Over-Education and Unobserved Skills: Evidence from a Matched Panel[#]

Pedro Martins[*]
University of St Andrews & University of Warwick

This version: 19 January 2004
First version: 3 December 2001

Abstract: I examine the wage implications of schooling/occupation mismatches, motivated by the evidence of a poor synchronisation between education systems and labour markets. My main contribution is in studying the role of unobserved heterogeneity in over- and under-education, drawing on a matched employer-employee panel. Although the identification of the key parameters is difficult due to the lack of variability in the regressors, some interesting results include the evidence found of selectivity in the transitions experienced by workers and of a "stepping-stone" interpretation of schooling mismatches.

JEL codes: C23, I21, J24, J31.

Keywords: Over-Education; Unobserved Heterogeneity; Matched Employer-Employee Panels; Fixed Effects.

[#] I thank comments by Martyn Andrews, Arnaud Chevalier, Joop Hartog, Reamonn Lydon, Robin Naylor, Christopher Pissarides and Ian Walker, and seminar participants at the Universities of Copenhagen, Birmingham, Warwick, Azores and Leicester. I also acknowledge financial support from *Fundação para a Ciência e a Tecnologia* (POCTI/ECO/33089/99 and SFRH/BD/934/2000) and from the British Council and logistical support from *Banco de Portugal*. All errors are my own.
[*] Address: Department of Economics, University of St Andrews, St Andrews, Fife, KY16 9AL, UK. Email: pm31@st-and.ac.uk; Web: www.st-and.ac.uk/~pm31; Phone: +44/0 1334 462423; Fax: +44/0 1334 462444.

1. Introduction

Two striking general trends have characterised the link between education systems and labour markets throughout the last decades. On the one hand, there has been a clear increase in the supply of highly-*schooled* individuals.[1] On the other hand, there has been a considerable increase in the *skills* requirement of some jobs, given the large technological changes that have been taking place during that period.

Although at a first look these two trends suggest that education systems and labour markets are evolving in harmony, one should acknowledge that there is considerable scope for serious mismatches, with relevant economic consequences. Indeed, such mismatches, such as over- and under-education, may have sizeable economic costs: not only the sheer waste of resources placed on unnecessary schooling paths but also the costs inherent to not meeting the economic potential of an economy due to the lack of appropriate human capital.[2]

Previous results in the over-education literature suggest there is indeed a large amount of schooling mismatches. For instance, the meta-analysis by Groot and Maassen van den Brink (2000) points out to an average of 23% and 14% of a country's workers that are over- and under-educated, respectively.

Moreover, there is a large body of evidence documenting the wage impacts of schooling mismatches. When controlling for worker's occupation, the over-educated workers are found to be better off than appropriately allocated workers (the opposite applies to the under-educated). Conversely, if controlling for the worker's schooling, the over-educated are worse off than the appropriately allocated workers (again, the opposite applies to the under-educated).

This latter fact is particularly hard to fit in a human capital framework, where supply side features (namely characteristics such as schooling, experience and tenure) are supposed to explain the wage outcomes of different workers.[3] Oliveira et al. (2000) argue that "it is [difficult] to explain why employers hire under-educated workers at a premium over their market rate, or why over-educated workers accept positions that pay below their market rate".

Several explanations have been put forward to deal with this puzzle and they are surveyed in the next section. However, an explanation that has been ignored so far concerns the role of *unobserved heterogeneity*. In fact, it has been an implicit assumption in this literature that there are no differences between the unobservables of workers with and without good schooling/occupation matches.

However, it may be the case that the over-educated have a weaker endowment of unobservable characteristics that influence wage determination, such as motivation or ability.

---

[1] See OECD (2000) for a thorough analysis of trends in the schooling attainment and other schooling-related variables of the OECD countries' populations.

[2] This issue has been receiving attention by high-level policy makers. In the conclusions of the Stockholm European Union Summit (Council of the European Union, 2001) it is argued that "Improving basic skills, particularly IT and digital skills, is a top priority to make the Union the most competitive and dynamic knowledge-based economy in the world. This priority includes education policies and lifelong learning as well as overcoming the present shortfall in the recruitment of scientific and technical staff." See also Freeman (1978a) for an early study of over-education.

[3] The key references in the human capital literature are Becker (1964) and Mincer (1974).

