OVEREDUCATION AND LOCAL LABOUR MARKETS IN SPAIN

RAÚL RAMOS* & ESTEBAN SANROMÁ**

*AQR-IREA, University of Barcelona; Facultat d'Economia i Empresa, Universitat de Barcelona, Avda. Diagonal 690, 08034 Barcelona, Spain. E-mail: rramos@ub.edu

**IEB, University of Barcelona; Facultat d'Economia i Empresa, Universitat de Barcelona, Avda. Diagonal 690, 08034 Barcelona, Spain. E-mail: esanroma@ub.edu

Abstract: The objective of this paper is to analyse the influence of individual variables and some spatial mobility characteristics of regional labour markets on overeducation in Spain. To achieve this aim, we use microdata from the Spanish Budget Household Survey to estimate a probit model for the probability of overeducation while accounting for sample selection and the presence of data at different levels (individuals and territory). The obtained results permit us to conclude that the sizes of local labour markets and the possibility of extending the job search to other labour markets by commuting are relevant factors in explaining overeducation in the Spanish labour market. In spite of differences in labour market institutions, our results are very similar to those obtained for other countries, particularly when working with more recent data.

INTRODUCTION AND OBJECTIVES

Recent studies have found that spatial factors can play an important role in explaining educational mismatches in the labour market (Büchel and Battu, 2003; Büchel and Van Ham, 2003; Hensen et al., 2009; Quinn and Rubb, 2011; and Jauhiainen, 2011). In particular, if high-skilled workers have limited capacity to migrate or commute, they will search for jobs in local rather than global labour markets. In this case, the probability of finding jobs matching their education levels decreases significantly because larger labour markets provide a greater number and more diverse employment opportunities than smaller ones. Thus, high-skilled workers facing spatial constraints only have two options: unemployment or accepting a job for which they are overeducated (the worker's education level is higher than the job requires). Therefore, the incidence of overeducation is expected to be higher in small, isolated labour markets and among workers with limited spatial mobility.

The literature has highlighted the role of two different types of factors limiting spatial flexibility: the first are those related to individual characteristics (i.e., mobility constraints are greater for married women, particularly those with children, and for individuals without access to adequate transport options) and, second, those linked to regional characteristics (i.e., rural areas located far from urban labour markets). Regarding the first type of factors, the argument is that overeducation will primarily affect married women because their job searches are restricted to the local labour market in which they live, while their husbands could search for jobs that better match their education levels in a wider labour market. This is known as the "hypothesis of differential overeducation," and it was first proposed by Frank in 1978 for the United States. More recent studies considering this issue have also suggested other relevant individual factors, such as having small children (Büchel and van Ham, 2003), or factors related to mobility decisions, such as access to private means of transportation (Büchel and Battu, 2003; Hensen et al., 2009). With respect to regional characteristics, Büchel and van Ham (2003) also observed that the size of a labour market had a significant effect on the probability of overeducation, while the regional unemployment rate seemed to have no effect on educational mismatches (although it does affect the probability of employment). However, Van Ham (2002) found that overeducation can also occur in large labour markets, such as metropolitan areas, due to increased competition for jobs among highly educated workers; consequently, it is unclear whether there is a clear relationship between labour market size and the incidence of overeducation.

The objective of this paper is to test the influence of individual variables, including the spatial mobility characteristics of the labour force and the characteristics of regional labour markets, on overeducation in the Spanish economy. To achieve this goal, we use microdata from two different waves of the *Encuesta de Presupuestos Familiares* (Household Budget Survey). In particular, we use data from the 1990-91 and the 2006 surveys. Although the primary purpose of this survey is to analyse the consumer expenditures of Spanish families, it also provides information on wages and individual and workplace characteristics. The 1990-91 dataset provides greater territorial detail compared to more recent databases with similar content. In particular, the 1990-91 wave provides information at the provincial level - the

territorial administrative unit corresponding to the NUTS-III level of classification. This territorial unit is sufficiently small to allow us to assume that provinces approximate local labour markets. Moreover, although the 2006 wave only provides information at the *Comunidad Autónoma* level (NUTS-II), it contains more detail on town populations and population density, two key variables regarding the size of local labour markets. Taking these different features of the datasets into account, we conducted our empirical analysis using the two sources

The analysis of the Spanish case in this context is particularly interesting for several reasons. First, previous studies have analysed gender differences in the spatial pattern of overeducation in a context in which both genders equally participate in the labour market. In particular, studies have focused on countries with high female labour market participation and where it seems rational to assume that household location decisions are based on joint decision-making. This is the case in the United States, the United Kingdom, Germany or Finland. However, analyses from countries with low rates of female labour force participation are scarce. In this sense, it is worth noting that, according to the Labour Force Survey, the labour force activity rate of women in Spain at the beginning of the nineteen-nineties was half that of men (35% and 70% for women and men, respectively). The situation has clearly changed over the last twenty years. Although there is still a significant gap, the activity rate of women in 2006 was nearly 50%, much closer to the 70% figure observed for men. A second interesting feature of the Spanish economy in the period under study is its lower levels of interregional mobility compared to other developed countries. According to Huber (2004), the gross interregional migration rate in Spain during the nineteen-nineties was approximately 0.5%, while this figure was considerably higher for other European countries (i.e., 1.3% in Germany and 1.7% in the Netherlands and the United Kingdom). For this reason, analysing regional spatial constraints seems particularly interesting. Finally, an additional contribution of this paper is that we will consider the role of regional transport infrastructure as a determinant of commuting decisions. Between 1990 and 2000, the number of households that had access to a car increased significantly, from 74% to 92% (Matas and Raymond, 2008). In a parallel way, over the last 20 years Spain has embarked on an ambitious motorway construction programme that has provided the country with important transport infrastructure linking its major cities (Holl, 2007). It is interesting to determine whether this investment improved access to larger labour markets between 1990 and 2006 and, as a result, whether transportation infrastructure plays a major role in diminishing regional differences in overeducation.

