch phenomenon as a consequence of the low ability and skills possessed by some workers can be overly simplistic. Indeed, workers with higher unobserved ability are precisely those who are more heavily penalized in mismatched jobs. We claim, therefore, that educational mismatches are to a large extent the result of real inefficiencies in which the worker's productivity potential is constrained by the job class.

A limitation of the paper is that schooling levels are assumed to be exogenous. This simplification is based on several considerations. First, the main hypothesis outlined in this study is inspired in consistent evidence reported in Buchinsky (1994), Pereira and Martins (2002), Fersterer and Winter-Ebmer (2003), Martins and Pereira (2004), Machado and Mata

(2005). On the account of the extra complication of quantile regression, these authors abstract from selection effects, and so we do. Re-assessing the existing evidence and re-testing our hypotheses in a context of endogenous schooling is beyond the scope of the present paper. And second, there is evidence to suggest that assuming exogenous schooling is not crucial for the results. Several authors have shown that standard returns to education do not change by more than one third when ability and selectivity are taken into account (Card, 2001, Trostel et al., 2002). Moreover, there are results in the literature showing that adding explicit controls for ability in a quantile regression framework produces similar levels of within-education-groups wage dispersion (Chernozhukov et al., 2007).

Finally, there are several policy implications arising from the analysis. First, education pays off in terms of wages, although the dispersion of this payoff increases as individuals invest in more education. We have shown that the increased dispersion is independent from the prevalence of educational mismatches in the labour market. This finding warns prospective students that even if they end up in matched jobs the returns to their educational investment are subject to a substantial amount of uncertainty. As investors tend to avoid risks, it is likely that this uncertainty exerts a negative effect on the demand for further education among certain individuals. In this respect, policies oriented to reduce the variability in the private returns to schooling by easing the school-to-work transition and by relating the private costs of education to the individual future earnings may be of particular importance.

Second, countries make substantial investments from both public and private sources in education. It is important to ensure that the education programmes they support are effective and that the benefits are distributed equitably. However, the mismatch phenomenon points to existing rigidities in labour markets that limit the capacity of societies to fully utilise and reward highly educated workers. In this respect, policies aimed to improve the integration between the schooling system and the changing demand for different types of skills in the society seem to be in order.

An agenda for further research is, first, gathering new data with detailed information on the individuals' abilities, skills and educational background and, second, improving the measurement of mismatch. The phenomenon requires a multidimensional approach that simultaneously considers all the knowledge, capabilities and skills that can be needed in a job and compares them to those possessed and effectively used by the individual. These extensions

would allow us to obtain a more refined view on how workers who are mismatched in several ways and to different degrees are penalized in the labour market. Such analysis would help the profession in the task of disentangling what individual and institutional factors and to what extent are responsible for the existence of mismatched workers.

Appendix A. Definition of variables.

Net hourly wage. ECHP and PLFS: monthly net salary in the main job (in euros) divided by four times the weekly hours worked in the main job.

Education. Maximum level of completed schooling. ECHP: three categories based on the ISCED-97 classification: less than upper secondary, upper secondary, and tertiary education. PLFS: ten categories, each of them paired with the corresponding years of schooling.

Training. Dummy variable. ECHP: activated if the employer provided training to the worker during the previous year. PLFS: activated if the worker has ever participated in a training activity.

Experience. ECHP and PLFS: age minus age of first job.

Tenure. ECHP and PLFS: difference between the year of the survey and the year of the start of the current job. Three categories were constructed: from 1 to 4 years, from 5 to 9 years, and 10 years or more.

Permanent contract. PLFS: dummy variable. Takes the value 1 if the individual has a permanent contract, zero otherwise.

Single. ECHP and PLFS: dummy that takes the value 1 if the individual is single (including widow and divorced), zero otherwise (married or living in a couple).

Immigrant. ECHP and PLFS: dummy activated if the individual was born in a foreign country.

Services. ECHP: dummy that takes the value 1 if the individual works in the services sector, zero if he works in the industry sector.

Firm size. ECHP and PLFS: decomposed into four categories, from 1 to 19 employees, from 20 to 99 employees, from 100 to 499 employees, and 500 employees or more .

Bad health. ECHP: individuals report their health status using a scale that ranges from 1 (very good) to 5 (very bad). The dummy 'bad health' takes value one if the answer is 4 or 5.

Past unemployment. ECHP: dummy variable, activated in case of unemployment experience before current job.

Occupation. ISCO-88 classification disaggregated at the 2-digit level up to 25 occupations (ECHP) and National Classification of Occupations disaggregated at the 2-digit level up to 29 occupations (PLFS). In Tables 1 and 3 this variable has been aggregated into 9 broader categories.

