Human Capital and Regional Wage Gaps.

Enrique López-Bazo♂
Elisabet Motellón♀

♂ AQR—IREA, Universitat de Barcelona, Avda Diagonal 690
Tel: +34 93 4037041FAX: +34 93 4021821
Email: elopez@ub.edu

♀ Universitat Oberta de Catalunya, Av. Tibidabo 39-43
Tel: + 34 932537501 FAX: +34 934176495
Email: emotellon@uoc.edu
&
AQR—IREA, Universitat de Barcelona
Tel: + 34 93 4021010 FAX: +34 93 4021821
Email: emotellon@ub.edu

Abstract: This paper uses micro-level data to analyse the effect of human capital on regional wage differentials. The results for the set of Spanish regions confirm that they differ in the endowment of human capital, but also that the return that individuals obtain from it varies sharply across regions. Regional heterogeneity in returns is especially intense in the case of education, particularly when considering its effect on the employability of individuals. These differences in endowment and, especially, in returns to human capital, account for a significant proportion of regional wage gaps.

Keywords: Education, Experience, Regional disparities, Returns to human capital, Wage gap decomposition

JEL: C24, J31, R11, R23

1. INTRODUCTION

The effect of human capital – an intangible asset embodied in individuals – on regional growth and development has been examined by regional scientists and economists in recent decades. The assumption has been that the human capital endowment of a regional economy is an essential element in explaining its level of development and long-run economic growth. Besides its effect as an additional factor of production, it has been argued that human capital allows and encourages the generation and adoption of technological innovations that improve productivity. Almost all the studies to date have used aggregate data for a set of regions, and so the key variables considered have been the average of the measure used to proxy for the endowment of human capital (e.g. average years of schooling or the share of population with a certain educational attainment) in each region and some measure of aggregate economic activity, such as income or output per capita. In addition, previous studies have only considered the possibility that regional differences in levels of development and growth are due to different human capital endowments across regions (Rodríguez-Pose and Vilalta-Bufí, 2005; Di Liberto, 2008; López-Bazo and Moreno, 2008; Bronzini and Piselli, 2009). That is, no attention has been paid to the possibility that regional heterogeneity in the effect of human capital may be the cause of some of the economic disparities observed across regions. This paper argues that regions may differ in both the endowment and the return to human capital accumulated by individuals. Accordingly, both should be considered when explaining regional differences in levels of economic activity.

To complement the evidence obtained by using aggregate data, this paper proposes the use of micro-data at the regional level. Micro-data provide additional evidence on the effect of human capital in explaining regional disparities and, in turn, a more appropriate control of regional differences in the distribution of individuals' characteristics. In particular, the use of individual data makes it possible to quantify the degree of regional differences in human capital endowment and also to measure its specific effect in each region, that is, to check whether the regions are also heterogeneous in the returns they obtain from human capital investments made by individuals. This has obvious implications for assessing policies designed to increase human capital endowment in order to promote growth in the less developed regions, as the effectiveness of such policies largely depends on the particular effect that human capital has in each region.

The use of information at the individual level allows a consideration of two different effects of human capital on regional economic performance. The first is the immediate effect on productivity from those in employment. The second is an indirect effect that is likely to occur through the increased employability of individuals endowed with a certain level of human capital. Studies using aggregate regional data have focused only on the first of these effects, although there is evidence to support a positive effect of human capital on labour market participation and a negative influence on the likelihood and duration of episodes of unemployment. Our hypothesis is that the two types of effect may differ across regions, thus contributing to regional disparities.

Using reliable individual data on wages obtained from a representative survey for each Spanish region, this paper assesses the effect of human capital within the framework of a Mincerian wage equation. Under the human capital theory, the higher a worker's human capital endowment, the higher the wage she will earn, since it is assumed that education and experience (the two traditional components of individuals' human capital) have a positive effect on her productivity.[1] Within this context, we analyse the contribution of human capital to regional wage gaps, the hypotheses being that i) in addition to the effect associated with regional differences in human capital endowment, heterogeneity in terms of its return across regions may play a key role in explaining regional wage gaps, and ii) there is a direct effect of human capital, since it affects productivity of employees, and an indirect effect, by increasing the employability of all individuals. Aggregating over the individuals in a given region, this means that human capital stimulates aggregate productivity and the employment rate, thus contributing to increasing regional income per capita.

From a methodological point of view, the paper provides a framework for assessing regional differences in the conditional (being in employment) and the unconditional returns to education and experience. In a second step, it proposes a detailed decomposition of regional wage gaps to isolate the particular contribution of individuals' human capital. The approach followed here has been common practice for decomposing wage gaps across different groups of workers (e.g. gender or racial gap) in the labour market literature. But its application to the

[1] An alternative is to estimate the effect of human capital on firms' productivity. However, the lack of firm-level data from a representative survey for each Spanish region prevented us from considering this approach. In any case, under well-known assumptions, the marginal productivity explanation of wage determination establishes the link between wages and productivity. The assessment of the return to human capital based on the estimation of a wage equation is standard in the labour market literature.

analysis of wage differentials across regions has been limited so far, and constrained to models that do not control for individuals' decision to participate in the labour market (Reilly, 1991; García and Molina, 2002). The results for the set of Spanish regions confirm differences in terms of human capital endowment, and also in the return that individuals obtain in each region. Regional heterogeneity in returns is especially intense in the case of education, particularly when they incorporate the indirect effect. The decomposition of the wage gap between each region and the rest of the country shows that these differences in the endowment and in the returns to human capital account for a significant portion of the gap.

The rest of the paper is organized as follows. The next section introduces the dataset and discusses the results of the descriptive analysis. The empirical wage model and the derivation of the returns to the components of human capital are sketched in Section 3, which also discusses the results obtained for the set of Spanish regions. Section 4 presents the method proposed to obtain the detailed decomposition of the regional wage gaps and discusses the results of the contribution of human capital. Finally, section 5 concludes.

2. DATASET AND DESCRIPTIVE ANALYSIS

This paper uses the micro-data from the Spanish sample of the European Community Household Panel (ECHP).[2] The ECHP is a standardized survey conducted in the Member States of the European Union under the auspices of the Statistical Office of the European Communities (EUROSTAT). The survey involved annual interviewing of a representative panel of households and individuals in each country. The analysis in this paper exploits the 2000 extended sample of the ECPH because it was specifically designed for cross-sectional studies and above all because it is the only wave that provides representative samples at the NUTS II regional level in Spain. NUTS is the French acronym for Nomenclature of Territorial Units for Statistics, a hierarchical classification established by EUROSTAT which provides comparable regional breakdowns of EU Member States. In Spain, the NUTS II

[2] The ECHP has frequently been used in wage studies for the Spanish labour market and for other EU Member States (Montuenga et al, 2003; Rodríguez-Pose and Vilalta-Bufí, 2005; García-Pérez and Jimeno, 2007). Although the Earnings Structure Survey (a dataset also produced in the EU countries under the auspices of EUROSTAT) contains the most complete information on wages workers, jobs and firms' characteristics, it does not provide information on the non-employed. This prevents us from controlling for sample selection and computing the indirect effects of human capital on wages through its effect on the probability of employment, which is one of the objectives of this paper.

regions correspond to the 17 Autonomous Communities, which are historical geographical and administrative regions with a high level of political and financial autonomy.[3] The ECHP offers detailed information on the personal characteristics of the individuals, including the particularities of the household, as well as on the labour conditions of those employed. For the analysis of the effect of human capital on regional wages, the sample of individuals between 16 and 65 years in all the Spanish regions, except for the two city-regions in the north of Africa (Ceuta and Melilla), has been selected.[4]

A first insight into the amount of regional wage differentials in Spain is obtained from the simple description of the sample in Table 1, which shows the average gross hourly wage, its standard deviation and the number of workers contained in the sample for each one of the regions and for Spain as a whole. Large differences in average wages across regions are observed. For instance, the average wage in Extremadura, the region with the lowest wage level, was only 69.75% of the average wage in the Basque Country, the region with the highest. And the ratio between the top five regions and the five bottom regions is 1.29. This evidence confirms that the amount of regional wage disparities is of the same order of magnitude as those existing in other key economic variables such as income per capita and labour productivity.

