Working papers series

WP ECON 09.17

Estimating Human Capital Externalities: The Case of Spanish Regions

Walter García-Fontes (U. Pompeu Fabra y CREA)

Manuel Hidalgo (U. Pablo Olavide)

JEL Classification numbers: I21, J31, O47

Keywords: externalities, human capital, constant composition.

Department of Economics

Estimating Human Capital Externalities: The Case of Spanish Regions

Walter Garcia-Fontes*
Manuel Hidalgo†

September 2009 ‡

*Universitat Pompeu Fabra and CREA
†Universidad Pablo de Olavide
‡Walter Garcia-Fontes thanks partial support of grants PR08305 and PR00504 from the Spanish Ministry of Education. We thank the comments from Antonio Ciccone. Corresponding author: Manuel Hidaldo. Universidad Pablo de Olavide, Ctra de Utrera, km1, Sevilla. CP 41013, Spain. Email: mhidper@upo.es.

Abstract

We estimate the strength of schooling externalities for Spanish regions over the 1981-2001 period. Our empirical work employs both main approaches available in the literature. Both methodologies yield significant externalities. Using a growth accounting exercise, we find that human capital externalities account for one half of the increase in real wages for the period between 1981 and 2001.

JEL Codes: I21, J31, O47

1 Introduction

Estimating human capital externalities—the difference between the social and the private marginal returns to human capital—is important for various reasons. First, the strength of such externalities determines the optimal subsidies to education and to immigration of highly qualified workers. Second, human capital externalities have been emphasized as a key for understanding the process of economic growth (e.g. Lucas (1988)). It is therefore not too surprising that there are a variety of estimation approaches and estimates in the literature (e.g. Rauch (1993); Black and Henderson (1999); Acemoglu and Angrist (2001); Rudd (2000); Moretti (2004a); Moretti (2004b); Ciccone and Peri (2006)).

For Spain there is much less work however. The available estimates of the return to human capital almost all reflect private returns (e.g. Alba and Segundo (1995); Barceinas, Oliver, Raymond, and Roig (2000); Raymond (2002); De la Fuente (2003) and De la Fuente, Domenech, and Jimeno (2003)). As far as we know, there are only three attempts to estimate social returns to education or externalitities. De la Fuente and Domenech (2005) estimate social marginal returns to education in the nineties while Alcala and Hernandez (2005) estimate human capital externalities at the firm and industry level. García-Fontes and Hidalgo (2008) find externalities using aggregated regional data for the period 1980-2000. All these papers find evidence of positive significant externalities for Spain.

There are currently two approaches in the literature to estimate human capital externalities. The first approach augments standard Mincerian wage equations with variables that measure the level of human capital at the regional level (e.g. Rudd (2000); Acemoglu and Angrist (2001); Moretti (2004a)). This methodology estimates the strength of human capital externalities by looking at the effect of regional human capital levels on individual wages. The basic idea is that human capital externalities should show up in individual wages once all relevant individual characteristics are controlled for. A key assumption of this (Mincerian) approach is that workers with different levels of human capital are perfect substitutes in production. If different human capital levels are imperfect substitutes, the Mincerian approach yields a positive effect of aggregate human capital on individual wages even if the social return to human capital equals the private return (e.g. Ciccone and Peri (2006)). The intuition is that with imperfect substitution, an increase in the number of skilled workers implies an increase of the wages of unskilled workers that more than offsets the decrease in the wage of skilled workers. The empirical evidence for the United States (e.g. Katz and Murphy (1992); Ciccone and Peri (2006)), as well as other countries (e.g. Angrist (1995)) including Spain (Hidalgo (2009)), indicates that different levels of human capital are imperfect substitutes. Therefore the Mincerian approach must be complemented with the so-called constant composition approach, which yields consistent estimates of the wedge between the social and the private return to human capital even if skilled and unskilled workers are imperfect substitutes (Ciccone and Peri (2006)). This approach estimates the strength of human capital externalities as the marginal effect of aggregate human capital levels on average wages holding the labor force composition constant. Ciccone and Peri show that this bias is directly related to the wage difference between skilled and unskilled workers, and inversely related to the elasticity of substitution. The smaller the wage premium of skilled workers the smaller the bias introduced by the Mincerian approach.