This difference in endowments would then be picked up by the over-education control variable, thus biasing it downwards. It may be this potential bias which typically leads to the rejection of the Human Capital (competitive) model in the over-education literature.[4]

An important and recent related contribution is Bauer (2002). This author draws on a panel data set and, by estimating the traditional models in the over- and under-education literature, uncovers evidence of *unobserved heterogeneity*. In particular, he finds that, accounting for that, the over- and under-education coefficients become insignificantly different from the required schooling coefficient, as the human capital (supply side) model predicts.[5]

This paper also aims at considering the impact of *unobserved heterogeneity* upon the link between schooling mismatches and earnings. In particular, I draw on a particularly rich panel data set, which includes information on different firms (and all its workers) across time. I then employ fixed effects models, which difference out individual-specific, time-invariant variables, thus implicitly controlling for unobservable heterogeneity (provided it is time invariant and not differently rewarded across time).

With respect to the structure of the paper, I start by surveying the methodology and econometric models considered in the over-education literature. In Section 3, the data set used is described, and in Section 4, the results are presented. Section 5 presents the conclusions.

2. Survey

The over-education literature aims at studying the quality of the match between the worker's skills and the firm's job requirements. Duncan and Hoffman (1981) took the first steps in this field. These authors extended the traditional Mincer models by distinguishing between attained and required schooling.[6] In particular, they decomposed total schooling into required, surplus and deficit schooling and allowed for each component to have different impacts upon workers' earnings, putting forward the following specification:

(1) $$y_i = \beta_0 + \beta_1 req_i + \beta_2 overed_i + \beta_3 undered_i + X_i\beta_4 + u_i \quad (i=1, \ldots, N)$$

In this specification, the dependent variable, "y", corresponds to the logarithm of earnings. The first independent variable, after the constant, is "req", which corresponds to the required level of schooling for the job the worker (to be determined under different methods discussed below). The "overed" regressor corresponds to the years of excess schooling (over-education) of the worker (the difference between his or her attained schooling and that considered in the relevant "req" variable). Finally, "undered" corresponds to the years of deficit schooling (under-education), computed similarly to the "overed" variable. X represents all other control variables (eg, experience).

[4] This rejection arises from the fact that, as against what Becker suggested, not only the workers' supply variables seem to be influencing the workers' earnings: their placement across different jobs also matters.

[5] However, his data set (the German Socio-Economic Panel) includes few observable controls and is of a relatively small size (less than 3,000 individuals). These two factors are likely to influence the results obtained: few individual-specific controls raise the importance of unobserved heterogeneity while a small sample size decreases the precision of the estimated coefficients.

[6] In this work the concepts of "Schooling" and "Education" will be used interchangeably.

This is an attractive specification, given its relatively straightforward estimation and interpretation.[7] It has been labelled as "ORU" specification (from "Over-Education, Required Education, and Under Education") in Hartog (2000).

A striking difference of the "ORU" specification with respect to the Mincer approach is the role played here by the demand side, disregarded in the human capital approach. In this latter model, it is implicitly assumed that only the supply side variables matter, since human capital variables such as schooling, general experience and firm-specific experience should be sufficient to determine worker's earnings.

The "ORU" specification, instead, accommodates additional views of the process of wage determination, as it includes the traditional Mincer specification as a particular case. In equation (1), if $\beta_1 = \beta_2 = -\beta_3$, then one obtains the Mincer equation, given that the level of attained schooling corresponds to the sum of required schooling, excess schooling and (minus) deficit schooling.

However, for cases in which those equalities do not hold, the model allows for demand side variables to play a role, through the role of "required schooling". An extreme case of this situation would be when $\beta_2 = \beta_3 = 0$, which would mean that only the amount of schooling required for the particular job would be valued, regardless of the specific schooling attainment of the worker. This case has been linked to the "Job Competition" or "Job Allocation" models (see Thurow, 1975).

2.1 Definitions of Required Education

An important step in the implementation of this approach is defining "required education". This is probably a debatable concept, as it does not apply to some occupations. Whereas a lawyer, for instance, must necessarily have completed a given number of schooling years, the threshold level for the schooling attainment of a secretary is less clear. Another criticism is that the "required education" concept, as it will be seen below, abstracts from qualitative differences across degrees that exhibit the same quantity of schooling.

Having acknowledged these shortcomings in the over-education literature, there are several methods that have been used, in particular the following three:

1. *Job analysis*. This method involves the study of the schooling requirements of different occupations by experts on the matter. This may be the most rigorous approach, as it provides reliable information on the number of schooling years that each worker is supposed to have, given his or her job. The main drawbacks to this approach involve the limited availability of data sets that include this information and the somewhat subjective nature of the evaluation.