In order to control for the effect of price differentials, an estimate of the relative level of regional prices has been used to compute real wages in each region.[5] The average and standard deviation of real wages are shown in the last two columns in Table 1. Taking account of price differentials causes some changes in the ranking of regions, the most significant case being Extremadura, which moves from bottom to eighth place. Additionally, wage differentials are somewhat lower in real terms. For instance, the average real wage in the five bottom regions increases by around 2% due to their lower relative prices, whereas the average

[3] The regional representativeness of the sample for the entire panel of the ECHP is only guaranteed at the NUTS I level, which corresponds in Spain to an artificial grouping of regions based on geographical criteria alone.

[4] Individuals working less than 15 hours a day were removed from the sample, given that in this case the ECHP does not provide information on some variables that are important for our analysis (e.g. tenure).

[5] This information was kindly provided by the Catalan Institute for Statistics (IDESCAT), which estimates the parity power standards for the 17 Spanish regions from the aggregate Spanish figures used by the Statistical Office of the EU, EUROSTAT, to produce a data net of the cost of living differences across the Member States. Note that, given the common currency for the spatial units under analysis, parity power standards only account for differences in the cost of living.

in the five upper regions falls by the same percentage as a result of their higher prices. However, most of the regional disparities remain after controlling for differences in prices across regions: for instance, the average real wage in Murcia (the region with the lowest value) is still under 75% of the real wage in Madrid, the region with the highest level in real terms.

Real wages may differ between regions because of what is known as the composition effect, that is to say, because workers' characteristics differ across regions.[6] In this case, the real wage paid to each class of workers should be interregionally invariant, and wage differentials would be merely an illusion caused by the failure to distinguish between types of labour (Farber and Newman, 1989). A simple look at the amount of regional differences in workers' characteristics in the sample can be obtained from Table 2, which shows the average value for the characteristics observed in the sample for the whole of Spain, and for the two regions with the highest (Madrid) and lowest (Murcia) average real wage.[7] In each case, the figures refer both to the sample of employees and non-employees (unemployed workers and non-participants). Focusing on the measures of human capital, the results reveal notable differences in education (measured by years of schooling) and in tenure between the two regions. On average, employees in Madrid spent three years longer at school than those in Murcia; the difference is not so high among non-employees, but it remains non-negligible (more than 1.2 years). As regards tenure, most of the differences correspond to the categories of less than one year and more than 15 years. This is to do with regional differences in the number of fixed-term contracts; which is much higher in Murcia than in Madrid (for a further discussion of this issue, see Motellón, 2008). In contrast, there do not seem to be significant differences across regions in labour market experience.

Table 2 shows differences between regions for other individual and household characteristics, such as gender, age and household composition, for both employees and non-employees. Therefore, wages may differ across regions because regions have different human capital endowments and because of other characteristics that are believed to affect wages directly and

[6] It can be argued that jobs' and firms' characteristics also differ across regions. And as far as wages vary within these characteristics, the composition effect should include them as well. However, here the focus is on individuals' characteristics, given our interest in the effect of human capital. In any case, a great deal of the wage variability associated with different jobs and firms is likely to be captured by differences in workers' human capital if there is a process of sorting across jobs and firms depending on the endowment of human capital.

[7] Results for the 17 regions are not reported here to save space, though they are available upon request.

indirectly, through the probability of employment, but also because of regional differences in the return to human capital and in the price of other characteristics.

This seems to be supported by the wage differences observed within categories of levels of schooling, tenure and experience, as reported in Table 3. This table shows the average real wage for the sample of workers in each of the categories of the human capital variables, for Spain as a whole and for the regions with the highest and lowest average real wages. Observe that the wage gap between Madrid and Murcia at each level of schooling decreases somewhat, although the average wage in Murcia was still some 20% lower. The only exception is the regional gap for workers with a university degree, in which case the average wage in Murcia was 92% of that in Madrid. The cases of tenure and experience are quite similar, as the regional gap within categories decreases only marginally (the wage in Murcia being between 70% and 80% of that in Madrid for most of the categories).

Taking this preliminary descriptive evidence into consideration, our hypothesis is that not only the endowments but also the returns to human capital vary across regions, thus contributing to wage differentials, both directly and indirectly through the impact that human capital has on the probability of employment. The next section presents results for the estimates of direct and indirect effects of human capital obtained when conditioned to other factors that are also likely to affect the wage earned by each worker. The estimates of the returns to schooling, tenure and experience obtained for each Spanish region will allow us to check for the regional heterogeneity in the returns to human capital.

3. REGIONAL RETURNS TO HUMAN CAPITAL

3.1. Empirical framework

The framework for the empirical analysis is a model in which the wage for an individual i in region r is given by:

$$W_{ir} = X_{ir}\beta_r + \varepsilon_{ir} \tag{1}$$

$$C^*_{ir} = Z_{ir}\gamma_r + \upsilon_{ir} \tag{2}$$

where W_{ir} is the log of the wage of individual i in region r, X_{ir} denotes the set of characteristics that affect the wage of this individual in a direct way (education, experience, tenure, and gender), and β_r is the vector of prices or returns associated with the characteristics.[8] C^*_{ir} is a latent and unobservable process that assigns the individual i in region r to the sample of employees or to the sample of non-employees, Z_{ir} being the vector of observations for characteristics that determine the process of selection (education, gender, age, marital status, chronic disease, proxies for household composition, and household income other than the wage of the individual)[9] and γ_r the corresponding parameters. ε_{ir} y υ_{ir} are i.i.d errors following a bivariate normal distribution $(0, 0, \sigma_{\varepsilon_r}, \sigma_{\upsilon_r}, \rho_r)$, with ρ_r the correlation coefficient for both error terms in region r.

Only the result of the selection process in (2) is observed, the indicator variable C_{ir}, that equals 1 when $C^*_{ir}>0$, and 0 otherwise. Then, the probability of employment (selection) of individual i in region r is given by:

$$C_{ir} = \mathrm{Prob}\left(C^*_{ir} > 0\right) = \mathrm{Prob}\left(\upsilon_{ir} > -Z_{ir}\gamma_r\right) = \Phi\left(Z_{ir}\gamma_r\right) \tag{3}$$

where $\Phi(\cdot)$ is the standard normal distribution function.