In García-Fontes and Hidalgo (2008) human capital esternalities are estimated for Spain at the regional level. Using the constant composition approach, human capital externalities are estimated to be positive and significant. In the current paper human capital externalities are estimated using both the Mincerian and constant composition approach, and the bias of the Mincerian approach is quantified. This bias is expected to be significant and large in the presence of imperfect substitutability between workers of different skill levels[1].

Both approaches yield evidence of significant human capital externalities. In line with theory, the Mincerian approach yields larger externalities than the constant composition approach. The difference in the point estimate of the externalities of human capital between the two approaches depends on the particular specification adopted, but has an approximate average value of 11%. It can be therefore considered that the Mincerian approach provides an upper bound for the estimation of human capital externalities, while the constant composition approach provides a lower bound. Taking this into account, and through a growth accounting exercise, we find that human capital externalities account for more or less one third of the increase in wages for the period between 1981 and 2001.

A key issue when estimating human capital externalities at the regional level is that changes in aggregate human capital levels are endogenous as regions with higher productivity and wages may attract more skilled workers. This makes it desirable to implement an instrumental-variables approach. We present instrumental-variables estimates of the strength of human capital

[1]Hidalgo (2009) and Hidalgo, O'Kean y Rodríguez (2008) find evidence in favor of imperfect substitutability of workers of different skill levels for Spain.

externalities that instrument the regional increase in human capital levels by the initial demographics of each region. In particular, we will show that shares of young and old population age groups help to predict future increases in human capital over the period we analyze, as older, retiring workers had considerably lower schooling levels than young workers entering the labor force.

The rest of the paper is organized as follows. Section 2 summarizes the relevant literature, while Section 3 reviews the two main empirical methodologies used and how the bias might be calculated. Section 4 presents the data sources used. The main results are in Section 5 and Section 6 concludes.

2 Related Literature

The strength of human capital externalities is defined as the difference between the social and private marginal return to an additional unit of human capital. Most empirical work focuses on the return to an additional year of (formal) schooling. There is a very large literature on the private return to an additional year of schooling, which has found the return to lie between 5 and 12% depending on the country and time period considered (e.g. Card (1999)). There is less work estimating the strength of schooling externalities. Rauch (1993) estimates schooling externalities in the US in 1980 using a Mincerian wage equation augmented for state-level schooling measures. The idea behind the Mincerian approach is that if there are externalities, individual wages should be increasing in aggregate schooling levels controlling for individual characteristics, such as education, experience, gender, etc. Rauch find schooling externalities between 3 and 5%. Later contributions refine the Mincerian approach by using panel data to control for region fixed effects and by employing instruments for the change in aggregate schooling levels, see Acemoglu and Angrist (2001), Rudd (2000), Conley, Flier, and Tsang (2003), Moretti (2004a), and Moretti (2004b)). The result vary with the time period, the level of spatial aggregation, the country, and the specification. For example, while Acemoglu and Angrist (2001) do not find state-level average schooling externalities in the US over the 1960-1980 period, Moretti (2004b) finds externalities from the share of college workers in the US to be significant at the city level for 1981-1991.

The Mincerian approach to human capital externalities assumes that workers with different human capital levels are perfect substitutes in production. Perfect substitutability simplifies identification because it implies that changes in the relative supply of human capital do not affect the relative wages of the different human capital groups, holding total factor productivity constant. Consequently, all the effects that human capital supply changes have on workers with a given level of human capital have to come through total factor productivity and can be interpreted as externalities. Ciccone and Peri (2006) show that when workers with different human capital are imperfect substitutes, the Mincerian approach overestimates the strength of schooling externalities. They propose an alternative methodology that estimates externalities as the marginal effect of human capital on log average wages holding labor-force composition constant.