2. *Worker self-assessment*. In this case, it is the worker that describes the educational attainment that he or she considers to be required for the job. Although on the one hand the worker is supposed to know in detail the characteristics of the job, this method obviously involves a considerable amount of subjectivity and may not be comparable across jobs.

[7] Verdugo and Verdugo (1989) have suggested a different specification, which involves extending the Mincer regression (which controls for schooling attainment) with dummies for over- and under-education.

3. *Realised matches*. This index defines the amount of required schooling as some range (say, a one-standard-deviation) around the mean (or median or modal) level of schooling within occupations disaggregated at some level. The over-educated would then be, for instance, those workers whose schooling exceeds the mean level of schooling for their occupation plus one standard deviation whereas the under-educated have a level of schooling below that given by the mean level of schooling for their occupation minus one standard deviation.

The survey and meta-analysis by Groot and Maassen van den Brink (2000) points out that these different measurement methods impact upon the share of schooling mismatches in each country. In particular, the realised matches method leads to the lowest incidence of both over- and under-education, whereas the worker self-assessment and the job analysis methods lead to the highest incidence of over- and under-education, respectively.

2.2 Stylised Facts

Some stylised facts can be outlined from the estimation of this specification across several countries and throughout the last two decades. Firstly, there is evidence of large schooling mismatches, either in terms of over- or under-education, across a large number of countries and time periods. The specific figures appear to be somewhat sensitive to the specific definition of over- and under-education but, in general, they tend to be rather high. As mentioned before, according to Groot and Maassen van den Brink (2000), who analyse 25 studies, the shares of a country's workers that are either over- or under-educated average at 23% and 14%, respectively.[8]

Secondly, the return to required schooling is larger than the return to attained schooling. Moreover, the return to over-education is positive and the return to under-education is negative (controlling for required schooling – or occupation). This means that there is a premium upon correct matching but those who are over-qualified are still better off than their colleagues who hold the same occupation but are appropriately qualified (i.e., have less schooling). Moreover, those workers whose schooling is lower than expected, given their occupation, earn less than the comparison group (their colleagues who hold the same occupation but are appropriately qualified).

Finally, both the return to over-schooling and the absolute value of the return to under-schooling are lower than the return to required schooling.[9] This latter fact is particularly hard to fit in a Human Capital framework, where supply side features are supposed to explain the wage outcomes of different workers.

2.3 Explanations

---

[8] These sizeable shares of schooling mismatch may be a driving force behind the results presented in Martins and Pereira (2004). By estimating Mincer equations using quantile regression, the authors find that returns to education are systematically higher at the top than the bottom of the wage distribution. This dispersion in the returns to education may have serious implications in terms of wage inequality and may be closely linked to school mismatches.

[9] Usually, the absolute value of the return to under-schooling is also lower than the return to over-schooling.

Three main theoretical approaches have been put forward in order to interpret these findings. A first view concerns *searching and matching*. This line of research – see Diamond (1982) and Pissarides (1990) for an outline of the search literature – highlights the frictions associated to the matching process, giving rise to a large number of costs that may hamper an "efficient" matching of workers with jobs and firms.

A second approach stems from the *human capital* theory. Here, over-education may be an interesting investment, if it allows the worker to accumulate skills that would allow him or her to move to a better job later on his career. Tests of this "stepping-stone" view are presented in Sicherman and Galor (1990), Sicherman (1991) and Kiker et al. (1997), who suggest that over-education is indeed due to a trade-off between schooling and other types of human capital.

Finally, *assignment models* have also been put forward to explain the above-mentioned results (see Sattinger, 1993, for a survey). Here, the demand side of the labour market plays an important role in the wage determination process, which arises from the match between heterogeneous workers and heterogeneous jobs. Moreover, these models present a general equilibrium view of the labour market and typically attempt at explaining the economy's distribution of earnings.

However, except for Bauer (2002), to my knowledge no research has addressed the possibility that the stylised facts mentioned before are a product of unobserved heterogeneity. However, given the increasing supply of highly schooled labour, it is likely that new worker cohorts are characterised by an increasingly diverse background.[10]

3. Data

I use the Personnel Records ("Quadros de Pessoal" data set), which is an employer-based survey on both firm and employee characteristics, which has been covering the Portuguese economy since 1982 on an annual basis. This survey is run by the Ministry of Employment, based on a law that makes it compulsory for every Portuguese firm to hand out the required data.

This data involves an extensive set of characteristics concerning the firm, the establishment (if relevant) and the firm's employees. Moreover, since 1991 an individual identifier (which stems from the worker's National Insurance number) has been added to the data. A similar identifier is available for each firm. Furthermore, each set of characteristics of each individual includes a reference to the firm for which the individual is working in each year. By assembling these different pieces of information, it is therefore possible to build a matched employer-employee panel data set.