Appendix B. Estimates with a restricted set of controls

Table B1. Returns to schooling and mismatch effects - ECHP, restricted set of controls

	OLS	$\theta = .10$	$\theta = .20$	$\theta = .30$	$\theta = .40$	$\theta = .50$	$\theta = .60$	$\theta = .70$	$\theta = .80$	$\theta = .90$	$\Delta(.90-.10)$	F-test
Panel 1. Standard model with education levels												
Tertiary	0.957 *** 0.023	0.697 *** 0.029	0.799 *** 0.027	0.857 *** 0.030	0.921 *** 0.021	0.979 *** 0.039	1.019 *** 0.023	1.077 *** 0.029	1.077 *** 0.027	1.112 *** 0.030	0.415 ***	30.73 ***
Upper Secondary	0.308 *** 0.013	0.155 *** 0.021	0.205 *** 0.014	0.214 *** 0.014	0.247 *** 0.016	0.284 *** 0.017	0.323 *** 0.018	0.365 *** 0.017	0.408 *** 0.018	0.442 *** 0.020	0.288 ***	38.41 ***
R-squared	0.400	0.179	0.180	0.179	0.182	0.192	0.207	0.228	0.264	0.297		
Panel 2. Subjective model												
Tertiary	0.851 *** 0.024	0.628 *** 0.031	0.710 *** 0.034	0.784 *** 0.029	0.821 *** 0.031	0.880 *** 0.034	0.897 *** 0.029	0.953 *** 0.027	0.947 *** 0.027	0.998 *** 0.039	0.370 ***	26.10 ***
Upper Secondary	0.255 *** 0.013	0.134 *** 0.021	0.170 *** 0.016	0.188 *** 0.013	0.206 *** 0.014	0.243 *** 0.014	0.262 *** 0.018	0.303 *** 0.016	0.332 *** 0.019	0.376 *** 0.022	0.242 ***	19.86 ***
<i>Overqualification</i>	0.020 0.017	0.038 * 0.022	0.041 ** 0.019	0.035 ** 0.017	0.043 * 0.022	0.036 * 0.020	0.021 0.025	-0.017 0.020	-0.011 0.022	-0.026 0.032	-0.064 *	2.17 *
<i>Skills mismatch</i>	-0.143 *** 0.014	-0.084 *** 0.018	-0.085 *** 0.016	-0.092 *** 0.014	-0.106 *** 0.016	-0.122 *** 0.016	-0.148 *** 0.022	-0.190 *** 0.018	-0.190 *** 0.017	-0.213 *** 0.026	-0.129 ***	9.12 ***
<i>Strong mismatch</i>	-0.099 *** 0.014	-0.061 *** 0.017	-0.055 *** 0.017	-0.058 *** 0.015	-0.076 *** 0.018	-0.087 *** 0.017	-0.111 *** 0.024	-0.123 *** 0.020	-0.127 *** 0.017	-0.132 *** 0.026	-0.071 ***	3.71 ***
R-squared	0.418	0.187	0.188	0.190	0.193	0.204	0.219	0.243	0.280	0.313		
Panel 3. Subjective model with interactions												
Tertiary	0.943 *** 0.040	0.661 *** 0.062	0.781 *** 0.075	0.901 *** 0.078	0.995 *** 0.063	0.996 *** 0.032	0.988 *** 0.047	0.989 *** 0.045	1.041 *** 0.067	1.112 *** 0.066	0.451 ***	9.37 ***
Upper Secondary	0.209 *** 0.038	-0.015 0.038	0.075 0.047	0.099 *** 0.032	0.141 *** 0.041	0.190 *** 0.039	0.245 *** 0.071	0.328 *** 0.054	0.394 *** 0.056	0.435 *** 0.048	0.450 ***	13.97 ***
<i>Overqualification</i>												
Tertiary	-0.099 ** 0.047	-0.044 0.079	-0.060 0.084	-0.108 0.085	-0.156 ** 0.064	-0.111 ** 0.049	-0.079 0.056	-0.035 0.057	-0.101 0.067	-0.136 * 0.073	-0.092	0.90
Upper Secondary	0.106 *** 0.040	0.185 *** 0.048	0.154 *** 0.052	0.169 *** 0.042	0.182 *** 0.044	0.151 *** 0.041	0.096 0.070	0.006 0.055	-0.035 0.055	-0.066 0.051	-0.251 ***	5.17 ***
Less than Upper Sec.	0.012 0.020	-0.004 0.021	0.012 0.023	0.017 0.016	0.021 0.026	0.021 0.019	0.015 0.028	-0.021 0.030	-0.025 0.028	-0.008 0.041	-0.004	1.16
<i>Skills mismatch</i>												
Tertiary	-0.284 *** 0.101	-0.007 0.077	-0.196 ** 0.093	-0.367 *** 0.096	-0.420 *** 0.080	-0.450 *** 0.075	-0.500 *** 0.080	-0.608 *** 0.111	-0.628 *** 0.159	-0.760 *** 0.194	-0.753 ***	10.59 ***
Upper Secondary	-0.052 0.040	0.062 0.043	-0.013 0.047	-0.017 0.034	-0.055 0.043	-0.107 ** 0.044	-0.147 ** 0.065	-0.262 *** 0.054	-0.293 *** 0.064	-0.256 *** 0.070	-0.318 ***	5.97 ***
Less than Upper Sec.	-0.093 *** 0.016	-0.111 *** 0.016	-0.107 *** 0.016	-0.100 *** 0.012	-0.104 *** 0.015	-0.114 *** 0.015	-0.125 *** 0.022	-0.177 *** 0.023	-0.181 *** 0.019	-0.198 *** 0.029	-0.087 ***	4.39 ***
<i>Strong mismatch</i>												
Tertiary	-0.401 *** 0.068	-0.439 ** 0.214	-0.291 * 0.150	-0.405 *** 0.157	-0.284 * 0.146	-0.263 *** 0.099	-0.274 *** 0.095	-0.376 *** 0.134	-0.318 0.205	-0.265 *** 0.317	0.174	0.51
Upper Secondary	-0.108 *** 0.042	0.087 * 0.050	0.042 0.046	0.026 0.030	-0.025 0.046	-0.048 0.045	-0.081 0.069	-0.146 ** 0.057	-0.201 *** 0.058	-0.230 *** 0.057	-0.317 ***	6.88 ***
Less than Upper Sec.	-0.137 *** 0.015	-0.087 *** 0.016	-0.085 *** 0.016	-0.066 *** 0.012	-0.074 *** 0.016	-0.078 *** 0.016	-0.087 *** 0.025	-0.115 *** 0.026	-0.116 *** 0.021	-0.109 *** 0.033	-0.022	1.55
R-squared	0.420	0.190	0.016	0.192	0.196	0.207	0.221	0.244	0.281	0.313		
No. of observations	8,319											