This paper complements the existing literature for Spain. De la Fuente and Domenech (2005) relate regional productivity growth to human capital and other variables. They estimate a 16% elasticity of human capital, which is larger than the elasticity estimated in previous works which is around 8%. They attibute the difference to the social return to education. Using an internal rate of return approach, they estimate the social rate of return of education to be between 10% and 12%. Taking the difference between the private and the social return, externalities are estimated to be between 4% and 5%. Alcalá and Hernández (2005) estimate the externalities of human capital at the firm level. Their estimates show a private return of 8% and externalities equal to 4,7%.

3 Empirical Approaches to Human Capital Externalities

We now turn to the two approaches that we will use to estimate regional schooling externalities in Spain.

3.1 The Mincerian approach

The main idea of the Mincerian approach to schooling externalities is to introduce aggregate schooling levels as an additional explanatory variable in an otherwise standard Mincerian wage equation. To estimate regional schooling externalities, the approach can be implemented in two steps. In the first step, we regress the log of individual wages on individual characteristics, like education, experience, and gender, and on region-time fixed effects. The goal of this first step is to estimate average regional wages net of the private returns to the characteristics that determine individual productivity. Formally, the first step estimates

$$\log(w_{irt}) = \alpha_{rt} + \gamma s_{irt} + \sum_{k=0}^{K} \beta_t^k z_{irt}^k + u_{irt}, \tag{1}$$

where w_{irt} is the wage of individual i in region r during year t, s_{irt} is individual years of schooling, z_{irt} are K additional individual characteristics that we may want to control for, and u_{irt} captures the effect of unobservable variables. Region-time effects are captured by α_{rt}. These region-time effects capture "cleaned" average regional wages at a given point in time.

In the second step, we regress the change in estimated "cleaned" average regional wages over time on the change in regional schooling levels,

$$\Delta\hat{\alpha}_{rt} = \hat{\alpha}_{rt} - \hat{\alpha}_{r\tau} = controls + \theta\Delta\hat{h}_{rt} + v_{rt}, \tag{2}$$

where Δh_{rt} represents the change in schooling in region r between t and τ. The strength of the schooling externalities in the Mincerian approach is equal to θ. It is easy to add additional regional controls to this approach and to use instrumental-variables techniques (if appropriate instruments are available). Notice that all factors causing permanent differences in productivity and wages across regions are differenced out when using this approach.

We will instrument the change in regional schooling levels over the 1981-1991 period by 1981 demographic structure and the change in regional schooling levels over the 1991-2001 period by 1991 demographic structure. The idea is that, as younger generations got considerably more schooling than those who retired over this period, regional schooling levels in the absence of migration should have been rising faster the larger younger relative to older age groups in the regional population. With regional migration, regional schooling levels also depend on the inflow and outflow of workers, but we will show that regional demographics continue to matter.

3.2 The Constant Composition Approach

The Mincerian approach assumes that workers with different levels of human capital are perfect substitutes. The empirical evidence suggests that this is often not the case (e.g. Katz and Murphy (1992) and Ciccone and Peri (2006) for the for the United States; Angrist (1995) for Israel; and Hidalgo (2009) for Spain). It is therefore interesting to also estimate schooling externalities with an approach that can identify externalities whether different human capital levels are perfect or imperfect substitutes. Ciccone and Peri (2006) develop an approach that can also be implemented in two steps.