I draw on a subset of the data set made of firms from the chemical industry for the years of 1991 up until 1995. This sector was chosen because it is perceived to be a modern, skills-intensive industry where the role of schooling is probably particularly important. Moreover, by considering only one specific industry I ensure a high degree of homogeneity between the

[10] See Dolton and Vignoles (1997) and Chevalier (2001) where this point is addressed in particular detail for the case of the recent cohort of graduate students in Britain. See Sloane et al. (1999) with further evidence for Britain.

different job levels and therefore control for any sector-specific heterogeneity that may bias the results.

The sample considered includes approximately 70% of the total number of firms in the industry in each year, amounting to a total of 870 different firms covered throughout the five-year period considered. However, all workers of each firm sampled are present. This structure results in an unbalanced panel, given that workers of firms not sampled in some given year will not be available in such year. More importantly, some workers are likely to change their industry affiliation or their job/labour-market status (i.e. become public servants, self-employed, unemployed or inactive) so that the leave the survey structure.

As to the creation of the schooling mismatch variables, I consider the “realised matches” definition. In particular, I focused on eight aggregate job level categories available in the data set. I then computed the modal schooling for each category (disaggregated into a two-digit classification) and for each one of the five years studied (1991 up until 1995). The over-education variable is then defined as the difference between the individual’s schooling and this modal value, if the difference is positive, or zero, otherwise. The under-education variable is defined in a similar way, i.e. the difference between the worker’s occupation average schooling minus the modal value, if the difference is positive, or zero, otherwise.

I present in Table 1 the descriptive statistics of each group (the over-educated, the under-educated and the appropriately educated), with a total number of observations of approximately 137,000, referring to approximately 44,000 different workers. A first result is that the share of educationally mismatched workers is considerable: approximately 33% are over-educated whereas 17% are under-educated.

Moreover, substantial differences can be noticed among the different sub-groups. In particular, the over-educated have a clearly higher schooling attainment level (almost nine years, against six, for both the required- and under-educated workers). Moreover, the over-educated are younger (born on average in 1958, against 1951 and 1948 for the other two groups, respectively) and have both less tenure and experience (10 and 20 years of tenure and experience against 15, 30 and 18, 33).

Table 2 focuses on the required levels of schooling considered for different job-level/year combinations. It is found that a large majority of such combinations is characterised by a very low level of modal schooling: only four years of education. Moreover, the table shows a great deal of stability in this measure, mostly when assessed along time but also across different job level categories. This means that there is not much variability in the education variables, given that even if workers change their job levels, this will probably not imply changes in their required-, over- or under-education variables.

4. Results

In this section I present some results concerning the impact of schooling mismatches upon wages. My main interest is assessing the role of unobserved heterogeneity in such link. In particular, I test the hypothesis that over- (under-) education is negatively (positively) correlated with unobserved ability. If this were the case, then strong doubts would be cast upon the traditional results of the impact of over- and under-education upon earnings.

Moreover, these unbiased results would also shed light on the contrasting views of the link between schooling and earnings, namely the human capital and the job allocation theories.

A second goal is to evaluate the identification strategy involved in the use of panel data. As mentioned before, I rely on individuals who have their over- or under-education status changed in order to disentangle the wage impacts of unobserved heterogeneity from those of schooling mismatches. Indeed, it is impossible to distinguish between the impact of these two factors for the individuals who preserve their over- or under-education status unaltered across the years.

Concerning the first goal, I draw on the "ORU" specification, mentioned before:

(1) $y_{it} = \beta_0 + \beta_1 req_i + \beta_2 overed_{it} + \beta_3 undered_{it} + X_{it}\beta_4 + v_i + u_{it}$ (i=1, …, N; t=1, …, T(i))

The variables' definition is similar to those in Section 2. The differences are that this is a panel data specification, which gives rise to a subscript for year (t). I also allow for the panel's unbalanced nature, making the total number of years that each individual is covered dependent on the specific individual. Moreover, I introduce an individual-specific effect (which may be assumed to be either random or fixed).

The other control variables, included in X_{it}, are experience and tenure (both in quadratic form), log hours worked, a female dummy, firm ownership (dummies for public or foreign firms), wage bargaining regime (dummies for two different types of wage contracts), and firm size (number of workers). All these variables are time-variant, except for the worker's schooling and gender (and the v_i disturbance).