Our results do not support a hypothesis of differential overeducation, but they confirm that the size of the local labour market and the possibility of searching for a job in a wider area by commuting are important factors in explaining overeducation in the Spanish labour market. However, it is unclear whether transport infrastructure reduces the probability of overeducation.

The rest of the paper is structured as follows. First, the next section provides a brief summary of the literature on the topic. In the third section, the database is described and a measure of overeducation is calculated for the Spanish case. In the fourth section, the applied econometric methodology is explained and the empirical results are presented. Finally, the paper concludes with a summary of the main results.

LITERATURE REVIEW

Although an extensive body of research has analysed overeducation[1] since the seminal contributions of Freeman (1976) and Duncan and Hoffman (1981), only a few recent studies have considered the role of regions. The link between both strands of the literature is the differential overeducation hypothesis. This hypothesis was first proposed by Frank in 1978 for the United States. The objective of this study was to explain wage differences between men and women while accounting for the role of overeducation. Frank (1978) assumes that individuals attempt to maximise their incomes when they search for jobs and minimise overeducation (the difference between their educational levels and the educational requirements of the job). This search process takes place in the global labour market for single individuals –as they can easily migrate- but not for couples. In the latter case, a joint decision must be taken in which the family member with a higher educational level and who works more hours (the individual with greater income is usually the husband in the context of Frank's study) searches for a job in the global labour market that minimises his/her overeducation. However, the other family member (usually the wife) searches for the best possible job given the limitations of the local labour market where the family lives. The wife's ability to minimise her overeducation in the case of an optimal match is purely governed by chance. This spatial restriction would generally imply that the wife's level of overeducation is higher than that of the husband. Moreover, in this case, the wife's level of overeducation will be related to the size of the local labour market because the smaller it is, the smaller the number of available jobs that she can accept. Based on this hypothesis and according to Frank (1978), wage differences by gender would be related to overeducation because overeducation would predominantly affect married women. The empirical evidence obtained by Frank (1978) supports this hypothesis: the smaller the local labour market, the greater the wage differences between genders.

Two decades later and in the context of a resurgence of interest in overeducation, McGoldrick and Robst (1996) directly tested Frank's hypothesis instead of estimating wage equations. Using 1985 data for the United States, they calculated overeducation measures using three different methods and, in three cases, found that married women had higher levels of overeducation than their husbands, but the size of the local labour market did not have any significant effect on this difference. In contrast to Frank (1978), they did not find evidence to support a hypothesis of differential overeducation

Büchel and Battu (2003) open a new research line that extends Frank's model to include the possibility of commuting. Using this new framework, they test the hypothesis of differential overeducation using data for the German economy from 1995 to calculate a subjective measure of overeducation and considering the role of spatial mismatching. Their results seem to confirm the hypothesis because the probability of being overeducated is higher for married women living in rural areas (less than 20,000 inhabitants). According to these authors, the probability of a woman finding a job that matches her educational level not only depends on the size of the local labour market (Frank's original contention) but also on the possibility of accessing other local labour markets by commuting. The probability of being overeducated decreases for higher commuting distances. Moreover, and more relevant to the present

discussion, there is no evidence to support the differential overeducation hypothesis. The results reveal that, controlling for commuting distance, the risk of being overeducated is higher for couples living in rural areas for both members of the couple and not specifically for woman.

Based on these previous results, Van Ham (2002) and Büchel and Van Ham (2003) develop a theoretical framework that relates the presence of overeducation at the individual level (for both men and women) to the availability of job opportunities. Following Simpson (1992), they highlight that an individual searching for a job in a particular local labour market has three options this market lacks an appropriate job: the first option is not to accept the job and continue the search (unemployment); the second option is to accept a job in the local labour market that requires a lower educational level than the individual possesses (overeducation); and the third option is to accept a job in a different local labour market, likely entailing a commuting distance that is greater than desired. The central aspect of the analysis by Büchel and van Ham (2003) is an attempt to discover whether job opportunities in local labour markets (unemployment rates) or commuting (availability of private transport and commuting time) explains the probability of being overeducated. As in Frank (1978), geographical restrictions play a key role in explaining overeducation, but Büchel and Van Ham (2003) extend the effect to all workers (and not only women), and they consider the possibility of extending the job search to other local labour markets through commuting. Their results for the German labour market show that regional variables related to the spatial distribution of jobs explain overeducation. Moreover, increased individual mobility (from owning a car or making a longer commute) permits an increase in the "effective" size of the labour market, which decreases the probability of being overeducated. A similar result is found by Hensen et al. (2009), although they did not explicitly consider gender differences when analysing the relationship between geographic mobility and education-employment mismatch in the Netherlands[2].