The first step consists of estimating "cleaned" average wages at the region-year level, just as in the case of the Mincerian approach. But now this estimation must be done separately for each schooling level. Formally, we therefore estimate,

$$\log(w_{isrt}) = \alpha_{srt} + \sum_{k=0}^{K} \beta_{st} z_{isrt} + u_{isrt}; \quad (3)$$

that is, we estimate (1) separately for each schooling level. As a result, the fixed effects (α_{srt}) are now specific to a region, a year, and a level of schooling (s). The region-year-schooling effects α_{srt} are then aggregated using the labor-force composition weights of a base year T, l_{srT}, to obtain a constant-labor-force-composition "cleaned" average wage for each region and year,

$$\log(\hat{w}_{rt}^{F}) = \log\left(\sum_{s=1}^{S} l_{srT}\alpha_{srt}\right). \quad (4)$$

Ciccone and Peri (2006) show that the base year is best chosen in the middle of the time period being analyzed. The log change in fixed-composition "cleaned" average wages is then regressed on the change in regional schooling levels,

$$\log(\hat{w}_{rt}^{F}) - \log(\hat{w}_{r\tau}^{F}) = controls + \theta\Delta h_{rt} + v_{rt}. \quad (5)$$

Just like in the case of the Mincerian approach, it is easy to add controls or implement an instrumental-variables approach.

3.3 Mincerian estimation bias

According to Ciccone and Pieri (2006), the bias of the Mincerian approach when skilled and unskilled workers are substitutable, can be estimated as:

$$Bias = \frac{1}{\sigma}(\frac{w_s - w_u}{w}),$$

where σ is the elasticity of substitution between more (skilled, s) and less (unskilled, u) educated workers and the term in brackets is the wage premium of skilled workers. If we use Hidalgo, O'Kean and Rodriguez (2008) average estimate for Spain of the elasticity of substitution between colleged-educated workers and the rest of workers of 1.6 and assume a wage premium between 1980 and 2001 of approximately 40% as estimated by Hidalgo (2008), the bias of the Mincerian estimation is approximately 25% (0.4/1.6).

4 Data and Instruments

We use several data sources to implement the empirical approaches presented above. Individual data on wage earnings and schooling come from the Survey of Family Budgets[2] corresponding to 1981 and 1991 (EPF-81 and EPF-91) and the Continuous Survey of Family Budgets[3] corresponding to 2000/01 (ECPF-00-01), which we pool. Our regional schooling data comes from the Human Capital Project of the Instituto Valenciano de Investigaciones Econmicas (IVIE). For robustness analysis, we also employ total regional physical capital and ICT capital data also from the IVIE database. The demographic variables used for instrumentation comes from the Population Census for the years 1981 and 1991.[4] All this data was collected for 17 Spanish Autonomous Communities (all except Ceuta and Melilla).

4.1 Individual Data

The Survey of Family Budgets (1981 and 1991) and the Continuous Survey of Family Budgets (2000 and 2001) contain earnings information and other characteristics of family members. To gain accuracy, we restrict the sample to heads of households aged 16 to 65 who work more than 15 hours per week as employees. We also eliminate individuals with earnings below the minimum wage. This yields 7027 individuals for 1981, 8193 for 1991, and 2057 for the years 2000 and 2001. The exact data that we use is:

- *Wage earnings:* Labor earnings of employees (excluding self-employment).
- *Years of schooling:* We use individual years of schooling in the Mincerian wage regression to estimate the private return to an additional year of schooling. This variable is constructed by assigning to each schooling level the minimum amount of years needed to complete the degree. Table 1 shows the different schooling groups that we define and the information of schooling levels contained in the surveys. It can be seen that there are 8 schooling levels in EPF-81; 10 schooling levels in EPF-91; and 7 schooling levels in ECPF-00-01. Our results show that between 1981 and 2001 average of schooling in Spain increased from 6 to nearly 10 years, with the largest increase occurring during the 1990s.

[2] *Encuesta de Presupuestos Familiares, Instituto Nacional de Estadstica,*
[3] *Encuesta Continua de Presupuestos Familiares, Instituto Nacional de Estadstica.*
[4] *Censo de la poblacin, Instituto Nacional de Estadstica.*

[table 1 about here]

- *Schooling level:* To implement the constant composition approach we need to define schooling levels. We use primary education or less, secondary education (lower and upper), and college.
- *Gender:* Always used as a control in the first step wage regression.
- *Experience:* Experience is defined as age minus years of schooling minus 6.
- *Agricultural workers:* The first step also controls for whether workers work in agriculture or not.
- *Married:* The first step wage regressions also control for whether workers are married or not.