As mentioned before, the two competing views on wage determination would predict contrasting results under each specification. On the one hand, the Human Capital approach would predict that $\beta_1 = \beta_2 = -\beta_3$. The Job Allocation approach, on the other hand, would result in $\beta_2 = \beta_3 = 0$.

Equation (1) was estimated under different specifications. The first column amounts to the benchmark case within the over-education literature. I find results that fit the stylised facts described before: a return to required schooling of approximately 11%, a pay premium of about 9% per year of excess schooling and a pay penalty of about 7% per year of deficit schooling.[11]

However, when testing for the adequacy of the poolability assumption, by using the standard Breusch-Pagan (1980) Lagrange-Multiplier test, I find a particularly high statistic, that leads to the rejection of the null hypothesis under any reasonable confidence level. This means that disregarding the specific panel nature of the data set is not adequate.

I then consider the alternative Random Effects model (see column 2). By acknowledging an individual-specific effect (and assuming it to be random), I find no clear differences with respect to the previous specification. However, the Hausman test[12] produces a particularly

[11] For an in-depth analysis of Mincer returns to education in Portugal see Pereira and Martins (2001).
[12] This test takes as the null hypothesis the adequacy of the Random Effects model and as the alternative hypothesis that the Fixed Effects model is the correct one.

high statistic, that leads to the rejection of the null hypothesis under any reasonable confidence level.

The final specification is the fixed effects model (see column 3). Here the individual-specific time-invariant variables (schooling and gender in this context) are not identifiable, as the individual-specific dummy absorbs them. Moreover, given that changes in required schooling are identically equal to changes in over-education plus changes in under-education, perfect multicollinearity prevents the estimation of the three coefficients simultaneously. This identity results from the fact that schooling is a time-invariant variable.

Moreover, the results in column 3 show that there is very little variability in both the required- and under-education variables. Given that a fixed-effects model relies only on individual-specific time changes in the variables of interest, the lack of variability outlined above explains the very low coefficients estimated.

Having identified the problem in terms of lack of variability, which prevents the results from a fixed-effects model from being satisfactory, I move on to characterise in more detail the transitions which give rise to the little variability available. In Table 4, I present results from probit analysis of the probability of a worker changing occupation. The first set of columns refer to any changes in the occupational ladder whereas the second refer only to upward changes.

Table 5 presents a probit model of the probability of workers having their over- or under-education status changed. I find a statistically significant positive relationship between the level of over- or under-education upon the probability of the worker decreasing his or her over- or under-education level, respectively.

The results suggest a statistically robust correlation between educational mismatches and changes in the job level ranking, even after controlling for variables such as schooling, tenure and experience. However, it is found that both over- and under-education impact upon mobility positively. This finding, similar to the one by Sicherman (1991), provides mixed support to stepping-stone explanations of the over-education phenomena.

Some further analysis of mobility is presented in Table 6. Here I present results from similar pooled "ORU" equations but focusing on individuals who change their over- or under-education characteristics.[13] Column 1 presents the benchmark case: pooled analysis of all workers, as in Table 3, column 1. Column 2 presents results for the sub-sample of workers who were school-mismatched in the previous period. It is noticeable that this sub-sample produces very similar results to those of the full sample.

Column 3 of Table 5 considers then this same sub-sample but extends the "ORU" specification by including dummies that pick up the effect of lower levels of over- and under-education. The results suggest that those individuals that experienced a decrease in their over-education level are in the following period earning a small premium (2%) when compared to other workers who have the same observable characteristics but faced no changes in their

[13] A similar approach adopted here can be found in the union wage premia literature (see Mellow, 1981, and Jakubson, 1991) and in the inter-industry wage differentials literature (see Gibbons and Katz, 1992, and Krueger and Summers, 1988).

over-education level. Moreover, the similar premium for those whose over-education fell is particularly high (22%).

A clear interpretation of these changes is difficult, as one should take into account the impact of changes in required and over- or under-education upon the worker's earnings. Moreover, the variables that pick up changes in over- or under-education are dummies and therefore disregard the amount of over- or under-education that changed between the two years. However, to the extent that those whose over-education decreased are those who moved up in the job level, the first result suggests that they become paid similarly to their co-workers. The second result suggests that the less under-educated, who probably fell in the job level hierarchy, are being paid more than their co-workers with a stable path.