Estimates of returns based on the wage equation in (1), leaving aside the selection equation in (2), are biased and inconsistent if $\rho_r \neq 0$. Consistent estimates can be obtained by maximum likelihood considering the information from the two equations or, alternatively, by applying the two-step method proposed in Heckman (1976). The Heckit method includes the inverse Mills ratio in the wage equation as an additional regressor to obtain wages conditional on being employed:

$$W_{ir} \mid C^*_{ir} > 0 = X_{ir}\beta_r + \theta_r\lambda_{ir} + \varepsilon_{ir} \tag{4}$$

where

$$\lambda_{ir} = \frac{\phi(Z_{ir}\gamma_r)}{\Phi(Z_{ir}\gamma_r)} \tag{5}$$

is the inverse Mills ratio for individual i in region r computed from the probabilistic model in (3), and $\theta_r = \rho_r\sigma_{\varepsilon_r}$ is the coefficient that measures its effect on wages.

[8] Note that, as is usual in this type of analysis, a simple specification of the Mincerian wage equation is used to obtain a better insight into the global effects of the human capital variables on wages (see Pereira and Silva, 2004).

[9] The full list of characteristics included in X and in Z is shown in Table 2. See section 3.2 for further details.

From the specification of the model of wage determination in (1) and (2), and the one for conditional wages in (4), different types of returns to characteristics can be defined.[10] In the case of education – S – the conditional return is defined as:

$$\mathrm{CRS}_{ir} \equiv \partial E\left[W_{ir} \mid C^*_{ir} > 0\right] / \partial S_{ir} = \beta^S_r - \theta_r \gamma^S_r \delta_i \tag{6}$$

where $E\left[W_{ir} \mid C^*_{ir} > 0\right] = X_{ir}\beta_r + \theta_r \lambda_{ir}$ and $\delta_i = \left(Z_{ir}\gamma_r + \lambda_{ir}\right)\lambda_{ir}$. Then, CRS_{ir} is the marginal effect of S_{ir} on the conditional expected value of W_{ir}. The second term is the correction that takes into account that only the effect of S_{ir} on W_{ir} for employed individuals should be considered. That is to say, CRS_{ir} is a measure of the effect that a year of education has on the wage received by employees. Notice that the conditional return to education will be different for each individual in each region, as it depends on the regional coefficients β^S_r, θ_r, and γ_r, and on the value of δ_i. As is usual in these cases, the conditional return to education for each region r – CRS_r – will be computed as the average for the sample of employees in that region.

In addition, the expected value of the wage earned by a randomly selected individual from the entire population (employees and non-employees) is of interest as well:

$$E\left[w_{ir}\right] = \Phi\left(Z_{ir}\gamma_r\right) E\left[w_{ir} \mid C^*_{ir} > 0\right] = \Phi\left(Z_{ir}\gamma_r\right) \exp\left(X_{ir}\beta_r + \theta_r \lambda_{ir} + 0.5\sigma^2_{\varepsilon_r}\right) \tag{7}$$

where w_{ir} is the wage level of individual i in region r. That is, for any individual the unconditional expected wage is the one obtained in the case of being employed, multiplied by the probability of being employed. The marginal effect of education on the unconditional expectation in (7) is then defined as the unconditional return to education (provided that the function is evaluated at a point with $E\left[w_{ir}\right] \neq 0$):

$$\mathrm{URS}_{ir} \equiv \frac{\partial E\left[w_{ir}\right] / E\left[w_{ir}\right]}{\partial S_{ir}} = \frac{\partial \ln E\left[w_{ir}\right]}{\partial S_{ir}} = \beta^S_r - \theta_r \gamma^S_r \delta_i + \gamma^S_r \lambda_i = \mathrm{CRS}_{ir} + \gamma^S_r \lambda_i \tag{8}$$

[10] See Greene (2003) and Cameron and Trivedi (2005) for the derivation of the expressions and the discussion of these marginal effects. Hoffmann and Kassouf (2005) and Arrazola and De Hevia (2008) used these expressions to compute different types of returns to education.

The second term in the unconditional return in (8) reflects the effect that education has on the probability of employment, which is an indirect effect on wages. As this effect is likely to be positive (more education will decrease the episodes of unemployment and non-participation), the URS_{ir} is expected to be higher than the CRS_{ir}. As stressed in Arrazola and De Hevia (2008) individuals take this indirect effect into account when they decide on their investment in education. As in the case of the conditional return, URS_{ir} depends on regional coefficients and on individual values for the characteristics that determine the process of participation, Z_{ir}. Accordingly, the unconditional return to education for each region r – URS_r – will be computed as the average for the total sample of individuals (employees and non-employees) in that region.

As for the other two components of human capital, experience and tenure, note that they are not included in the list of determinants of the probability of employment. As a consequence, they only exert a direct influence on wages through their inclusion in the wage equation. This means that the unconditional effects of these characteristics equal the conditional ones, which are simply a function of the corresponding elements in the vector of coefficients of the wage equation, β.[11]

3.2. Results

The conditional and unconditional returns defined above were computed based on the estimation of the coefficients in the empirical wage model defined by (1) and (2). As already indicated, a simple specification for the wage equation was used to fully account for the effects of the human capital variables. It includes the number of years of schooling, the years of experience and its square, a set of dummies that account for tenure, and the gender of the individual. As for the participation equation in (2), in addition to the measure of education, it includes proxies for the individual and family characteristics that are supposed to affect the chance of being employed: the individual's gender, age, and marital status, presence of chronic disease, the household income other than the wage earned by the individual, and variables of household composition such as its size, the number of children under 15 years, and the presence of children under 6 years.

[11] As usual in the specification of wage equations, a quadratic form is used for experience ($\beta_{EXP}\cdot EXP_{ir} + \beta_{EXP2}\cdot EXP^2_{ir}$). As a result, the return to experience (conditional and unconditional) is $\beta_{EXP}+2\cdot\beta_{EXP2}\cdot EXP_{ir}$. In the case of tenure, its return will be measured by the estimation of the coefficients of each of its categories.

An instrumental variables estimator (IV) was used to avoid the bias of the traditional estimates due to the likely endogeneity of education. Suitable instruments should capture exogenous factors that affect the choice of the individuals' degree of education but not their current wages. Immediate information on variables of this kind (such as family background and ability) is not readily available from surveys like the one used in this study. So we follow the suggestion made in the recent related literature and use as instruments variables that reflect whether the education of the individual was affected by profound changes in the educational system and by extraordinary historical events such as a war (see for instance Harmon and Walker, 1995; Ichino and Winter-Ebmer, 1999 and 2004; Arrazola et al, 2003). Specifically, a dummy variable was defined to account for the effect of the change in the regulation of the Spanish educational system brought in by the 1970 General Education Act, which established free, compulsory education for children between 6 and 14 years old. The instrument is a dummy variable that takes a value of 1 for individuals aged 6 or under in 1971, that is, members of the sample whose period of schooling was affected by the reform. An instrument related to the Spanish Civil War (1936 to 1939) was also defined to capture its effects on individuals who were of school age during that period; in this case the corresponding dummy variable takes a value of 1 for individuals born in or before 1945. In addition, following the suggestion in Wooldridge (2002), the variables in Z, that is, the ones that affect the probability of employment, were included in the list of potential instruments for education in the wage equation.[12]

IV estimates for the parameters of the wage system in (1) and (2) were obtained for each region and for Spain as a whole. They are not shown here for reasons of space, although some comments are in order. First, the coefficients in both equations were jointly significant in all cases, particularly for the human capital variables. For all the regions, education increases the wage earned and the probability of receiving a wage. Experience and tenure also exert a significant positive effect on wages. Second, the coefficient associated with the inverse Mills ratio, θ_r, was statistically significant for 9 out of the 17 regions and for Spain as a whole. It was positive in all cases excepting two regions, in which it was not significant. This means that, in general, shocks that increase the probability of employment also increase the expected

[12] Several statistics of the validity of instruments and the Sargan test of over-identification (to check for exogeneity of the set of instruments) were obtained for the IV estimation of each region and Spain as a whole. The final set of instruments used in each case (which always includes the ones for the change in the educational system and the Civil War) was defined to fulfil both criteria. Details of these results are available upon request.

wage of employees. This shows that the expected wage of employed individuals is higher than that of individuals selected randomly from the entire population.