More recently, Quin and Rubb (2011) used the Panel Study on Income Dynamics (PSID) for the United States, a longitudinal dataset that makes it possible to control for the timing of migration and its results in terms of the education-employment match. Their results suggest that migration often leads wives to exit full-time paid employment in larger numbers than husbands. In fact, migration tends to lower the level of overeducation for men more robustly than for women, a result supporting the differential overeducation hypothesis. In fact, their analysis suggests that an overeducated wife may be more willing to migrate to improve her husband's career and/or to exit the labour force rather than remain overeducated.

Finally, Jauhiainen (2011) uses a 7% random sample drawn from the Finnish census of 2001 with a regional detail of 85 sub-regions (NUTS-IV) to test the differential overeducation hypothesis. The results demonstrate that living in a large labour market decreases the probability of overeducation for both men and women; consequently, no evidence of differential overeducation is found for Finland.

Empirical evidence on the differential overeducation hypothesis and its relationship with regional characteristics is scarce for Spain. Only García-Serrano and Malo (1997) have tested this hypothesis for the Spanish case, using data from the *Encuesta de Estructura, Conciencia y Biografía de Clase* of 1991. Their results do not completely confirm the differential overeducation hypothesis. They do not find clear evidence of gender effects on educational mismatch, although this variable is statistically significant when interacted

with age or with a variable capturing cohabitation. Therefore, they find that there are differences in Spain in terms of educational mismatch, but this mismatch is not explained by the size of the local labour market. In fact, the size of the local labour market and being a women living in the couple negatively affect the probability of being undereducated, but not the probability of being overeducated.

In light of these contributions, the objective of this paper is to identify the factors that explain overeducation and, in particular, to analyse the effect of geographic factors. The first stage tests differential overeducation hypothesis suggested by Frank (1978), while the second stage examines the effects of different territorial variables, such as local labour market size or the possibilities of commuting. In the Spanish case, the empirical literature on commuting decisions would, a priori, support the differential overeducation hypothesis. Artís et al. (2000) estimate a multinomial logit model to analyse commuting using data for Catalonia from 1991. Their results reveal that the probability of commuting is clearly lower for wives, especially those who are also mothers. Romaní et al. (2001) analyse the interactions between housing location and workplace location, confirming the previous results regarding marital status and motherhood. Casado (2000) obtains similar results using data for *Comunidad Valenciana* from 1991. According to this author, the probability of commuting declines when the number of sons increases. These results are in line with the asymmetric domestic work sharing found in Spain at the time. Moreover, it is possible to imagine that commuting distances would also differ across gendersbecause there is evidence that being the primary income provider in the household increases the probability of commuting and of commuting using private transport (car) (Matas, 1991 and Matas et al., 2009). This evidence regarding differential commuting, combined with the substantial investments in transport infrastructure made over the last twenty years (Holl, 2007), makes the analysis in the paper all the more relevant.

Before providing the results of the econometric analysis, the next section describes the data and provides the results of calculating a measure of overeducation for the Spanish economy using these data.

OVEREDUCATION IN THE SPANISH LABOUR MARKET

The Spanish Household Budget Survey - The estimates presented in this study are based on individual data from two different waves of the *Encuesta de Presupuestos Familiares* (Household Budget Survey) conducted by the INE (the Spanish Institute of Statistics) in 1990/91 and 2006. Although the main objective of these surveys is to analyse Spanish family consumption expenses, these two waves provide individual information on similar characteristics that are of substantial interest for labour microeconomics.

In particular, the 1990-91 wave provides information on individual and family characteristics, their relationships with the labour market and their dwelling characteristics with a greater territorial detail (*provincia*, NUTS-III) than other similar databases. In this paper, we have used information for individuals with ages between 16 and 65, and we have not considered individuals living in Ceuta and Melilla and those working in the army or unclassified occupations. This sample has been divided into two subsamples. The first is formed by the 21,359 employed individuals living in Spain and not engaged in formal studies, and it was used to calculate the incidence of overeducation in the Spanish labour market in 1990/91. The second

subsample is formed by 5492 individuals that had at least a secondary school education (FP2, COU or university degrees) and did not continue studying. Of the total, 4889 of these individuals were employed, while the other 603 were unemployed.

As previously mentioned, the 2006 wave does not provide information at the provincial level (NUTS-III) but at the *Comunidad Autónoma* level (NUTS-II). Although this regional level is most likely not the most appropriate to proxy for the concept of local labour markets, it contains greater detail on some variables related to the territory than the 1990-91 wave. In particular, there is more detailed information on town population size and on population density, two key variables regarding the size of local labour markets. Using the same filters as above (ages between 16 and 65, not living in Ceuta and Melilla and not working in the army or unclassified occupations), there are 23,025 employed individuals living in Spain and not engaged in formal studies. This sample was used to calculate the incidence of overeducation in the Spanish labour market in 2006. The second subsample is formed by qualified individuals and consists of 12,481 employed workers and 1129 unemployed individuals. These different samples are used for the econometric analysis in the following section.

The measurement of the educational mismatch - As the *Encuesta de Presupuestos Familiares* (Household Budget Survey - HBS) does not contain subjective information about overeducation but provides detailed information on educational levels and occupations. We have used statistical methods[3] to obtain our overeducation measure. The HBS contains 14 different schooling levels and ten of them can be applied to individuals older than 16. It also provides detailed information on 81 occupations, a 2-digit disaggregation that is below the optimum (3-digits of occupational classifications). For this reason, it is not possible to apply the corrected mode criterion, as the number of categories where the mode will not contain at least 60% of the total number of individuals is quite high. For this reason, we have used the average criterion.