Another sub-sample is considered in tables 4 and 5: here I address the individuals who will have their over- or under-education status changed in the following period to that in which they are considered. By drawing on similar dummies, referring to changes between the current and the following period, it is found that those individuals whose over-education will decrease are ex-ante earning more than those observably equal but who over-education status will be constant. Moreover, those workers whose under-education will fall (presumably those who will move down the job level hierarchy) are already ex-ante earning less than their co-workers whose under-education status will remain unaltered.

This suggests an element of selectivity in the transitions that underlie the identification of the role of over-education upon wages in a fixed-effects model. Such estimation relies on individuals who are already earning more, when compared to their colleagues similarly over-educated but whose over-education does not change. This selectivity in the transitions process is likely to give rise to downward biased estimates of the impact of over-education on earnings, as it considers as base wages (before over-education changes) those of individuals whose performance is higher, given their observables.

5. Conclusions

This paper addresses the wage implications of schooling mismatches, paying particular attention to the role of unobserved heterogeneity. Indeed, over- and under-educated individuals may not to be a random sample of the population, even after controlling for their observable differences: not accounting for that may lead to biased estimates of the role of over- and under-education upon earnings.

I deal with this issue by drawing on a sample from a large Portuguese matched employer-employee panel. Using the modal definition of schooling mismatches and job levels as the classification group, I find that approximately 33% of the workers are over-educated and 17% are under-educated. Moreover, each group exhibits very different observable characteristics. In particular, the over-educated tend to be younger, more educated, and less tenured.

Regression results for the pooled data are in line with the stylised facts from the literature: required education and over-education are positively related to earnings, whereas under-education influences earnings negatively. Moreover, the return to required education is greater than that of over-education, which is also greater than the (negative) return to under-education.

A new set of results concerns the estimation of this equation on panel data and, in particular, in a fixed-effects model: the lack of variability in the key variables prevents the estimation of their effects. Since there is a large clustering for different years and job levels on a specific schooling attainment level (four years), even individuals who change their job levels cannot be used in many cases for identification purposes.

Other results address the characteristics of those workers whose over- or under-education status changes across time. It is found that these two variables are good predictors of changes in occupation. This provides some support to a stepping-stone interpretation of schooling mismatches.

I also compare individuals whose over- or under-education status changes with other workers. The evidence suggests that, for instance, the over-educated whose over-education levels fall are already better off before that happens. This suggests that the standard estimates of the impact of over-education that rely in such unrepresentative workers are biased.

References

Baltagi, B. H. (1995) "Econometric Analysis of Panel Data", John Wiley.

Bauer, T. (2002) "Educational Mismatch and Wages in Germany", Economics of Education Review, 21, 221-229.

Becker, G. (1964) "Human Capital: A Theoretical Analysis with Special Emphasis on Education", NBER.

Breusch, T. and Pagan, A. (1980) "The LM Test and its Application to Model Specification in Econometrics", Review of Economic Studies, 47, 239-254.

Chevalier, A. (2001) "Measuring Over-Education", mimeo, Centre for the Economics of Education, London School of Economics.

Council of the European Union (2001) "Presidency Conclusions – Stockholm European Council", Council of the European Union, Brussels.

Dolton, P. J. and Vignoles, A. F. (1997) "How Long Did Graduates in the 1980s Take to Get a Graduate Level Job?", mimeo, University of Newcastle.

Duncan, G. J. and Hoffman, S. D. (1981) "The Incidence and Wage Effects of Over-Education", Economics of Education Review, 1, 75-86.

Freeman, R. B. (1978a) "The Overeducated American", Academic Press.

Gibbons, R. and Katz, L. (1992) "Does Unmeasured Ability Explain Inter-Industry Wage Differentials?", Review of Economic Studies, 59, pp. 515-535.

Groot, W. and Maassen van den Brink, H. (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, pp. 131-147.

Hausman, J. A. and Taylor, W. E. (1981) “Panel Data and Unobservable Individual Effects”, Econometrica, 49, pp. 1377-1398.

Hsiao, C. (1986) “Analysis of Panel Data”, Cambridge University Press.

Jakubson, G. (1991) “Estimation and testing of the Union Wage Effect Using Panel Data”, Review of Economic Studies, 58, pp. 971-991.

Kiker, B. F., and Santos, M. C. (1991) “Human Capital and Earnings in Portugal”, Economics of Education Review, 10, pp. 187-203.

Kiker, B. F., Santos, M. C. and de Oliveira, M. (1997) “Over-Education and Under-Education: Evidence for Portugal”, Economics of Education Review, 16, pp. 111-125.

Krueger, A. B. and Summers, L. H. (1988) “Efficiency Wages and the Inter-Industry Wage Structure”, Econometrica, 56, pp. 259-293.