Tables 4 and 5 reproduce the returns to the different types of human capital computed using the estimates above. As for the returns to education, the first and second columns of results in Table 4 show the conditional and unconditional returns. Remember that the conditional return for each region was computed with the sample of employees, whereas the unconditional one was calculated with the whole sample (employees, unemployed and non-participants). Both types of returns were statistically significant at 1% in Spain and in all regions (for this reason, asterisks are not included alongside the figures in Table 4). The conditional return to education in the entire country was above 6%, which means that an additional year of education increased the expected wage of those actually earning a wage by more than 6%. But this figure for the country as a whole hides significant regional heterogeneity in the conditional effect of education. The conditional return in Cantabria, the highest, almost doubled that in Madrid, the lowest. Among the regions with the highest conditional return are some of the traditionally less developed regions, such as Galicia, Murcia, Castile-León, Castile-La Mancha, and Extremadura. These are also among the regions with the lowest endowments of education. In contrast, the return was below the country average in the most advanced regions, which are the ones with the highest endowment of that type of human capital (such as Madrid, Catalonia, Valencia and the Basque Country).

In view of the positive influence of education on the probability of employment and the positive sign of the estimate of θ_r, the unconditional return to education in Spain as a whole was far above the conditional return. An increase of one year of schooling represented an increase of more than 16% in the expected wage of an individual randomly drawn from the Spanish active population. This result confirms the importance of considering the indirect effect of education when analysing its connection with wage expectations. Actually, the estimate for Spain suggests that the second term of the unconditional return defined in (8) – the indirect effect — is far larger than the direct effect of education on employees' productivity. The same argument applies to almost all the regions under analysis, although once again the results for the estimates of the unconditional returns at the regional level confirm our hypothesis of the strong spatial heterogeneity in the effect of education. The unconditional return in Navarre (12.4%) is half that in Extremadura, which is as high as 24.5%. And regardless of some changes in the ranking, the association between returns and the level of development (and in this case of employment rates) is also observed for unconditional returns.

All in all, these estimates confirm the positive (direct and indirect) effect of education on wages and the existence of substantial regional variability in the return to investments in this type of human capital. In addition, the results in Table 5 show that there was also regional heterogeneity in the return to the other types of human capital considered in this study: general experience in the labour market and specific experience in the firm (tenure). In the country as a whole, an additional year of general experience caused an increase of around 1% in the expected wage. The return to experience is much higher in regions such as Extremadura and Galicia (1.76% and 1.63% respectively) and substantially below the country average in others like Baleares (0.68%) and Cantabria (0.78%). The case of returns to tenure is quite similar, as the profile of wage increases associated with the defined intervals of years of specific experience varies widely across regions. For instance, in the case of Madrid there was a substantial gain linked to workers' tenure: employees with more than 15 years' experience in the firm earned as much as 41% more than those with one or less than one year. This gain was far lower in Extremadura and Galicia (14% and 15% respectively.

The evidence presented so far thus not only confirms that regions differed in the human capital endowment of their employees and the rest of their labour force but also shows sizeable regional variability in the return that individuals obtain from their accumulated human capital. As the final step in this study, the next section assesses the contribution of this variability in regional endowments and returns to the wage gap across regions.

4. HUMAN CAPITAL AND REGIONAL WAGE GAPS

4.1. Methodology

This section briefly describes the method proposed to obtain a detailed decomposition of the average wage gap between any two regions (A and B), or between a region and the rest of the country, under the presence of a selection process such as the one described in (2). Technical details of the derivation are sketched in the appendix. From expression (4), the average of conditional wages in regions A and B can be expressed as:

$$\overline{W}_A = \overline{X}_A \hat{\beta}_A + \hat{\theta}_A \overline{\lambda}_A \tag{9}$$

$$\overline{W}_B = \overline{X}_B \hat{\beta}_B + \hat{\theta}_B \overline{\lambda}_B \tag{10}$$

where the "over bar" represents the value of the sample's average. Defining the average of a counterfactual inverse Mills ratio for region B as:

$$\overline{\lambda}_B^A \equiv \overline{\frac{\varphi(Z_B \hat{\gamma}_A)}{\Phi(Z_B \hat{\gamma}_A)}} \tag{11}$$

the difference between the second terms in the RHS of equations (9) and (10) can be expressed as:

$$\left(\hat{\theta}_A \overline{\lambda}_A - \hat{\theta}_B \overline{\lambda}_B\right) = \hat{\theta}_A \left(\overline{\lambda}_B^A - \overline{\lambda}_B\right) + \hat{\theta}_A \left(\overline{\lambda}_A - \overline{\lambda}_B^A\right) + \left(\hat{\theta}_A - \hat{\theta}_B\right) \overline{\lambda}_B \tag{12}$$

Building on (12), Neuman and Oaxaca (2004) proposed an extension of the traditional decomposition as follows: [13]

$$\overline{W}_A - \overline{W}_B = \left(\overline{X}_A - \overline{X}_B\right)\hat{\beta}_A + \hat{\theta}_A \left(\overline{\lambda}_A - \overline{\lambda}_B^A\right) + \overline{X}_B \left(\hat{\beta}_A - \hat{\beta}_B\right) + \hat{\theta}_A \left(\overline{\lambda}_B^A - \overline{\lambda}_B\right) + \left(\hat{\theta}_A - \hat{\theta}_B\right) \overline{\lambda}_B \tag{13}$$

The first two terms in the RHS of (13), $\left(\overline{X}_A - \overline{X}_B\right)\hat{\beta}_A + \hat{\theta}_A \left(\overline{\lambda}_A - \overline{\lambda}_B^A\right)$, correspond to differences in the endowment of characteristics between regions A and B, both those directly affecting wages and those determining the probability of employment. The third and fourth terms, $\overline{X}_B \left(\hat{\beta}_A - \hat{\beta}_B\right) + \hat{\theta}_A \left(\overline{\lambda}_B^A - \overline{\lambda}_B\right)$, measure the contribution to the wage gap of regional heterogeneity in returns, through the direct and the indirect effect respectively. Finally, $\left(\hat{\theta}_A - \hat{\theta}_B\right) \overline{\lambda}_B$ is a sort of residual term related to the regional difference in the impact of the process of selection on wages.