Notes to Table B1: i) dependent variable: log hourly wages; ii) standard errors are in smaller type; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors have been calculated using a bootstrap method of 500 replications; v) * denotes significant at the 10% confidence level, ** denotes significant at the 5% confidence level, *** denotes significant at the 1% confidence level; vi) the last column reports, for each covariate i , the F-statistic of the test $H_0 : \beta_{0.1,i} = \beta_{0.2,i} = \dots = \beta_{0.9,i}$, $H_1 : \beta_{j,i} \neq \beta_{h,i}$ for some $j \neq h$; vii) Controls: labour market experience and year dummies.

Table B2. Returns to schooling and mismatch effects - PLFS, restricted set of controls

	OLS	$\theta = .10$	$\theta = .20$	$\theta = .30$	$\theta = .40$	$\theta = .50$	$\theta = .60$	$\theta = .70$	$\theta = .80$	$\theta = .90$	$\Delta(.90-.10)$	F-test
Panel 1. Standard model with years of schooling												
Actual years of schooling	0.069 *** 0.002	0.043 *** 0.002	0.050 *** 0.002	0.054 ** 0.002	0.059 *** 0.002	0.063 *** 0.002	0.069 *** 0.002	0.073 *** 0.002	0.079 ** 0.001	0.084 *** 0.002	0.041 **	75.12 **
R-squared	0.158	0.056	0.085	0.102	0.111	0.129	0.146	0.168	0.185	0.198		
Panel 2. ORU model												
Years of overschooling	0.042 *** 0.003	0.027 *** 0.003	0.031 *** 0.002	0.035 ** 0.002	0.036 *** 0.002	0.038 *** 0.002	0.041 *** 0.002	0.043 *** 0.003	0.047 ** 0.002	0.051 *** 0.004	0.024 **	8.78 **
Years of required schooling	0.080 *** 0.002	0.056 *** 0.002	0.063 *** 0.002	0.068 ** 0.001	0.073 *** 0.002	0.078 *** 0.002	0.082 *** 0.002	0.086 *** 0.001	0.091 ** 0.002	0.096 *** 0.002	0.041 **	63.33 **
Years of underschooling	-0.057 *** 0.004	-0.044 *** 0.003	-0.050 *** 0.003	-0.053 ** 0.003	-0.055 *** 0.003	-0.056 *** 0.003	-0.057 *** 0.002	-0.056 *** 0.003	-0.057 ** 0.003	-0.061 *** 0.004	-0.017 **	4.03 **
R-squared	0.179	0.069	0.100	0.120	0.133	0.154	0.172	0.194	0.211	0.221		
No. of observations	11,947											

Notes to Table B2: i) dependent variable: log hourly wages; ii) standard errors are in smaller type; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors have been calculated using a bootstrap method of 500 replications; v) * denotes significant at the 10% confidence level, ** denotes significant at the 5% confidence level, *** denotes significant at the 1% confidence level; vi) the last column reports, for each covariate i , the F-statistic of the test $H_0 : \beta_{0.1,i} = \beta_{0.2,i} = \dots = \beta_{0.9,i}$, $H_1 : \beta_{j,i} \neq \beta_{h,i}$ for some $j \neq h$; vii) Controls: labour market experience and dummies for quarter.

References

- Alba-Ramírez, A. and M. Blázquez (2002), Types of Job Match, Overeducation, and Labour Mobility in Spain, in *Overeducation in Europe: Current Issues in Theory and Policy*, Büchel, F., A. de Grip and A. Meitens (eds), 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.
- Ashenfelter, O. and C. Rouse (2000), Schooling, Intelligence, and Income in America, in *Meritocracy and economic inequality*, K. Arrow, S. Bowles and S. Durlauf, (eds), Princeton University Press, Princeton, N.J.
- Battu, H., Belfield, C. and P. Sloane (2000), How well can we measure graduate overeducation and its effects?, *National Institute Economic Review* 171, 82–93.
- Bauer, T. (2002), Educational mismatch and wages: a panel analysis, *Economics of Education Review* 21, 221–229.
- Bonin, H, T. Dohmen, A. Falk, D. Huffman and U. Sunde (2007), *Cross-sectional Earnings Risk and Occupational Sorting: The Role of Risk Attitudes*, *Labour Economics* 14(6), 926-937.
- 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.
- Budría, S. and A. Moro-Egido (2008), Education, Educational Mismatch, and Wage Inequality: Evidence for Spain, *Economics of Education Review* 27(3), 332-341.
- Card, D. (2001), Estimating the return to schooling: progress on some persistent econometric problems, *Econometrica*, 69(5), 1127-1160.
- Carneiro, P., K. Hansen, and J. J. Heckman (2003), Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice, *International Economic Review* 44(2), 361-422.
- Chernozhukov, V., C. Hansen and M. Jansson (2007), Inference approaches for instrumental variable quantile regression, *Economics Letters* 95, 272–277.
- Cohn, E. and P. Khan, (1995), The wage effects of overschooling revisited, *Labour Economics* 2, 67–76.