The results of applying this procedure to employed individuals that were not engaged in formal studies during the 1990-1991 period show that 14.6% were overeducated, while in 2006 this figure was 13.8%. Comparing these results with those of previous studies available for the two considered periods (table 1) reveals that, when the objective method is applied to the Spanish Labour Force Survey (LFS), the results are clearly lower, while the subjective method applied to the *Encuesta de Estructura, Conciencia y Biografía de Clase* (EECBC) provides a higher value. The results of the statistical procedure will be in the middle. The studies applying this method –with the average criterion- to data for 1991 find a similar percentage to that found here: 15.9% for the EECBC and 8.9% for the LFS. The corrected method criterion developed by Oliver and Raymond (2002) also provides a similar value to that found here. For 2006, the comparison yields similar results. Therefore, the results found when using data from the *Encuesta de Presupuestos Familiares* are in line with those found by previous studies on the Spanish economy.

Table 1. *Percentage of overeducated workers in Spain according to different studies*

Authors	Survey	Year	Method	Overeducated (%)
Alba-Ramírez (1993)	Encuesta de Condiciones de Vida y de Trabajo	1985	Subjective (indirect)	17.00%
García Montalvo (1995)	Labour Force Survey (LFS)	1985	Objective	3.70%
García Montalvo (1995)	Labour Force Survey (LFS)	1989	Objective	6.30%
Oliver and Raymond (dirs.) (2002)	Labour Force Survey (LFS)	1990	Statistic (mode)	17.20%
García Serrano and Malo (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Subjective (indirect)	28.40%
García Serrano and Malo (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Subjective (direct)	29.40%
García Montalvo (1995)	Labour Force Survey (LFS)	1991	Statistic (average)	8.90%
Beneito et al. (1996)	Encuesta de Conciencia y Biografía de Clase	1991	Statistic (average)	15.90%
García Montalvo (1995)	Labour Force Survey (LFS)	1993	Objective	7.70%
Murillo et al. (2011)	Structure of Earnings Survey (SES)	1995	Statistic (mode)	35,30%
Own estimates	Household Budget Survey (HBS)	1990-91	Statistic (average)	14.60%
Budría and Moro-Egido (2008)	European Community Household Panel (ECHP)	2001	Subjective (indirect)	34.60% men; 40.10% women
Oliver and Raymond (2002)	Labour Force Survey (LFS)	2001	Statistic (mode)	32.50%
Murillo et al (2011)	Structure of Earnings Survey (SES)	2002	Statistic (mode)	31.88%
Murillo et al (2011)	Structure of Earnings Survey (SES)	2006	Statistic (mode)	37.21%
Nieto and Ramos (2011)	Adult Education Survey	2006	Statistic (average)	19.90%
Nieto and Ramos (2011)	Adult Education Survey	2006	Statistic (mode)	24.10%
Own estimates	Household Budget Survey (HBS)	2006	Statistic (average)	13.80%

Source: Blanco (1997) and own elaboration.

An empirical regularity in the literature on overeducation is that the probability of being overeducated is higher for more educated workers. We have calculated the relevance of overeducation in the samples of employed individuals who have at least completed secondary education[4]. The incidence of overeducation is four percentage points higher for more educated individuals than for the general population in 1990/91, 18.7% and 14.6%, respectively, and the percentages are 17.2% and 13.8%, respectively, in 2006. The results also differ across gender because overeducation is slightly more common for women than men: 19.1% and 18.5% in 1990/91 and 20.9% and 13.9% in 2006 for women and men, respectively. Because unemployment is higher among women, one could imagine that a higher unemployment rate could drive women to accept jobs with educational requirements below their own education levels. Both phenomena (employment opportunities and overeducation) could clearly be interrelated, an aspect that is further developed in the next section.

OVEREDUCATION, LOCAL LABOUR MARKETS AND COMMUTING

Methodology - The objective of this paper is to identify the factors that explain being overeducated, and, in particular, to analyse the possible effects of territorial variables. In particular, we seek to explain the values of a dichotomous variable taking value 1 when the individual is overeducated and 0 otherwise. In this context, multiple linear regression models are not appropriate. It is much more interesting to answer questions such as: What is the probability of being overeducated considering characteristics of the individual, of his/her job or the territory where he/she lives? These types of questions can easily be answered using probit models that make it possible to estimate (using maximum likelihood techniques) the increase in the probability of being overeducated given marginal variations in the explanatory variables.

However, an additional aspect that should be considered prior to the empirical analysis is that it is impossible to know whether an unemployed worker is overeducated because they cannot be assigned to an occupation. However, it is possible that an unemployed worker prefers to extend his/her job search when he/she is offered a job that requires a lower educational level than his/her own. If this possibility is not considered in the empirical analysis, the results could be incorrect. This problem is known in the literature as "selection bias". To solve this problem, we have conducted the analysis using maximum likelihood probit models with sample selection. It is important to note that, to properly identify the model, there must be a variable in the employment

equation that is excluded from the overeducation equation. Finding a variable that does not affect overeducation but affects employment is difficult because most of the variables influence both employment and overeducation. In this paper, we have used marital status. Another methodological aspect that should be taken into account is the presence of the explanatory variables related to individual characteristics and others related to regional characteristics. If the presence of data from different levels is not considered, the inferences obtained from the model could be seriously affected. As Moulton (1986) shows, estimating this type of equation using ordinary least squares of will introduce upward bias in the significance values for this variable. In other words, the statistical significance of the regional variables (i.e., the regional unemployment rate) may not be rejected because of the estimation procedure applied. For this reason, we have applied the correction proposed by White (1984) for arbitrary intra-region correlation.[5]

Results - This section presents the results of estimating different models to identify the determinants of being employed and being overeducated in the Spanish labour market in the two time periods under study. The results of testing the differential overeducation hypothesis and the influence of territorial variables on overeducation in the Spanish economy are shown in tables 2, 3 and 4.