Martins, P. S. and Pereira, P. T. (2004) “Does Education Reduce Wage Inequality? Quantile Regression Evidence From Fifteen Countries”, Labour Economics, forthcoming.

Mellow, W. (1981) “Unionism and Wages: A Longitudinal Analysis”, Review of Economics and Statistics, 63, pp. 43-52.

Mincer, J. (1974) “Schooling, Experience and Earnings”, NBER.

OECD (2000) “Education at a Glance”, OECD, Paris.

de Oliveira, M., Santos, M. C. and Kiker, B. F. (2000) “The Role of Human Capital and Technological Change in Over-Education”, Economics of Education Review, 16, 2, 111-125.

Pereira, P. T. and Martins, P. S. (2001) “Portugal”, Chapter 12, in Harmon, Colm; Ian Walker and Niels Westergard-Nielsen (eds) “Education and Earnings in Europe: A Cross-Country Analysis of Returns to Education”, Edward Elgar.

Pissarides, C. A. (1990) “Equilibrium Unemployment Theory”, Blackwell.

Sattinger, M. (1993) “Assignment Models of the Distribution of Earnings”, Journal of Economic Literature, 31, pp. 851-880.

Sicherman, N. and Galor, O. (1990) “A Theory of Career Mobility”, Journal of Political Economy, 98, pp. 169-192.

Sicherman, N. (1991) “Over-Education in the Labor Market”, Journal of Labor Economics, 9, pp. 101-122.

Sloane, P. J., Battu, H. and Seaman, P. T. (1999) “Over-Education, Under-Education and the British Labour Market”, Applied Economics, 31, pp. 1437-1453.

Thurow, (1975) “Generating Inequality”, Basic Books.

Verdugo, R. and Verdugo, N. (1989) “The Impact of Surplus Schooling on Earnings: Some Additional Findings”, Journal of Human Resources, 24, 629-643.

Table 1 - Descriptive Statistics

By schooling mismatch type, all years pooled

	Over-educated		Required educated		Under-educated	
Variable	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
Schooling	8.80	2.71	5.88	3.48	6.15	3.45
Wages	140,621	96,216	131,976	106,500	167,202	116,666
Female	0.26	0.44	0.30	0.46	0.30	0.46
Birth Year	57.8	9.5	51.0	9.8	47.8	9.9
Tenure	9.6	8.2	14.6	9.3	18.0	9.9
Experience	20.3	9.8	30.1	10.9	33.0	11.1
Hours	174	20	174	21	169	20
Public	9.6%		0.1%		0.2%	
Foreign	20.9%		12.4%		21.5%	
JL1	0.0%		5.2%		15.0%	
JL2	5.7%		4.2%		12.4%	
JL3	15.0%		7.3%		11.1%	
JL4	12.7%		17.0%		24.0%	
JL5	38.2%		37.0%		25.0%	
JL6	17.3%		19.8%		6.8%	
JL7	9.2%		8.0%		2.4%	
JL8	2.1%		1.5%		3.2%	
N. Obs.	45708		67779		23525	
	33.3%		49.4%		17.1%	

Table 2 - Required Schooling Levels

By Job Level and Year, Mode classification

	Year					
Job Level	**91**	**92**	**93**	**94**	**95**	**Description**
10	16	16	16	16	16	Top executives
21	9	9	9	9	9	Intermediary executives
22	16	16	16	16	16	
30	9	4	4	4	9	Supervisors, team leaders
41	11	11	11	11	11	Highly-skilled professionals
42	4	4	4	4	4	
50	4	4	4	4	4	Skilled professionals
51	9	9	9	9	9	
52	4	4	4	4	9	
53	4	4	4	4	4	
54	4	4	4	4	4	
61	4	4	4	4	4	Semi-skilled professionals
62	4	4	4	4	4	
71	4	4	4	4	4	Non-skilled professionals
72	4	4	4	4	4	
80	6	6	6	6	6	Apprentices
81	4	4	4	4	6	

Table 3 - "ORU" specification, Panel Models						
Dependent Variable: Log Monthly Earnings						
	1 (Pooled)		**2 (RE)**		**3 (FE)**	
	Coeff.	**P. Value**	**Coeff.**	**P. Value**	**Coeff.**	**P. Value**
Required Ed.	0.110	0.000	0.114	0.000	0.003	0.000
Over Ed.	0.087	0.000	0.100	0.000	--	
Under Ed.	-0.067	0.000	-0.090	0.000	0.000	0.837
Exp	0.032	0.000	0.036	0.000	0.018	0.000
Exp^2	-0.0004	0.000	-0.0005	0.000	-0.0003	0.000
BP-LM			51,931			
Hausman			17,851			
N. Obs.	136,944		136,944		136,944	
R^2	0.580		0.562		0.079	
Note: Other controls: Tenure (and square), Bargaining Regimes, Log Hours, Female dummy, Public and Foreign Ownership dummies, Regions, Years						