- Cunha, F., J. Heckman and S. Navarro (2005), Separating Uncertainty from Heterogeneity in Life Cycle Earnings, *Oxford Economic Papers* 57(2), 191-261.
- Cunha, F. and J. J. Heckman (2007), Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling, *Labour Economics* 14(6), 870-893.
- Daly, M., F. Buchel, and G. Duncan (2000), Premiums and penalties for surplus and deficit education: evidence from the United States and Germany, *Economics of Education Review* 19, 169–178.
- Decker, R., A. De Grip and H. Heijke (2002), The effects of training and overeducation on career mobility in a segmented labour market, *International Journal of Manpower* 23(2), 106–125.
- De Grip, A., H. Bosma, D. Willems and M. van Boxtel (2008), Job-worker mismatch and cognitive decline, *Oxford Economic Papers* 60(2), 237-253.
- Dolton, P. and A. Vignoles (2000), The incidence and effects of over-education in the UK graduate labour market, *Economics of Education Review* 19, 179-198.
- Dolton, P. and M. Silles (2008), The Effects of Overeducation on Earnings in the Graduate Labour Market, *Economics of Education Review* 27(2), 125-139.
- Dolton, P., R. Asplund and E. Barth (2009), *Education and Inequality Across Europe*, Edward Elgar.
- Duncan, J. and S. Hoffman (1981), The incidence and wage effects of overeducation, *Economics of Education Review* 1(1), 75–86.
- Eurostat (2007), *Structural indicators*, Luxembourg. Available at: <http://epp.eurostat.ec.europa.eu>
- 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. (1996), The incidence of, and returns to overeducation in the UK, *Applied Economics* 28, 1345-1350.
- Groot, W. and H. Maassen van den Brink (2000), Overeducation in the Labor Market: A Meta-Analysis, *Economics of Education Review* 19, 149-158.
- Halaby, C (1994), Overeducation and skill match, *Sociology of Education* 67(1), 47–59.
- 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.
- Hartog, J., E. Plug, L. Serrano and J. Vieira (2003), Risk compensation in wages, a replication, *Empirical Economics* 28, 639-647.

- Hartog, J. and L. Serrano (2007), Earnings Risk and Demand for Higher Education, *Journal of Applied Economics* 10(1), 1-28.
- Hartog, J. and W. Vijverberg (2007), On compensation for risk aversion and skewness affection in wages, *Labour Economics* 14(6), 938-956.
- Hogan, V. and I. Walker (2007), Education choice under uncertainty: implications for public policy, *Labour Economics* 14(6), 894-912.
- Kiker, B.F., M.C. Santos and M. Mendes de Oliveira (1997), Overeducation and undereducation: evidence for Portugal, *Economics of Education Review* 16(2), 111-125.
- 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.
- Maier, M., F. Pfeiffer und W. Pohlmeier (2003), Overeducation and Individual Heterogeneity, in *Overeducation in Europe: Current Issues in Theory and Policy*, Büchel, F., A. de Grip and A. Meitens (eds), Edward Elgar Publishing, Cheltenham, UK.
- Martins, P. and P. Pereira (2004), Does Education Reduce Wage Inequality? Quantile Regressions Evidence from Fifteen European Countries, *Labour Economics*, 11(3), 355-371.
- McGuinness, S. (2003), Graduate overeducation as a sheepskin effect: evidence from Northern Ireland, *Applied Economics* 35, 597-608.
- McGuinness, S. (2006), Overeducation in the Labour Market, *Journal of Economic Surveys* 20(3), 387-418.
- McGuinness, S. and J. Bennett (2006), Overeducation and the Graduate Labour Market: A Quantile Regression Approach, *Economics of Education Review* 26(5), 521-531.
- 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.
- Rubb, S. (2004), Overeducation in the labor market: a comment and re-analysis of a meta analysis, *Economics of Education Review* 22, 621-629.
- Sicherman, N. (1991), Mismatch in the Labor Market, *Journal of Labor Economics* 9(2), 101-122.
- 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 Mismatch Literature really Tell Us? Keynote Address, International Conference on Mismatch in Europe: What Do We Know?* 22-23 November, Berlin.

- Trostel, P., I. Walker and P. Woolley (2002), Estimates of the economic return to schooling for 28 countries, *Labour Economics* 9, 1-16.
- Verdugo, R. and N.T. Verdugo (1989), The impact of surplus schooling on earnings: Some additional findings, *Journal of Human Resources* 24, 629-695.
- Wasmer, E., P. Fredriksson, A. Lamo, J. Messina and G. Peri (2007), *The Macroeconomics of Education in Europe*, Oxford University Press.
- Wolfe, B. and R. Haveman (2001), Accounting for the Social and Non-market Benefits of Education, in J.F. Helliwell (ed.), *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-being: International Symposium Report*, Human Resources Development Canada and OECD.
- Verhaest, D. and E. Omey (2006), The impact of overeducation and its measurement, *Social Indicators Research*, 77(3), 419-448.
- Verhaest, D. and E. Omey (2009), The determinants of overeducation: different measures, different outcomes?, *International Journal of Manpower*, forthcoming.