The decomposition in (13) allows us to assess the contribution of characteristics and returns to the regional wage gap including the indirect effect coming from the process of selection. Therefore it is a decomposition of the gap in conditional wages. However, it does not allow us to obtain the contribution of each characteristic and each group of characteristics. This would be of particular interest when, as in this paper, we are interested in the effect of a set of

[13] Notice that in what follows it is assumed that the no-discrimination wage structure is that in region A.

variables such as those proxying for workers' human capital. The problem is how to assign the individual contribution to each variable when a non-linear term is involved; the actual and counterfactual inverse Mills ratios in equation (13). Our proposal to overcome this problem builds on Yun (2004)'s general decomposition of gaps in the first moments when the variable under analysis depends on a non-linear function which, however, has a linear function as argument. In this case, the decomposition in (13) can be expressed as:

$$\begin{aligned}\overline{W}_A - \overline{W}_B = & \left\{ \sum_{i=1}^{l_X} P^i_{\Delta X}\left[\left(\overline{X}_A - \overline{X}_B\right)\hat{\beta}_A\right] + \sum_{i=1}^{l_Z} P^i_{\Delta Z}\left[\hat{\theta}_A\left(\overline{\lambda}_A - \overline{\lambda}^A_B\right)\right]\right\} + \\ & \left\{ \sum_{i=1}^{l_X} P^i_{\Delta\beta}\left[\overline{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right)\right] + \sum_{i=1}^{l_Z} P^i_{\Delta\gamma}\left[\hat{\theta}_A\left(\overline{\lambda}^A_B - \overline{\lambda}_B\right)\right]\right\} + \\ & \left(\hat{\theta}_A - \hat{\theta}_B\right)\overline{\lambda}_B\end{aligned} \tag{14}$$

where

$$P^i_{\Delta X} = \frac{\left(\overline{X}^i_A - \overline{X}^i_B\right)\hat{\beta}^i_A}{\left(\overline{X}_A - \overline{X}_B\right)\hat{\beta}_A}, \qquad P^i_{\Delta Z} = \frac{\left(\overline{Z}^i_A - \overline{Z}^i_B\right)\hat{\gamma}^i_A}{\left(\overline{Z}_A - \overline{Z}_B\right)\hat{\gamma}_A}$$

$$P^i_{\Delta\beta} = \frac{\overline{X}^i_B\left(\hat{\beta}^i_A - \hat{\beta}^i_B\right)}{\overline{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right)}, \qquad P^i_{\Delta\gamma} = \frac{\overline{Z}^i_B\left(\hat{\gamma}^i_A - \hat{\gamma}^i_B\right)}{\overline{Z}_B\left(\hat{\gamma}_A - \hat{\gamma}_B\right)}$$

$$\sum_{i=1}^{l_X} P^i_{\Delta X} = \sum_{i=1}^{l_Z} P^i_{\Delta Z} = \sum_{i=1}^{l_X} P^i_{\Delta\beta} = \sum_{i=1}^{l_Z} P^i_{\Delta\gamma} = 1$$

are the weights that allow us to assign the contribution of each variable in X and Z to differences in characteristics ($P^i_{\Delta X}$ and $P^i_{\Delta Z}$) and in returns ($P^i_{\Delta\beta}$ and $P^i_{\Delta\gamma}$).[14] l_X and l_Z denote the number of characteristics included in X and in Z respectively.

4.2. Results

Instead of decomposing the wage gap for each pair of the 17 Spanish NUTS II regions, we computed the global and the detailed decomposition for the gap between the rest of the country and each region r, that is $\left(\overline{W}_{SP-r} - \overline{W}_r\right)$, where $\overline{W}_{SP-r}$ is the average (log) wage for the

[14] Notice that $P^i_{\Delta X}$ and $P^i_{\Delta\beta}$ are the weight in the standard linear decomposition.

sample of employees in Spain excepting those in region r, and $\overline{W}_r$ is the corresponding average for region r.[15] Then, following the notation in the previous section, A corresponds to SP-r, and B corresponds to r. To implement the decomposition of those gaps, we used the IV estimates of the coefficients in the wage and in the selection equation, β and γ, for each region, which were described in section 3.2. A set of IV estimates for the same coefficients was obtained corresponding to the samples of the rest of the country associated with each region. The characteristics of these estimates were similar to those discussed in section 3.2 in the case of the entire country.

As a first step, the results obtained for the global decomposition in (13) are summarized in Table 6. The first column of results shows the regional wage gap as defined above. It is positive when the average wage in the rest of the country exceeds the average wage in the region, and negative when the wage is higher in the region. The second and third columns of results correspond to the contribution of differences in endowments and returns to all the characteristics. Finally, the last column contains the contribution of the residual component in the decomposition which depends on the difference between the coefficient θ in the region and in the rest of the country: that is, the part of the wage gap attributed to differences in the particular impact of the probability of employment on the wage level.

For most of the Spanish regions, the contribution of returns is almost as large as that of endowment. Actually, it is particularly intense in regions with a positive gap. In these cases the contribution of returns clearly exceeds that of endowment (Galicia, Asturias, Cantabria, La Rioja) or both are of the same order of magnitude (Castile-La Mancha, Valencia, Baleares, and Murcia). Interestingly, in Extremadura and in the Canaries, the contribution of returns was so favourable that it counterbalanced part of the contribution of the other elements (endowment and residual term). In other words, if the returns to all the characteristics in those regions had been similar to the ones in the rest of the country, their wage gap would have been even larger. In sharp contrast, differences in endowments seem to explain most of the gap for regions with wages above the rest of the country (i.e. the Basque Country, Navarre, Aragon, Madrid, and to a lesser extent Castile-Leon and Catalonia). Finally, it should be

[15] It is impossible to summarize the results for the decomposition of the wage gap for all pairs of regions (17*16*0.5=136) in this type of publication. An alternative to the one in our study is to consider the gap with regard to a benchmark region (for instance, the one with the highest average wage), although this is subject to the criticism of the selection of the benchmark and slightly complicates the comparison of results across regions.

stressed that the contribution of the residual term in the decomposition in (13) is particularly intense for some regions, counterbalancing that associated with differences in returns in such cases (as in Asturias, La Rioja, Castile-La Mancha, Extremadura, and the Canaries).

The specific contribution of human capital to the wage gap in each region is summarized in Table 7. The first column of results reproduces the magnitude and sign of the wage gap, as in the previous table, in order to aid interpretation of the results. The effects of differences in endowment and in returns to human capital are shown in the second and third columns of results respectively. It is clear that both endowment and returns to human capital account for most of the gap in regions with wages above those in the rest of the country (e.g. Madrid, Basque Country, Aragon). Actually, if they had been the only source of regional differences, the wage gap in favour of those regions would have been much wider. Interestingly, the effect of differences in returns is even larger than that of endowments. For instance, in the case of the region with the highest average wage, Madrid, the actual wage gap was -0.17. However, differences in human capital endowment and returns with the rest of the country would have provoked a much higher gap (-0.36). About two-thirds of the gap corresponds to differences in returns (-0.23). The actual gap was much lower because differences in other characteristics partially counterbalanced the effect of human capital.

An interesting feature is observed for some regions with wages below those in the rest of the country. In Galicia, Castile-La Mancha, Andalusia, Murcia, and to a lesser extent in Baleares, the endowment of human capital contributed to the lower wages. But, in all cases, this effect was compensated by the contribution of returns. This was not so, however, for some other regions with low wages in which both effects worked in the same direction such as Extremadura, Valencia, and the Canaries. In any case, the results confirm that the contribution of differences in returns to human capital was greater than that of endowments for most of the regions.