The results of the test of the differential overeducation hypothesis for the Spanish case are shown in table 2 for both the 1990/91 and 2006 data. To test this hypothesis, we have included variables related to individual characteristics and those related to the local labour market, the size of the town where the individual lives and the regional unemployment rate[6]. Before describing the results related to the variables of interest, it is worth noting that educational level and experience have positive and significant effects on the probability of being employed, while being a woman and cohabitating or being male and living with one's parents reduce that probability. Local labour market conditions are also relevant for the probability of being employed because the regional unemployment rate is negative and statistically significant. These results hold for both datasets. The only difference is that the dummy variable associated with having children is only negative and statistically significant at the usual level in the 2006 dataset. The probability of being overeducated increases with education level. This is a common result in the international literature and for the Spanish economy. It is also worth noting that potential experience does not affect the probability of being overeducated. Although perhaps this is not a rigorous test, this result does not confirm the substitutability of education and other forms of human capital postulated by human capital theory[7].

It is also worth mentioning that the parameter capturing the interaction between the two equations is statistically significant at the usual levels in both samples, which can be interpreted as

evidence in favour of the previously mentioned strategy of not accepting an "overeducated" job but waiting for an optimal one in the Spanish labour market. This strategy is consistent with the fact that Spain has powerful family network effects that protect against unemployment and a generous unemployment benefit system (in 1990-91 and relative to other developed countries - OCDE, 1994). However, if unemployment is high and long-lasting, there will be limits to the validity of the strategy and to the significance of this parameter.

Turning to the variables related to the differential overeducation hypothesis, for the 1990/91 data, only a few interactions between gender, marital status and size of the town of residence (taking as a base category married women living in small towns) are statistically significant. The differential overeducation hypothesis would not be rejected if these interactions were significant and negative. This is exactly the case for the 2006 data. Therefore, the results for the 1990/91 dataset make it possible to reject the hypothesis proposed by Frank (1978), a similar result to those obtained by McGoldrick and Robst (1996) and by García Serrano and Malo (1997) for the Spanish case, while the opposite is true for the 2006 dataset.

However, as previously stated, the literature has also suggested that overeducation could be more related to the size of the local labour market (town) than to gender or marital status. This is the hypothesis proposed by Büchel and van Ham (2003). Following these authors, we have analysed the effect of the size of the local labour market and the possibility of commuting to a larger area to search for jobs on the probability of being overeducated. In particular, we have proxied the local labour market size by including in the model the number of inhabitants in the individual's town of residence. The information from the two waves of the Household Budget Survey has been classified into two categories: towns of more of 50,000 inhabitants, or provincial capitals, and the remaining towns. The possibility of geographically expanding the job search to achieve a better match that would avoid overeducation has been considered through the inclusion in the model of two additional variables: one at the individual level, the availability of private transport, and the other at the regional level, the number of road kilometres per car[8].

The results of this analysis are shown in table 3. We find that regional variables clearly affect overeducation and have the expected signs. Living in a small town (and, consequently, having fewer employment opportunities) increases the probability of being overeducated. It is especially interesting that the possibility of searching for a job in a wider area significantly reduces the probability of overeducation. The availability of private transport[9] is statistically significant and has the expected negative sign. The level of infrastructure in a given region (proxied by the number of kilometres of road per vehicle) has negative and significant effects for the 1990/91 dataset[10], but the effect has the opposite sign for the 2006 dataset. This anomalous result, together with the

negative sign on the unemployment rate (which is also counterintuitive) could be explained by the fact that the regional dimension in the 2006 dataset is less disaggregated than that of the 1990/91 dataset. In any case, it is worth mentioning that the coefficient associated with the interaction between gender and marital status is not statistically significant in either of the datasets, once the regional variables are included in the model. This result reinforces the notion that the differential overeducation hypothesis is not confirmed by the Spanish data.

Finally, because the 2006 dataset includes more detailed information on the area of residence than the 1990/91 dataset, we have estimated a new model in which local labour market size has been proxied by three different categories of the number of inhabitants in an individual's town of residence: towns with more than 50,000 inhabitants or provincial capitals (large towns), towns with more than 20,000 inhabitants but fewer than 50,000 (medium towns) and the remaining towns (small towns). The results are shown in the first column of table 4. The results are very similar to those discussed above. The coefficients associated with the dummy variables for living in a medium or in a small town are positive and statistically significant and do not exhibit statistically significant differences from one another. However, as the 2006 dataset also provides information on population density in the area of residence, we have disaggregated the sample into three different subsamples: those living in densely populated areas, those living in areas with medium population density and those living in low density areas. The results are shown in the last three columns of table 4. The most interesting result from this analysis is that town size is relevant in the three areas, although medium size towns seem to be worse off in high and medium density areas, while small towns experience a higher probability of overeducation in low density areas. The availability of private transportation is much more relevant in high-density areas, while the "anomalous" effect of transport infrastructure is only observed in medium- and low-density areas. From our point of view, these results could be interpreted following Van Ham (2002)'s line of reasoning. Van Ham (2002) argues that overeducation can also occur in large labour markets such as metropolitan areas due to increased competition for jobs among highly educated workers. However, if highly educated workers have access to private transportation and a good road network, it could be the case that they choose to commute to less populated areas where competition is less fierce, although the job opportunities will also be less attractive. If this is the case, improvements in the road network will not be beneficial to those living in medium and less populated areas because they will experience greater competition from those commuting from highly populated areas.