Tables

Table 1. Descriptive statistics by mismatch status - ECHP

Variables	Adequately educated	Overqualified	Skills Mismatched	Strongly mismatched
Proportion of the total sample	10.4	12.6	44.6	32.4
Log hourly wage	1.19	1.25	0.89	0.91
Log hours	3.73	3.74	3.74	3.74
Tertiary education	0.15	0.22	0.01	0.01
Upper secondary education	0.21	0.33	0.05	0.13
Less than upper secondary education	0.64	0.45	0.94	0.86
Training	0.21	0.32	0.03	0.05
Experience	15.4	12.2	17.4	15.7
Tenure < 5	0.52	0.53	0.52	0.54
5 ≤ Tenure < 10	0.28	0.29	0.28	0.28
Tenure ≥ 10	0.20	0.18	0.20	0.18
Single	0.40	0.41	0.39	0.42
Immigrant	0.06	0.06	0.02	0.03
Services	0.51	0.58	0.32	0.43
Firm size < 20	0.51	0.45	0.65	0.62
20 ≤ Firm size < 100	0.27	0.27	0.24	0.24
100 ≤ Firm size < 500	0.13	0.17	0.08	0.10
Firm size ≥ 500	0.09	0.11	0.03	0.04
Bad health	0.03	0.01	0.06	0.03
Past unemployment	0.27	0.23	0.24	0.24
Legislators, senior officials and managers	0.03	0.05	0.00	0.01
Professionals	0.12	0.14	0.00	0.01
Technicians and associate professionals	0.13	0.18	0.02	0.04
Clerks	0.08	0.13	0.06	0.10
Service workers and shop and market sales workers	0.08	0.11	0.10	0.15
Skilled agricultural and fishery workers	0.00	0.01	0.01	0.00
Craft and related trades workers	0.34	0.22	0.45	0.37
Plant and machine operators and assemblers	0.16	0.12	0.19	0.17
Elementary occupations	0.06	0.04	0.17	0.15

Table 2. The incidence of mismatch by education levels (%)

	Total Sample	Less than upper secondary	Upper secondary	Tertiary
Overqualified	12.6	6.9	32.3	58.5
Skills mismatched	44.6	51.1	18.3	2.8
Strongly mismatched	32.4	33.9	32.4	6.6
Adequately educated	10.4	8.1	17.0	32.2

Table 3. Descriptive statistics by mismatch status - PLFS

Variables	Adequately educated	Overqualified	Underqualified
Proportion of the total sample	44.6	36.8	18.7
Actual years of schooling	5.49	7.70	4.98
Years of overschooling	0.00	3.35	0.00
Years of underschooling	0.00	0.00	4.35
Log hourly wage	1.06	1.01	1.12
Log hours	3.71	3.72	3.70
Training	0.05	0.09	0.09
Experience	24.3	14.7	22.8
Tenure < 5	0.38	0.50	0.36
5 ≤ Tenure < 10	0.19	0.23	0.20
Tenure ≥ 10	0.42	0.27	0.44
Permanent contract	0.47	0.47	0.47
Single	0.20	0.43	0.28
Immigrant	0.01	0.02	0.01
Firm size < 20	0.26	0.28	0.29
20 ≤ Firm size < 100	0.06	0.06	0.07
100 ≤ Firm size < 500	0.03	0.02	0.03
Firm size ≥ 500	0.65	0.64	0.61
Executive civil servants, industrial directors and executives	0.01	0.00	0.03
Professionals and scientist	0.03	0.00	0.05
Middle management and technicians	0.05	0.02	0.20
Administrative and related workers	0.06	0.01	0.32
Service and sales workers	0.07	0.15	0.09
Farmers and skilled agricultural and fisheries workers	0.03	0.02	0.03
Skilled workers, craftsmen and similar	0.43	0.48	0.14
Machine operators and assembly workers	0.19	0.21	0.04
Unskilled workers	0.11	0.11	0.09