4.2 Regional Data

We use the following regional data:

- *Share of workers by schooling level:* Used to obtain average "cleaned" wages at the base year labor-force composition for the constant composition approach (Source: Human Capital Project Data).
- *Fraction of workers with secondary or college education:* Used to measure regional schooling levels (Source: Human Capital Project Data).
- *Total workers:* Used to measure the size of regions (Source: Active Population Survey[5]).
- *Physical capital stock and ICT capital stock:* Used for robustness analysis (Source: IVIE Database Data).
- *Population proportions for each age group:* These data come from the 1981 and 1991 Spanish population and housing censuses, published by the Spanish National Institute of Statistic (INE).

[table 2 about here]
[table 3 about here]

[5]Encuesta de Poblacin Activa.

4.3 Instruments

Migration across regions implies that regional schooling levels are endogenous. Higher productivity and wages in a region may lead to it attracting high skilled workers from elsewhere. Another factor that may work in the same direction is that high income regions may have amenities that are especially attractive for high skilled workers. Such concerns should be much attenuated by our panel data approach, which eliminates all permanent differences across regions. But residual endogeneity of regional schooling levels could lead to inconsistent least-squares estimates. We therefore implement a two-stage least-squares estimation approach with the beginning-of-sample population structure as an instrument for the change in regional schooling over the following decade(s). The underlying assumption is that a higher share of younger people implies a greater increase in average schooling levels in the region. To verify whether this has been the case empirically, we regress the log change of the population share with more than secondary school over the 1981-1991 and the 1991-2001 period on the share of the population ages 0-19 and 45-70 in 1981 and 1991 perspectively. The results in Table 4 show that the signs of the two age groups are as expected and highly (jointly) statistically significant (see also the F-statistic at the bottom of the table).

[table 4 about here]

5 Estimation and Results

We first discuss the results for the Mincerian approach, and then turn to the results of the constant composition approach.

5.1 The Mincerian approach

Table 5 contains the results of the first step wage regressions of the Mincerian approach. According to the table, the individual return to an additional year of schooling was 7.4% in 1981, 6.3% in 1991, and 5.6% in 2001. The return to education and experience are very similar to those found by others studies (e.g. Abadíe (1997)).

[table 5 about here]

The region-year specific fixed effects in Table 5 can now be used to obtain the change in "cleaned" average regional wages necessary to implement the second step of the Mincerian approach. The results are in Table 6. Columns (1) to (4) contain the results for the 1981-2001 period. Columns (5) to (8) show the results where we pool the 1981-1991 and the 1991-2001 period, with different intercepts for each decade to permit for different trends. . Columns (1), (3), (5) and (7) are obtained using OLS, while columns (2), (4), (6) and (8) are obtained using two-stage least-squares (2SLS). The strength of the externality for the pooled 1981-1991 and 1991-2001 period range from 0.153 to 0.261, with estimates statistically significant at the 10% and the 5% level. As we have two instruments per decade, we can test an overidentifying restriction, which we find cannot be rejected at the usual significance levels. The results for the 1981-2001 period yield highly statistically significant externalities of a similar magnitude, between 0.160 and 0.238. Hence, the Mincerian approach yields statistically significant externalities. One way to get a sense of the magnitude is to note that the share of college and secondary educated workers rose an average of 5,22% annually between 1981 and 2001. For example, taking (2) and (6) as our preferred estimations, the strength of externalities imply that 0.97-1.07% of annual wage growth are explained by schooling externalities.

[table 6 about here]

5.2 The constant composition approach

The results of the first step regression by schooling level are in Table 7.