Table 2. *Determinants of overeducation and employment (I)*

ML estimation of a probit model with sample selection	**1990/91** Coeff.	**2006** Coeff.
Overeducation probability		
(Married woman living in a small town)		
Married woman living in a big town	-0.060	-0.278***
Single woman living in a big town	-0.041*	-0.464***
Single woman living in a small town	-0.159	-0.282***
Married man living in a big town	-0.081*	-0.502***
Single man living in a big town	-0.105	-0.545***
Married man living in a small town	-0.014	-0.406***
Single man living in a small town	-0.131	-0.389***
Son living with their parents (not living)	-0.033	0.692***
Experience	0.005	-0.011
Experience squared	-0.000	-0.000
Schooling years	0.786***	0.600***
Part-time work (full time work)	0.164	
Regional unemployment rate	-0.007	-0.029
Employment probability		
Woman (man)	-0.288***	-0.476***
Woman*rest family income	0.000	0.000***
Woman with children (woman without children)	0.023	-0.208***
Married (single)	0.261**	0.411***
Married woman (single man)	-0.324***	-0.320***
Son living with their parents (not living)	-0.422***	0.005
Experience	0.072***	0.035**
Experience squared	-0.001**	-0.001**
Schooling years	0.012	0.046***
Regional unemployment rate	-0.018***	-0.004*
ρ	1.021*	2.291***
Wald test of indep. eqns.	3.550*	174.010***
Observations	5492	12324

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3. *Determinants of overeducation and employment (II)*

ML estimation of a probit model with sample selection	1990/91 Coeff.	2006 Coeff.
Overeducation probability		
Woman (man)	-0.053	0.167***
Woman with children (woman without children)	-0.056	-0.060
Married woman (single man)	0.074	0.064
Experience	0.015	-0.074***
Experience squared	-0.000	0.000**
Schooling years	0.778***	0.594***
Part-time work (full time work)	0.223	
Availability of private transport (not availability)	-0.332***	-0.106*
Living in a small town (living in a big town)	0.110**	0.189***
Regional unemployment rate	-0.010	-0.031*
Regional Km roads/car	-0.205*	0.748*
Employment probability		
Woman (man)	-0.267***	-0.455***
Woman*rest family income	-0.000	-0.000***
Woman with children (woman without children)	0.024	-0.215***
Married (single)	0.521***	0.422***
Married woman (single man)	-0.311***	-0.333***
Experience	0.083***	0.034**
Experience squared	-0.001**	-0.000***
Schooling years	0.013**	0.048***
Regional unemployment rate	-0.019***	-0.004*
ρ	4.876*	0.691***
Wald test of indep. eqns.	2.770*	20.670***
Observations	5492	12324

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4. *Determinants of overeducation and employment (III)*

2006	All Coeff.	High density Coeff.	Medium density Coeff.	Low density Coeff.
Overeducation probability				
Woman (man)	0.163***	0.130***	0.222**	0.242**
Woman with children (woman without children)	-0.061	-0.056	0.046	-0.117
Married woman (single man)	0.067	0.070	0.121	0.005
Experience	-0.073***	-0.067***	-0.092***	-0.079**
Experience squared	0.000**	0.000	0.001***	0.001
Schooling years	0.591***	0.572***	0.645***	0.628***
Availability of private transport (not availability)	-0.100*	-0.096*	-0.251	0.019
Living in a medium town (living in a big town)	0.168***	0.238***	0.119**	0.063
Living in a small town (living in a big town)	0.194**	-0.005	0.059	0.209*
Regional unemployment rate	-0.030*	-0.036*	-0.008	-0.046
Regional Km roads/car	0.715*	0.200	1.393*	1.006*
Employment probability				
Woman (man)	-0.446***	-0.396***	-0.405***	-0.631***
Woman*rest family income	-0.000***	-0.000***	-0.000***	-0.000***
Woman with children (woman without children)	-0.217***	-0.247***	-0.162	-0.104
Married (single)	0.423***	0.457***	0.431**	0.301
Married woman (single man)	-0.335***	-0.267**	-0.265	-0.502**
Experience	0.035**	0.039*	0.038	0.029
Experience squared	-0.000***	-0.001**	-0.001	-0.000
Schooling years	0.048***	0.035***	0.069***	0.081***
Regional unemployment rate	-0.004*	0.003	0.003	-0.024*
ρ	2.960***	4.648	2.367***	-12.818**
Wald test of indep. eqns.	13.550***	1.430	38.180***	3.950**
Observations	12324	6756	2676	2892

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In sum, the results permit us to conclude that the probability of being overeducated is partially explained by spatial factors, such as the size of the local labour market or the possibility of "increasing" its size through searching for jobs in more distant labour markets thanks to the availability of private transport, while the role of a good road network is unclear. These results are in line with those obtained for the German labour market by Büchel and Van Ham (2003) and for the United States by Quin and Rubb (2011), and they reinforce the notion that the spatial dimension of the problem is more relevant than the differential overeducation of married women. Moreover, the results can be interpreted to mean that differences in overeducation between men and women could be caused by the difficulties faced by women in commuting due to greater home responsibilities, especially when they have young children.