Table 4. Returns to schooling and mismatch effects - ECHP

	OLS	$\theta = .10$	$\theta = .20$	$\theta = .30$	$\theta = .40$	$\theta = .50$	$\theta = .60$	$\theta = .70$	$\theta = .80$	$\theta = .90$	$\Delta(.90-.10)$	F-test
Panel 1. Standard model with education levels												
Tertiary	0.519 *** 0.029	0.383 *** 0.044	0.461 *** 0.047	0.493 *** 0.034	0.504 *** 0.036	0.496 *** 0.033	0.498 *** 0.034	0.534 *** 0.043	0.550 *** 0.066	0.612 *** 0.071	0.230 ***	2.74 **
Upper Secondary	0.167 *** 0.012	0.104 *** 0.020	0.141 *** 0.015	0.147 *** 0.014	0.154 *** 0.012	0.164 *** 0.013	0.166 *** 0.015	0.180 *** 0.017	0.189 *** 0.026	0.218 *** 0.025	0.114 ***	4.47 ***
R-squared	0.504	0.243	0.246	0.249	0.257	0.271	0.286	0.305	0.333	0.367		
Panel 2. Subjective model												
Tertiary	0.494 *** 0.029	0.351 *** 0.040	0.442 *** 0.052	0.474 *** 0.034	0.463 *** 0.040	0.476 *** 0.033	0.481 *** 0.034	0.501 *** 0.038	0.524 *** 0.074	0.521 *** 0.063	0.170 **	3.76 ***
Upper Secondary	0.151 *** 0.012	0.087 *** 0.020	0.122 *** 0.016	0.145 *** 0.012	0.141 *** 0.013	0.149 *** 0.012	0.155 *** 0.014	0.171 *** 0.017	0.161 *** 0.024	0.164 *** 0.029	0.077 **	4.11 **
<i>Overqualification</i>	-0.001 0.016	0.038 0.020	0.010 0.016	0.019 0.015	0.012 0.016	0.012 0.019	-0.016 0.023	-0.026 0.022	-0.066 ** 0.026	-0.055 * 0.029	-0.093 ***	2.13 *
<i>Skills mismatch</i>	-0.073 *** 0.013	-0.051 *** 0.018	-0.047 *** 0.013	-0.046 *** 0.012	-0.046 *** 0.012	-0.047 *** 0.016	-0.078 *** 0.020	-0.083 *** 0.018	-0.140 *** 0.021	-0.147 *** 0.020	-0.097 ***	5.48 ***
<i>Strong mismatch</i>	-0.041 *** 0.014	-0.027 0.018	-0.014 0.012	-0.019 0.012	-0.016 0.014	-0.020 0.016	-0.050 *** 0.019	-0.052 *** 0.020	-0.093 *** 0.022	-0.091 *** 0.023	-0.063 **	3.00 **
R-squared	0.507	0.247	0.248	0.252	0.259	0.273	0.288	0.307	0.338	0.373		
Panel 3. Subjective model with interactions												
Tertiary	0.552 *** 0.045	0.395 *** 0.076	0.511 *** 0.068	0.549 *** 0.059	0.564 *** 0.057	0.569 *** 0.053	0.567 *** 0.046	0.541 *** 0.057	0.579 *** 0.107	0.648 *** 0.080	0.253 **	2.44 **
Upper Secondary	0.113 *** 0.035	-0.051 * 0.029	0.008 0.037	0.047 0.031	0.081 * 0.043	0.154 *** 0.037	0.173 *** 0.051	0.241 *** 0.050	0.216 *** 0.039	0.206 *** 0.042	0.257 ***	10.79 ***
<i>Overqualification</i>												
Tertiary	-0.068 0.042	-0.025 0.071	-0.070 0.072	-0.071 0.059	-0.074 0.053	-0.058 0.052	-0.075 0.048	-0.002 0.052	-0.092 0.086	-0.189 ** 0.084	-0.164 **	1.96 *
Upper Secondary	0.045 0.036	0.166 *** 0.043	0.109 *** 0.042	0.112 *** 0.037	0.077 0.047	0.035 0.039	-0.014 0.053	-0.073 0.054	-0.098 ** 0.047	-0.066 ** 0.030	-0.233 ***	4.66 ***
Less than Upper Sec.	-0.001 0.019	0.001 0.018	-0.005 0.020	0.016 0.018	0.013 0.019	0.026 0.021	-0.013 0.023	-0.023 0.026	-0.047 0.032	-0.017 0.040	-0.018	1.59
<i>Skills mismatch</i>												
Tertiary	-0.268 *** 0.055	0.058 0.106	-0.113 0.089	-0.203 *** 0.079	-0.258 *** 0.087	-0.246 *** 0.087	-0.292 *** 0.073	-0.315 *** 0.074	-0.486 *** 0.091	-0.660 *** 0.087	-0.718 ***	7.42 ***
Upper Secondary	-0.030 0.038	0.111 *** 0.039	0.089 ** 0.044	0.055 0.037	0.029 0.087	-0.039 0.038	-0.103 ** 0.051	-0.165 *** 0.055	-0.207 *** 0.046	-0.185 *** 0.050	-0.296 ***	7.76 ***
Less than Upper Sec.	-0.073 *** 0.014	-0.079 *** 0.015	-0.063 *** 0.014	-0.051 *** 0.013	-0.043 *** 0.012	-0.032 ** 0.016	-0.070 *** 0.018	-0.072 *** 0.017	-0.117 *** 0.024	-0.119 *** 0.024	-0.040 *	4.06 ***
<i>Strong mismatch</i>												
Tertiary	-0.166 * 0.098	-0.176 0.104	-0.298 ** 0.122	-0.332 ** 0.140	-0.319 *** 0.111	-0.192 0.123	-0.241 ** 0.112	-0.201 0.139	-0.320 0.253	0.091 0.309	0.268	0.93
Upper Secondary	0.005 0.036	0.107 ** 0.045	0.126 *** 0.038	0.088 *** 0.033	0.048 0.043	-0.011 0.037	-0.063 0.048	-0.111 ** 0.051	-0.120 *** 0.040	-0.104 ** 0.044	-0.211 ***	5.94 ***
Less than Upper Sec.	-0.042 *** 0.015	-0.056 *** 0.015	-0.032 ** 0.015	-0.025 * 0.014	-0.012 0.013	-0.007 0.016	-0.044 ** 0.018	-0.039 ** 0.019	-0.072 *** 0.025	-0.064 ** 0.025	-0.008	3.33 ***
R-squared	0.508	0.250	0.251	0.254	0.261	0.274	0.289	0.308	0.339	0.374		
No. of observations	8,319											