The next step in our analysis was to isolate the particular contribution of education to regional wage gaps. In the descriptive analysis and in the discussion of the estimated returns in section 3.2, it was observed that regional heterogeneity was more intense in education than in the two types of workers' experience. Correspondingly, we expected that most of the effect of human capital would come from the contribution of regional differences in education. The third and fourth columns of results in Table 7 show the contribution of endowments and returns to education respectively. These figures confirm that most of the effects mentioned above in reference to human capital are related to education. As for the endowment, differences in education accounted for at least two-thirds of the effect of human capital in 12 out of the 17 regions, and in two other regions the effect corresponding to education was above one-third of the global effect of human capital. In the case of returns, the effect attributable to education was at least two-thirds in all but two regions. Actually, the magnitude of the effect was higher for education than for the total contribution of human capital in 10 regions, meaning that differences in returns to experience and tenure to some extent counterbalanced the impact on wage disparities of regional heterogeneity in the return to education. In only one region (Cantabria) did the contribution of returns to experience exceed that of education, causing a change in the sign of the effect of human capital (from -0.06, corresponding to education, to 0.03, the total effect of human capital returns).

All in all, the results in this section support our hypothesis regarding the role played by differences in endowment and also in returns to human capital to explain regional wage gaps. Similarly, within human capital, the crucial elements are the endowment of individuals' education and the return that they obtain from it.

5. CONCLUDING COMMENTS

The results of this study confirm the usefulness of using micro-data in studies dealing with regional economic disparities and the impact of intangible assets in each region. They provide complementary empirical evidence to that obtained from aggregated regional data. In the specific case of human capital, the use of individual data allowed us to evaluate both the impact of differences in the endowment and the return obtained by individuals within each region.

We show that there are significant regional differences in the distribution of education and experience in Spain. We also provide evidence of the existence of strong disparities in the return to human capital, especially in the case of education. Actually, the results suggest that a large proportion of the total effect of education is related to an indirect effect, since the impact of education on employability varies considerably from region to region. The detailed decomposition of the regional wage gaps has allowed us to demonstrate that regional heterogeneity in the returns to human capital was the main factor explaining wage disparities across regions. Moreover, the detailed results suggest that most of this effect should be attributed to differences in the return to education, since the differences associated to returns to tenure and experience played a minor role in most regions.

An immediate implication can be drawn from these results. It appears that policies aiming to promote education are an effective tool in improving workers' productivity and in lowering the risk of unemployment and non-participation in the labour market. The effect of these policies is also likely to be stronger in regions with lower levels of development. Therefore, raising educational attainment in these regions would contribute to regional convergence in labour productivity and unemployment and participation rates. The overall effect would thus be an increase in the average income per capita of the less favoured regions and a reduction in regional disparities. Also worth noting is the suggestion that the promotion of education in less developed regions simultaneously meets the goals of equity and efficiency, given that the return of this policy is higher in less developed regions than in more advanced ones.

Finally, it must be emphasized that the conclusions are derived from a partial equilibrium exercise. As is usual in exercises of this kind, the counterfactual analysis in this paper did not predict the reaction of workers and firms, for instance, to the regional equalization of endowments and/or returns to human capital. On the other hand, the system of collective bargaining existing in Spain may be inducing wage differences between regions, independently of workers' characteristics; differences in returns may then be related to differences in sectoral minimum wages determined at subnational level (Simón et al, 2006). Still, our feeling is that the contribution of this element to the estimated regional differentials in the return to education is not as important as to invalidate the results discussed above. Nonetheless, a deeper analysis of this point is on our future research agenda.

APPENDIX

Evaluating the values of the inverse Mills ratios involved in the RHS of (12) using mean characteristics results in:

$$\left(\hat{\theta}_A \bar{\lambda}_A - \hat{\theta}_B \bar{\lambda}_B\right) = \hat{\theta}_A \left(\tilde{\lambda}_B^A - \tilde{\lambda}_B\right) + \hat{\theta}_A \left(\tilde{\lambda}_A - \tilde{\lambda}_B^A\right) + \left(\hat{\theta}_A - \hat{\theta}_B\right)\bar{\lambda}_B + R_M \qquad \text{(A.1)}$$

where $\tilde{\lambda}_r = \left(\varphi(\bar{Z}_r \hat{\gamma}_r) / \Phi(\bar{Z}_r \hat{\gamma}_r)\right)$, r= A, B, and $\tilde{\lambda}_B^A = \left(\varphi(\bar{Z}_B \hat{\gamma}_A) / \Phi(\bar{Z}_B \hat{\gamma}_A)\right)$. The error of approximation, R_M, is:

$$R_M = R_{MA} + R_{MB} + R_{MB}^A \qquad \text{(A.2)}$$

$$R_{Mr} = \bar{\lambda}_r - \tilde{\lambda}_r = \overline{\frac{\varphi(Z_r \hat{\gamma}_r)}{\Phi(Z_r \hat{\gamma}_r)}} - \frac{\varphi(\bar{Z}_r \hat{\gamma}_r)}{\Phi(\bar{Z}_r \hat{\gamma}_r)} \quad ; \quad r = A, B$$

$$R_{MB}^A = \bar{\lambda}_B^A - \tilde{\lambda}_B^A = \overline{\frac{\varphi(Z_B \hat{\gamma}_A)}{\Phi(Z_B \hat{\gamma}_A)}} - \frac{\varphi(\bar{Z}_B \hat{\gamma}_A)}{\Phi(\bar{Z}_B \hat{\gamma}_A)}$$

Using a first order Taylor expansion to linearize the terms that involve the inverse Mills ratios, $\hat{\theta}_A\left(\tilde{\lambda}_A - \tilde{\lambda}_B^A\right)$ and $\hat{\theta}_A\left(\tilde{\lambda}_B^A - \tilde{\lambda}_B\right)$, around $\bar{Z}_A \hat{\gamma}_A$ and $\bar{Z}_B \hat{\gamma}_B$ respectively:

$$\begin{aligned} \hat{\theta}_A\left(\hat{\lambda}_A - \hat{\lambda}_B^A\right) &= \hat{\theta}_A \bar{f}_A \left(\bar{Z}_A - \bar{Z}_B\right)\hat{\gamma}_A + R_{T1} \\ \hat{\theta}_A\left(\hat{\lambda}_B^A - \hat{\lambda}_B\right) &= \hat{\theta}_A \bar{f}_B \bar{Z}_B \left(\hat{\gamma}_A - \hat{\gamma}_B\right) + R_{T2} \end{aligned} \qquad \text{(A.3)}$$

where $\bar{f}_r = \partial(\lambda_r)/\partial(\hat{\alpha}_r) = -\lambda_r^2 + \hat{\alpha}_r \lambda_r$, r=A, B, $\hat{\alpha}_r = \bar{Z}_r \hat{\gamma}_r$, and R_{T1}, R_{T2} are the residuals of approximation.