CONCLUSIONS

The objective of this paper was to test the influence of local labour markets and other territorial and spatial mobility variables on overeducation in the Spanish economy.

To achieve this goal, we used microdata from two waves of the *Encuesta de Presupuestos Familiares* (1990/91 and 2006) and applied the most appropriate econometric techniques (a probit with sample selection and clustered robust standard errors). The results confirmed that the probability of being overeducated increases with education level, but this probability is not related to an individual's potential experience. This result makes it possible to reject the presence of substitutability between different human capital components. A second important result is that, in the Spanish labour market, workers may have preferred not to accept a job in which they would be overeducated and instead continued searching for an optimal job. Particularly for of the territorial variables, the differential overeducation hypothesis proposed by Frank (1978) has been rejected because the risk of overeducation is not greater for married women living in small towns once spatial factors are controlled for. This result is in line with the available evidence for other countries. However, the most relevant result of this study is that overeducation can be at least partially explained by spatial factors, such as the size of the town where an individual lives or the possibility of commuting to a larger labour market due to the availability of private transport that makes commuting easier. If this result is true, the individuals who are more affected by overeducation are not married women but individuals living in small towns who are geographically limited in their job search processes as a result of their low commuting ability. Another interesting result that could be explored in future research is the role of adequate transport infrastructure because it increases the possibility of searching for jobs in larger areas, but it can also increase local competition for qualified jobs.

In spite of the striking differences in labour market institutions, the results for Spain are in line with those obtained for the German labour market by Büchel and Van Ham (2003) and for the United States by Quinn and Rubb (2011), who highlight the relevance of the spatial dimension of overeducation as opposed to the differential overeducation of married women. In fact, possible differences in overeducation obtained

through the interaction of gender and marital status could perhaps be explained by the lower spatial mobility of married women.

NOTES

1. Surveys by Hartog (2000), Groot and Maassen van den Brink (2000a), Rubb (2003) and McGuiness (2006) have summarised the main finding of this literature.

2. Venhorst et al. (2010) also analysed the relationship between ability, field of study and spatial mobility using a micro-dataset on Dutch university and collage graduates. Their findings indicate that university graduates are more spatially mobile than vocational college level graduates and are more able minimise their job mismatch.

3. Hartog (2000) provides an excellent survey of the different methods (objective, subjective and statistical) to calculate educational mismatch and their main advantages and weaknesses. Verdugo and Verdugo (1989) and Kiker et al. (1997) defend the use of the statistical method: the first uses the average while the second uses the mode. Mendes de Oliveira et al. (2000) also recommend this method but using a "corrected mode".

4. Barone and Ortiz (2011), who focus their analysis on European graduates alone, adopt a similar approach.

5. The calculations in the paper have been obtained using STATA's *heckprob* command and the *cluster* option to estimate the variance-covariance matrix (VCE) corresponding to the parameter estimates.

6. The source for the provincial unemployment rate is the Encuesta de Población Activa (Labour Force Survey).

7. Sicherman (1991) for the United States, Mendes de Oliveira et al. (2000) for Portugal and García Serrano and Malo (1996) and Beneito et al. (1996) for Spain are examples of more detailed test of the human capital hypothesis.

8. The source of the number of cars and kilometres of road in each province is the *Anuario Estadístico del INE*.

9. A similar result to that found by Matas et al. (2009) for the Spanish labour market.

10. To take account for the possible endogeneity of the variable related to the availability of private transport, we have constructed a simultaneous equation model formed by the two considered equations and an additional one explaining car ownership at the household level following Bjorner and Leth-Petersen (2005). The results, which are available from the authors on request, are very similar to those presented here.

REFERENCES

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

ARTÍS, M., J. ROMANÍ. & J. SURIÑACH (2000), Determinants of Individual Commuting in Catalonia (1986-91): Theory and Empirical Evidence. *Urban Studies* 37 (8), pp. 1431-1450.

BARONE, C. & L. ORTIZ (2011), Overeducation among European University Graduates: A comparative analysis of its incidence and the importance of higher education differentiation. *Higher Education* 61 (3), pp. 325-337.

BLANCO, J. M. (1997), Comentarios Acerca del Desajuste Educativo en España, *Papeles de Economía Española* 72, pp. 117-128.

BENEITO, P., J. FERRI., M. L. MOLTÓ & E. URIEL (1996), Desajuste Educativo y Formación Laboral Especializada: Efectos sobre los Rendimientos Salariales, *Documento de Trabajo IVIE*, WP-EC 96-11.

BJORNER, T. B. & S. LETH-PETERSEN (2005), Dynamic Models of Car Ownership at the Household Level, *International Journal of Transport Economics* 32 (1), pp. 57-75.

BÜCHEL, F. & H. BATTU (2003), The Theory of Differential Overqualification: Does It Work?. *Scottish Journal of Political Economy* 50 (1), pp. 1-16.

BÜCHEL, F. & M. VAN HAM. (2003), Overeducation, Regional Labour Markets and Spatial Flexibility. *Journal of Urban Economics* 53 (3), pp. 482-493.