Notes to Table 4: i) dependent variable: log hourly wages; ii) standard errors are in smaller type; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors have been calculated using a bootstrap method of 500 replications; v) * denotes significant at the 10% confidence level, ** denotes significant at the 5% confidence level, *** denotes significant at the 1% confidence level; vi) the last column reports, for each covariate i , the F-statistic of the test $H_0 : \beta_{0.1,i} = \beta_{0.2,i} = \dots = \beta_{0.9,i}$, $H_1 : \beta_{j,i} \neq \beta_{h,i}$ for some $j \neq h$; vii) Controls: labour market experience, training provided by the employer, job tenure, unemployment experience, marital status, immigrant condition, health status, establishment size, industry, occupation and year dummies.

Table 5. Returns to schooling and mismatch effects - PLFS

	OLS	$\theta = .10$	$\theta = .20$	$\theta = .30$	$\theta = .40$	$\theta = .50$	$\theta = .60$	$\theta = .70$	$\theta = .80$	$\theta = .90$	$\Delta(.90-.10)$	F-test
Panel 1. Standard model with years of schooling												
Actual years of schooling	0.043 *** 0.002	0.029 *** 0.002	0.034 *** 0.002	0.036 *** 0.002	0.038 *** 0.002	0.040 ** 0.002	0.043 ** 0.002	0.045 *** 0.002	0.048 ** 0.002	0.049 ** 0.003	0.019 **	10.86 ***
R-squared	0.210	0.120	0.145	0.159	0.167	0.182	0.197	0.217	0.232	0.239		
Panel 2. ORU model												
Years of overschooling	0.039 *** 0.003	0.023 *** 0.003	0.029 *** 0.002	0.031 *** 0.002	0.035 *** 0.002	0.037 ** 0.002	0.039 ** 0.002	0.042 *** 0.003	0.044 ** 0.003	0.048 ** 0.004	0.024 **	8.65 ***
Years of required schooling	0.066 *** 0.009	0.040 *** 0.008	0.051 *** 0.009	0.049 *** 0.007	0.055 *** 0.008	0.059 ** 0.009	0.061 ** 0.009	0.067 *** 0.007	0.072 ** 0.010	0.093 ** 0.016	0.053 **	2.97 **
Years of underschooling	-0.048 *** 0.004	-0.037 *** 0.004	-0.040 *** 0.002	-0.043 *** 0.002	-0.042 *** 0.003	-0.044 ** 0.002	-0.046 ** 0.003	-0.047 *** 0.003	-0.051 ** 0.003	-0.049 ** 0.005	-0.012 **	1.57
R-squared	0.210	0.121	0.146	0.160	0.167	0.183	0.198	0.217	0.232	0.239		
No. of observations	11,947											

Notes to Table 5: i) dependent variable: log hourly wages; ii) standard errors are in smaller type; iii) OLS estimation is heteroskedastic-robust; iv) quantile standard errors have been calculated using a bootstrap method of 500 replications; v) * denotes significant at the 10% confidence level, ** denotes significant at the 5% confidence level, *** denotes significant at the 1% confidence level; vi) the last column reports, for each covariate i , the F-statistic of the test $H_0 : \beta_{0.1,i} = \beta_{0.2,i} = \dots = \beta_{0.9,i}$, $H_1 : \beta_{j,i} \neq \beta_{h,i}$ for some $j \neq h$; vii) Controls: labour market experience, training participation, job tenure, marital status, type of contract, immigrant condition, establishment size, occupation and dummies for quarter.

Figures

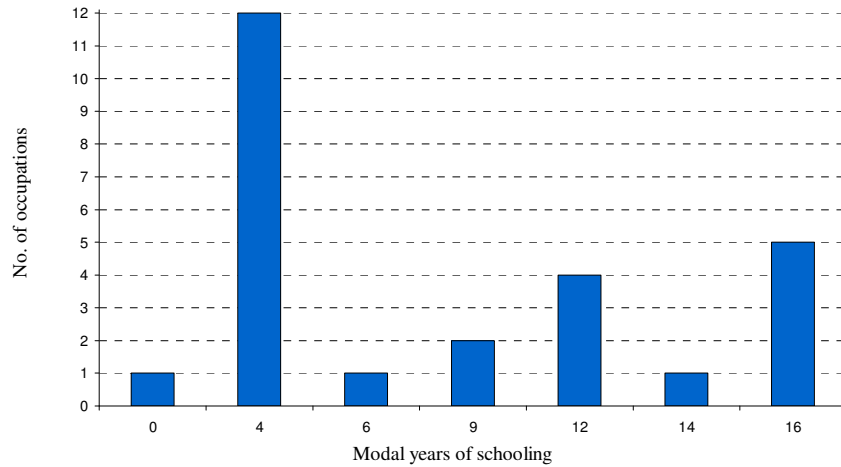


Figure 1. Modal years of schooling – Frequency. Based on the calculation of the modal years of schooling by occupations. The vertical bars represent the number of occupations in each education level. Source: Portuguese Labour Force Survey 2000-2002