Table 4 - Probability of changing occupation						
Dependent variable: Dummy taking value 1 if worker changes occupation						
	Moving up or down			**Moving up only**		
	Coeff.	**P.Value**	**Mg.Effect**	**Coeff.**	**P.Value**	**MgEffect**
Schooling	-1.384	0.000	-0.249	-1.954	0.000	-7,66e-06
Over-Educ.	0.067	0.000	0.012	0.053	0.000	2,09e-07
Under-Ed.	0.043	0.000	0.008	0.051	0.000	1,99e-07
Wages	0.000	0.000	1,24e-07	0.000	0.061	-7,25e-13
Female	-0.108	0.000	-0.019	0.048	0.012	1,99e-07
Birth Year	-1.379	0.000	-0.248	-1.906	0.000	-7,47e-06
Tenure	-0.005	0.000	-0.001	-0.011	0.000	-4,31e-08
Experience	-1.381	0.000	-0.249	-1.912	0.000	-7,49e-06
N. Obs.	114,597			114,597		
Pseudo R^2	0.0544			0.0933		
Note: Other controls added, as before.						

Table 5 - Probability of decreasing over- or under-education						
Dependent variable: Dummy taking value 1 if worker decreased...						
	Over-Education					
	Coeff.	**P.Value**	**Mg.Effect**	**Coeff.**	**P.Value**	**Mg.Effect**
Schooling	0.035	0.004	0.005	-0.051	0.000	-0.007
Over-Educ.				0.122	0.000	0.016
Wages	0.000	0.000	0.000	0.000	0.000	0.000
Female	0.012	0.671	0.002	0.095	0.001	0.013
Birth Year	-0.036	0.003	-0.005	-0.038	0.002	-0.005
Tenure	0.002	0.378	0.000	0.001	0.557	0.000
Experience	-0.030	0.010	-0.004	-0.033	0.007	-0.004
N. Obs.	28,511			28,511		
Pseudo R^2	0.2228			0.2368		
	Under-Education					
	Coeff.	**P.Value**	**Mg.Effect**	**Coeff.**	**P.Value**	**Mg.Effect**
Schooling	0.012	0.499	0.003	0.042	0.014	0.010
Under-Ed.				0.122	0.000	0.027
Wages	0.000	0.013	0.000	0.000	0.000	0.000
Female	-0.551	0.000	-0.116	-0.461	0.000	-0.095
Birth Year	0.029	0.100	0.007	0.044	0.009	0.010
Tenure	0.004	0.055	0.001	0.002	0.367	0.000
Experience	0.011	0.520	0.003	0.025	0.135	0.006
N. Obs.	13,778			13,778		
Pseudo R^2	0.1105			0.1401		
Note: Other controls added, as before.						

Table 6 - "ORU" specification, Selectivity

Dependent Variable: Log Monthly Earnings

	1		2		3		4		5	
	Coeff.	**P. Value**	**Coeff.**	**P. Value**	**Coeff.**	**P. Value**	**Coeff.**	**P. Value**	**Coeff.**	**P. Value**
Required Ed.	0.110	0.000	0.108	0.000	0.108	0.000	0.106	0.000	0.106	0.000
Over Ed.	0.087	0.000	0.090	0.000	0.095	0.000	0.098	0.000	0.094	0.000
Under Ed.	-0.067	0.000	-0.064	0.000	-0.058	0.000	-0.058	0.000	-0.056	0.000
Exp	0.032	0.000	0.031	0.000	0.031	0.000	0.035	0.000	0.035	0.000
Exp^2	-0.0004	0.000	-0.0004	0.000	-0.0004	0.000	-0.0005	0.000	-0.0005	0.000
Over Ed. Fell	--		--		0.021	0.000	--		0.132	0.000
Under Ed. Fell	--		--		0.218	0.000	--		-0.033	0.000
N. Obs.	136,944		44,164		44,164		44,175		44,175	
R^2	0.580		0.541		0.548		0.568		0.572	

Specifications 2 and 3: Sample refers to period *after* schooling mismatch changed

Specifications 4 and 5: Sample refers to period *before* schooling mismatch changed

Note: Other controls added, as before.