Using (A.3) and (A.1) the decomposition in (13) can be expressed as:

$$\begin{aligned} \bar{W}_A - \bar{W}_B = & \left[\left(\bar{X}_A - \bar{X}_B\right)\hat{\beta}_A + \hat{\theta}_A \bar{f}_A \left(\bar{Z}_A - \bar{Z}_B\right)\hat{\gamma}_A\right] \\ & + \left[\bar{X}_B\left(\hat{\beta}_A - \hat{\beta}_B\right) + \hat{\theta}_A \bar{f}_B \bar{Z}_B \left(\hat{\gamma}_A - \hat{\gamma}_B\right)\right] + \left(\hat{\theta}_A - \hat{\theta}_B\right)\bar{\lambda}_B + R_M + R_{T1} + R_{T2} \end{aligned} \qquad \text{(A.4)}$$

The expression in (A.4) is then used to obtain the weights for the contribution of each characteristic and return as shown in (14). To obtain $P_{\Delta Z}^i$ and $P_{\Delta\gamma}^i$ as in (14) it should only be noted that $\hat{\theta}_A \bar{f}_A$ and $\hat{\theta}_A \bar{f}_B$ do not vary across the variables in Z.

REFERENCES

Arrazola M., De Hevia J., Risueño M. and Sanz J. F. (2003) Returns to education in Spain: Some evidence on the endogeneity of schooling, Education Economics 11, 293–304.

Arrazola M. and De Hevia J. (2008) Three measures of returns to education: An illustration for the case of Spain, Economics of Education Review 27, 266-75.

Bronzini R. and Piselli P. (2009) Determinants of long-run regional productivity with geographical spillovers: The role of R&D, human capital and public infrastructure, Regional Science and Urban Economics 39, 187-99.

Cameron A. C. and Trivedi P. K. (2005) Microeconometrics: Methods and Applications. Cambridge University Press, New York

Di Liberto A. (2008) Education and Italian regional development, Economics of Education Review, 27, 94-107.

Farber S. and Newman R. (1989) Regional wage differentials and the spatial convergence of worker characteristic prices, Review of Economics and Statistics 71, 224-31.

García I. and Molina A. (2002) Inter-regional wage differentials in Spain, Applied Economics Letters 9, 209-15.

García-Pérez J. I. and Jimeno J. F. (2007) Public sector wage gaps in Spanish regions, The Manchester School 75, 501–31.

Greene W. H. (2003) Econometric analysis (5th ed.), Prentice-Hall, New Jersey

Harmon C. and Walker I. (1995) Estimates of the Economic Return to Schooling for the United Kingdom, American Economic Review, 85, 1278-86.

Heckman J. (1976) The common structure of statistical models of truncation, sample selecction, and limited dependent variables and a simple estimator for such models, Annals of Economic and Social Measurement 5, 475-492.

Hoffmann R. and Kassouf A. (2005) Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman's procedure, Applied Economics 37: 1303-11.

Ichino A. and Winter-Ebmer R. (1999) Lower and upper bounds of returns to schooling: an exercise in IV estimation with different instruments, European Economic Review 43, 889–901.

Ichino A. and Winter-Ebmer R. (2004) The long-run educational cost of World War II, Journal of Labor Economics 22, 57–86.

López-Bazo E. and Moreno R. (2008) Does human capital stimulate investment in physical capital?: Evidence from a cost system framework, Economic Modelling, 25, 1295-1305

Montuenga V., García I. and Fernández M. (2003) Wage flexibility: Evidence from five EU countries based on the wage curve, Economics letters 78, 169-74.

Motellón E. (2008) Un análisis de las diferencias regionales en el impacto de la contratación temporal en España, Investigaciones Regionales 12, 107-31

Neuman S. and Oaxaca R. (2004) Wage decompositions with selectivity-corrected wage equations: A methodological note, Journal of Economic Inequality 2, 3-10.

Pereira P. T. and Silva P. (2004) Returns to education and wage equations, Applied Economics 36, 525–31.

Reilly B. (1991) An analysis of local labour market wage differentials, Regional Studies 26, 257-64.

Rodríguez-Pose A and Vilalta-Bufí M (2005) Education, migration, and job satisfaction: The regional returns of human capital in the EU, Journal of Economic Geography, 5, 545-66.

Simón H., Ramos R. and Sanromà E. (2006) Collective bargaining and regional wage differences in Spain: an empirical analysis, Applied Economics 38, 1749-60.

Wooldridge J. (2002) Econometric analysis of cross section and panel data, MIT Press, Cambridge MA.

Yun M. (2004) Decomposing differences in the first moment, Economics letters 82, 275-80.

Table 1. Hourly wages in the Spanish regions.

	Obs	Hourly gross wage Mean	Hourly gross wage Std. Dev.	Real Hourly gross wage Mean	Real Hourly gross wage Std. Dev.
Galicia	795	6.42	3.830	6.51	3.888
Asturias	396	7.06	4.152	7.00	4.111
Cantabria	455	6.55	3.963	6.65	4.023
Basque Country	618	8.75	4.362	8.29	4.134
Navarre	496	7.91	3.543	7.36	3.296
La Rioja	358	6.67	2.932	6.57	2.886
Aragon	576	8.00	4.507	8.29	4.671
Madrid	1174	8.49	4.800	8.51	4.809
Castile-Leon	683	7.78	4.659	8.15	4.879
Castile-La Mancha	613	6.53	3.545	7.06	3.836
Extremadura	482	6.10	3.496	7.05	4.037
Catalonia	1513	7.92	4.490	7.44	4.216
Valencia	886	6.37	2.922	6.46	2.961
Baleares	379	6.80	3.291	6.50	3.143
Andalusia	1336	6.60	3.256	6.91	3.409
Murcia	558	6.23	3.498	6.37	3.580
Canary Isl.	848	6.49	4.224	6.65	4.327
Spain	12166	7.19	4.062	7.24	4.063

Table 2. Description of the variables in the empirical wage model for Spain and for the regions with the highest and lowest wage levels.

	SPAIN		MADRID		MURCIA	
	Employees	**Non-Employees**	**Employees**	**Non-Employees**	**Employees**	**Non-Employees**
WORKER'S HUMAN CAPITAL						
Education (years of schooling)	8.960	4.943	10.931	6.077	7.889	4.820
	(5.370)	(4.195)	(5.156)	(4.608)	(5.516)	(4.372)
Experience (years)	18.206	-	18.045	-	17.972	-
	(12.075)	-	(11.921)	-	(12.513)	-
Tenure						
≤ 1 year	29.41%	-	24.07%	-	29.38%	-
2-4 years	20.88%	-	22.38%	-	23.54%	-
5-9 years	9.53%	-	11.23%	-	9.56%	-
10-14 years	10.95%	-	10.98%	-	11.68%	-
≥ 15 years	28.31%	-	30.49%	-	24.60%	-
INDIVIDUAL AND FAMILY CHARACTERISTICS						
Age (years)	37.259	44.522	37.383	45.783	36.159	41.855
	(10.684)	(11.876)	(10.456)	(11.207)	(11.342)	(12.168)
Household size (persons)	3.697	3.813	3.542	3.676	3.933	4.000
	(1.276)	(1.382)	(1.205)	(1.168)	(1.320)	(1.501)
Other household income	1081.951	1465.931	1280.170	1774.817	1024.620	1349.367
(€ per month)	(972.951)	(913.266)	(1127.170)	(1089.941)	(1025.685)	(865.200)
Nº children ≤ 15 years	0.753	0.928	0.693	0.914	0.922	1.275
	(0.892)	(1.106)	(0.881)	(1.081)	(0.990)	(1.380)
Children 0-6 years	24.58%	29.10%	22.80%	27.54%	30.97%	36.24%
Gender						
Male	59.99%	12.54%	53.29%	7.57%	63.72%	12.60%
Female	40.01%	87.46%	46.71%	92.43%	36.28%	87.40%
Marital status						
Married	65.65%	84.49%	65.79%	88.12%	66.02%	85.08%
Other	34.35%	15.51%	34.21%	11.88%	33.98%	14.92%
Chronic disease	8.75%	25.16%	9.04%	22.72%	10.27%	24.61%

Note: Sample means and standard deviation in parentheses for the continuous variables. Share of each category for the discrete characteristics.