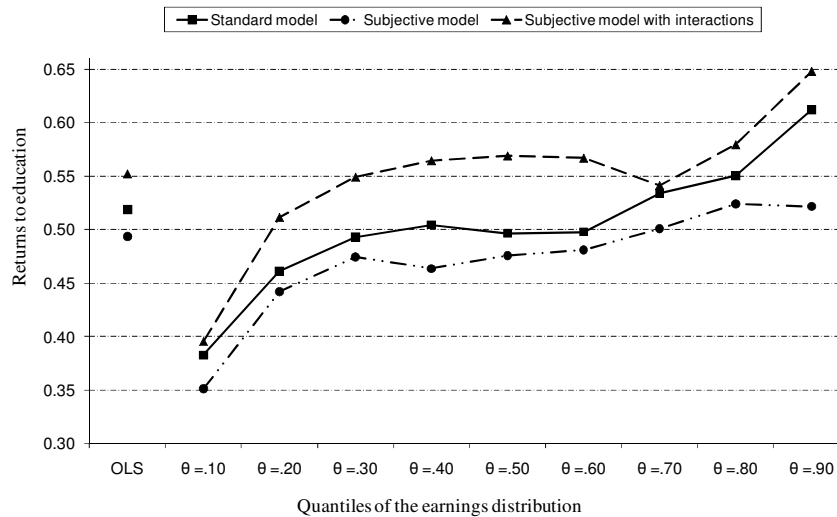


Figure 2. Quantile returns to tertiary education. This graph shows how the return to tertiary education changes when moving from the lower to the upper segments of the earnings distribution. X-axis: OLS and selected earnings quantiles. Y-axis: return to education. Source: European Community Household Panel 1994-2001.

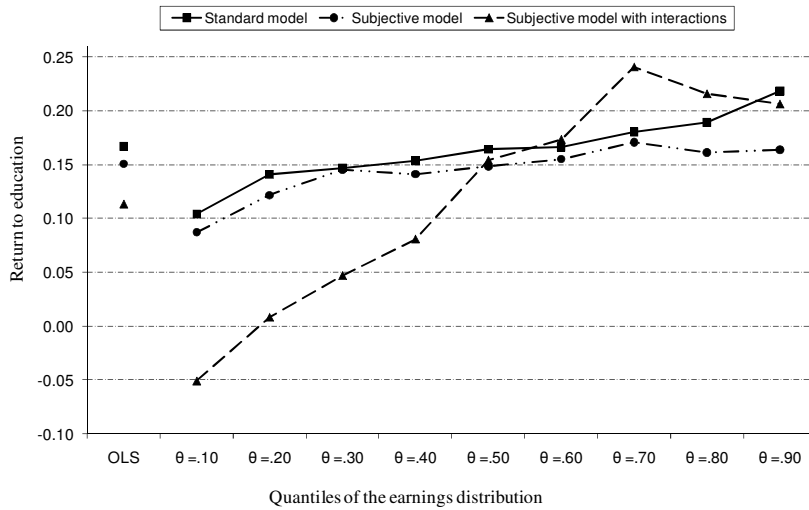


Figure 3. Quantile returns to upper secondary education. This graph shows how the return to upper secondary education changes when moving from the lower to the upper segments of the earnings distribution. X-axis: OLS and selected earnings quantiles. Y-axis: return to education. Source: European Community Household Panel 1994-2001.

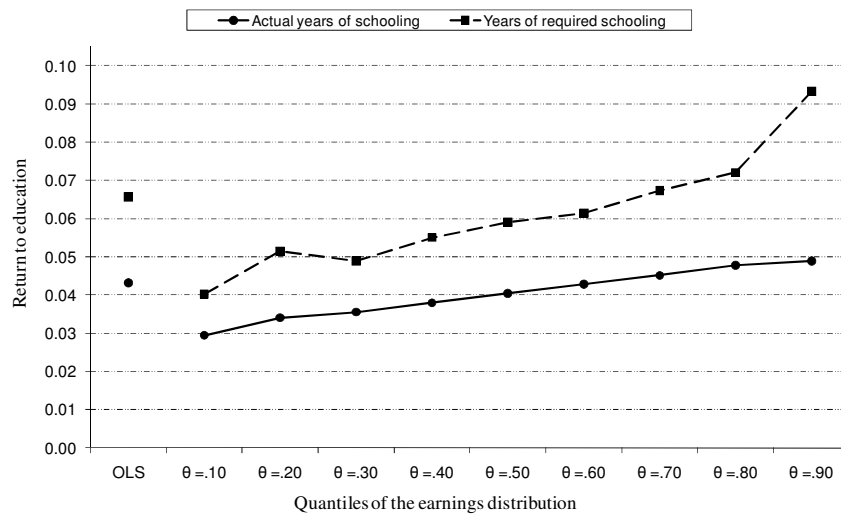


Figure 4. Quantile returns to actual and required years of schooling. This graph shows how the returns to schooling change when moving from the lower to the upper segments of the earnings distribution. X-axis: OLS and selected earnings quantiles. Y-axis: return to education. Source: Portuguese Labour Force Survey 2000-2002