BUDRIA, S. & A. MORO-EGIDO (2008), Occupational Mismatch and Wage Inequality: Evidence from Spain. *Economics of Education Review* 27, pp. 332-341.

CASADO, J.M. (2000), Diferencias de Género en los Desplazamientos Cotidianos por Razones Laborales. *Documento de Trabajo IVIE* WP-EC-2000-03.

DUNCAN, G. AND HOFFMAN, S. (1981), The Economic Value of Surplus Education. *Economics of Education Review* 1 (1), pp. 75-86.

FRANK, R. (1978), Why Women Earn Less: The Theory and Estimation of Differential Overqualification. *American Economic Review* 68 (3), pp. 360-373.

FREEMAN, R. B. (1976), *The Overeducated American*. London: Academic Press.

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

GARCÍA SERRANO, C. & M. A. MALO (1996), Desajuste Educativo y Movilidad Laboral en España. *Revista de Economía Aplicada* 4 (11), pp. 105-131.

GARCÍA SERRANO, C. & M. A. MALO (1997), ¿Es Diferente el Desajuste Educativo de las Mujeres?. *Información Comercial Española. Revista de Economía* 760, pp. 117-128.

HARTOG, J. (2000), Over-Education and Earnings: Where are We, Where Should We Go, *Economics of Education Review* 19, pp. 131-147.

HENSEN, M.M., M. R. DE VRIES, F. CÖRVERS (2009), The role of geographic mobility in reducing education-job mismatches in the Netherlands, *Papers in Regional Science* 88 (3), pp. 667-682.

HOLL, A. (2007), Twenty years of accessibility improvements. The case of the Spanish motorway building programme. *Journal of Transport Geography* 15 (4), pp. 286-297.

HUBER, P. (2004), Inter-regional mobility in Europe: a note on the cross-country evidence. *Applied Economics Letters* 11 (10), pp. 619-624.

JAUHIAINEN, S. (2011), Overeducation in the Finnish regional labour markets. *Papers in Regional Science* 90 (3), pp. 573–588.

KIKER, B., M. SANTOS & M. MENDES DE OLIVEIRA (1997), Overeducation and Undereducation: Evidence for Portugal. *Economics of Education Review* 16 (2), pp. 111-125.

MATAS, A. (1991), La Demanda de Transporte Urbano: Un Análisis de las Elasticidades y Valoraciones del Tiempo. *Investigaciones Económicas* XV (2), pp. 249-267.

MATAS, A., & J. L. RAYMOND (2008), Changes in the structure of car ownership in Spain. *Transportation Research Part A* 42, pp. 187–202.

MATAS, A., J. L. RAYMOND & J. L. ROIG (2009), Car ownership and access to jobs in Spain, *Transportation Research Part A: Policy and Practice* 43 (6), pp. 607-617.

MCGOLDRICK, K. & J. ROBST (1996), Gender Differences in Overeducation: A Test of the Theory of Differential Overeducation. *American Economic Review* 86, pp. 280-285.

MCGUINNESS, S. (2006), Overeducation in the labour market, *Journal of Economic Surveys* 20 (3), pp. 387-418.

MENDES DE OLIVEIRA, M. M. SANTOS & B. KIKER (2000), The Role of Human Capital and Technological Change in Overeducation. *Economics of Education Review* 19, pp. 199-206.

MOULTON, B. R. (1986), Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics* 32, pp. 385-397.

MURILLO, I., M. RAHONA & M. SALINAS (2011), Effects of Educational Mismatch on Private Returns to Education: An Analysis of the Spanish Case (1995–2006). *Journal of Policy Modelling* doi:10.1016/j.jpolmod.2011.07.012

OLIVER, J. & J. L. RAYMOND (dirs.) (2002), Educación Formal y Demanda de Cualificación de la Mano de Obra en España. Una Visión a Largo Plazo, *Centre d'Economia Industrial Document d'Economia Industrial* N. 14.

QUINN, M.A. & S. RUBB (2011), Spouse Overeducation and Family Migration: Evidence from the US. *Journal of Family and Economic Issues* 32 (1), pp. 36-45.

RAMOS, R. & S. NIETO (2010), Sobreeducación, Educación no Formal y Salarios: Evidencia para España. Fundación de las Cajas de Ahorro, Documento de Trabajo 577/2010.

ROMANÍ, J., M. ARTIS & J. SURIÑACH (2003), Are Commuting and Residential Mobility Decisions Simultaneous? The Case of Catalonia (Spain). *Regional Studies* 37 (8), pp. 813-826.

RUBB, S. (2003). Overeducation in the labor market: A comment and re-analysis of a meta-analysis. *Economics of Education Review* 22 (6), pp. 621-629.

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

SIMPSON, W. (1992), *Urban Structure and the Labour Market: Worker Mobility, Commuting and Underemployment in Cities*. Oxford: Clarendon Press.

VAN HAM, M. (2002), *Job access, workplace mobility and occupational achievement*. Ph. D. dissertation. Faculty of Geosciences, Utrecht University. Available at http://maartenvanham.nl/vanham2002.pdf

VENHORST, V., J. VAN DIJK & L. VAN WISSEN (2010), Do the best graduates leave the peripheral areas of the Netherlands?, *Tijdschrift voor Economische en Sociale Geografie* 101 (5), pp. 521-537.

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

WHITE, H. (1984), *Asymptotic Theory for Econometricians*, Academic Press, New York.