Table 3. Wage level within categories of worker human capital endowments.

	SPAIN	MADRID	MURCIA
Education			
Illiterate	5.18	5.60	4.49
Primary	5.73	6.33	5.11
Secondary	6.74	7.46	5.97
Tertiary	10.78	11.37	10.53
Experience			
≤ 1 year	4.70	5.05	3.94
2-9 years	5.92	6.80	5.41
9-19 years	7.47	8.87	6.75
19-29 years	8.12	9.08	6.94
≥ 30 years	8.06	10.12	7.11
Tenure			
≤ 1 year	5.37	6.04	4.56
2-4 years	6.32	7.50	5.70
5-9 years	7.28	8.69	5.79
10-14 years	8.38	9.78	7.66
≥ 15 years	9.38	10.67	8.80

Note: Sample mean of real wage per hour in euros within each category.

Table 4. Returns to education in the Spanish regions.

Region	Conditional Return	Unconditional Return
Galicia	0.0895	0.1903
Asturias	0.0625	0.1823
Cantabria	0.0968	0.1674
Basque Country	0.0596	0.1585
Navarre	0.0716	0.1240
La Rioja	0.0552	0.1312
Aragon	0.0664	0.1940
Madrid	0.0491	0.1625
Castile-Leon	0.0774	0.1848
Castile-La Mancha	0.0721	0.2022
Extremadura	0.0713	0.2450
Catalonia	0.0536	0.1253
Valencia	0.0501	0.1604
Baleares	0.0662	0.1732
Andalusia	0.0603	0.1943
Murcia	0.0795	0.1665
Canary Isl.	0.0680	0.1590
Spain	0.0633	0.1679

Note: All the returns in the table are statistically significant at 1%.

Table 5. Returns to experience and tenure in the Spanish regions.

Region	**Experience**	**Tenure**			
		2-4 years	5-9 years	10-14 years	≥ 15 years
Galicia	0.0163 ***	0.0773 ***	0.0255	0.0227	0.1515 ***
Asturias	0.0112 ***	0.0629	0.1210 *	0.3785 ***	0.3624 ***
Cantabria	0.0078 ***	-0.0478 *	0.0220	0.0475	0.1636 ***
Basque Country	0.0098 ***	0.0997 ***	0.1692 ***	0.2994 ***	0.3376 ***
Navarre	0.0106 ***	0.0989 ***	0.2134 ***	0.1773 ***	0.2757 ***
La Rioja	0.0085 ***	0.1886 ***	0.1019 **	0.2692 ***	0.2497 ***
Aragon	0.0093 ***	0.0955 ***	0.1916 ***	0.2269 ***	0.3758 ***
Madrid	0.0081 ***	0.1355 ***	0.2428 ***	0.3712 ***	0.4138 ***
Castile-Leon	0.0133 ***	-0.0271	0.1175 ***	0.2717 ***	0.2391 ***
Castile-La Mancha	0.0091 ***	-0.0318	0.1097 ***	0.0833 **	0.2978 ***
Extremadura	0.0176 ***	0.0908 **	0.0569	0.1436 ***	0.1381 ***
Catalonia	0.0109 ***	0.1250 ***	0.2062 ***	0.2629 ***	0.2818 ***
Valencia	0.0082 ***	0.0977 ***	0.1093 ***	0.2254 ***	0.3114 ***
Baleares	0.0068 ***	0.0295	0.0781 *	0.1310 ***	0.1791 ***
Andalusia	0.0127 ***	0.0912 ***	0.1401 ***	0.2653 ***	0.2298 ***
Murcia	0.0113 ***	0.1214 ***	0.1301 ***	0.2256 ***	0.2960 ***
Canary Isl.	0.0115 ***	0.0969 ***	0.1023 ***	0.2544 ***	0.3491 ***
Spain	0.0108 ***	0.0858 ***	0.1408 ***	0.2328 ***	0.2881 ***

Table 6. Regional wage gap decomposition.

	Wage gap	Endowment	Returns	Residual
Galicia	0.1201	0.0266	0.1067	-0.0134
Asturias	0.0415	0.0004	0.1771	-0.1357
Cantabria	0.0993	-0.0278	0.0598	0.0674
Basque Country	-0.1666	-0.0759	0.0273	-0.1181
Navarre	-0.0517	-0.0363	-0.0032	-0.0122
La Rioja	0.0541	-0.0259	0.1998	-0.1199
Aragon	-0.1352	-0.0968	0.0385	-0.0765
Madrid	-0.1699	-0.1196	0.0478	-0.0977
Castile-Leon	-0.1055	-0.0657	-0.0665	0.0267
Castile-La Mancha	0.0259	0.0793	0.0861	-0.1393
Extremadura	0.0307	0.0436	-0.1837	0.1708
Catalonia	-0.0456	-0.0124	0.0652	-0.0983
Valencia	0.0896	0.0355	0.0476	0.0062
Baleares	0.0845	0.0419	0.0559	-0.0137
Andalusia	0.0342	0.0651	-0.0255	-0.0057
Murcia	0.1366	0.0606	0.0722	0.0035
Canary Isl.	0.1146	0.0996	-0.0944	0.1094

Table 7. Contribution of human capital and schooling to regional wage gaps

		Human Capital		Schooling	
	Wage Gap	**Endowment**	**Return**	**Endowment**	**Return**
Galicia	0.1201	0.0255	-0.4095	0.0156	-0.2862
Asturias	0.0415	0.0035	-0.0097	0.0033	-0.0266
Cantabria	0.0993	-0.0309	0.0346	-0.0178	-0.0602
Basque Country	-0.1666	-0.0764	-0.2421	-0.0631	-0.2624
Navarre	-0.0517	-0.0343	-0.0950	-0.0104	-0.1315
La Rioja	0.0541	-0.0282	-0.0023	-0.0072	-0.0519
Aragon	-0.1352	-0.1010	-0.2135	-0.0608	-0.2515
Madrid	-0.1699	-0.1293	-0.2305	-0.1193	-0.2805
Castile-Leon	-0.1055	-0.0646	-0.2177	-0.0238	-0.1447
Castile-La Mancha	0.0259	0.0834	-0.3063	0.0620	-0.2930
Extremadura	0.0307	0.0462	0.1353	0.0306	0.1825
Catalonia	-0.0456	-0.0104	-0.0888	-0.0012	-0.0608
Valencia	0.0896	0.0353	0.1223	0.0304	0.0967
Baleares	0.0845	0.0356	-0.0273	0.0138	-0.0311
Andalusia	0.0342	0.0701	-0.1156	0.0446	-0.0437
Murcia	0.1366	0.0683	-0.0907	0.0557	-0.1167
Canary Isl.	0.1146	0.0954	0.0888	0.0586	0.1020