[table 7 about here]

In the second step we first aggregate the region-year and schooling level fixed effects to obtain region-year average wages holding the labor-force composition constant. In particular, we obtain $\log(\hat{w}_{rt}^{F}) = \log\left(\sum_{s=1}^{S} l_{srT}\alpha_{srt}\right)$ where l_{srT} is the average of the shares at the end points of the time period analyzed. The results of the second step of the constant composition approach are reported in Table 8. Columns (1)-(4) refer to the 1981-2001 period, while the remaining columns refer to the pooled 1981-1991 and 1991-2001 periods.

[table 8 about here]

Our results show significant schooling externalities, which are robust to the different methods used as well as the period we look at. The strength of the estimated externalities range from 0.134 to 0.214 for the 1981-2001, and 0.154 to 0.235 for the pooled pediod. The 2SLS coefficients are significant at the 5% level if we consider the longest possible period and the 5-10% level if we pool the 1980s and 1990s. As expected, externalities are weaker with the constant composition approach than the Mincerian approach. Again, our preferred estimations of the strength of externalities gives us a real wage growth around 0.80-1.00%. This amounts to about a half of real wage growth over this period.[6]

To compare our results with previous results obtained for Spain, for instance by De la Fuente and Domenech (2005) and Alcala and Hernandez (2005), we can do the following exercise. According to our results, we estimate the effect of human capital growth in Spain to be a yearly growth in real wages of 0.8%. This amounts to a total increase of 16% in real wages for the 1981-2001 period. On the other hand, the growth of our human capital proxy implies an increase of 4 years of schooling. Therefore the aggregate return attributable to externalities of one more year of schooling would be 4%. If we assume an average private return between 6% and 7%, we would have a social return betwen 10% and 11%. Our results are therefore similar to previous results.

5.3 Comparison of results: significance and bias

In the first and third column of table 9 the estimated coefficients using the Mincerian and constant composition approach, respectively, are presented, while the standard errors can be found in the second and fourth column. The t-statistic for the difference can be found in the fifth column, with the corresponding p-value in the fifth column. The sixth and final column shows the point estimate of the bias. As it can be verified at the table, the difference is not significant except for specifications 1 and 3, while specifications 2 and 4 are close to significant at the 10% level. Despite the fact that for the rest of specifications it cannot be assured that there is a positive and significant bias of the Mincerian approach, this may be attributed to a lack of power in the estimation (and consequently in the test).

The bias is larger for the first four specifications, with an average bias of 17.6%. For the rest of specifications the average bias it 5.6%. This has to be compared with the expected bias of 25% (see section 3.3).

6 Conclusions

The strength of human capital externalities in Spain is important for growth accounting and from a public-policy perspective. We have applied two different approaches to quantify the wedge between the social and the private return to schooling at the regional level for the 1981-2001 period. Our results yield evidence of significant schooling externalities. A growth accounting exercise yields that approximately a half of the average annual growth

[6]It is interesting to note that our estimates are closer to those of Moretti (2004a).

rate of wages over this period can be accounted for by schooling externalities. Ambos métodos ofrecen estimaciones puntuales diferentes pero sin embargo, dado el bajo sesgo estimado en función de la fóirmula desarrollada por Ciccone y Peri (2006) y por la escasa potencia de las estimaciones, estas diferencias no son estad´siticamente significativas. Future research could combine our estimates with the Spanish tax system and education subsidies to examine whether the incentives to human capital accumulation reflect social returns.

References

Abadíe, A. (1997): "Changes in Spanish Labor Income Structure During the 1980's: A Quantile Regression Approach," *Investigaciones Económicas*, XXI(2), 253–272.

Acemoglu, D., and J. Angrist (2001): "How Large are the Social Returns to Education: Evidence from Compulsory Schooling Laws," in *NBER Macroeconomic Annual 2000*, ed. by B. Bernanke, and K. Rogoff, pp. 9–59. Cambridge, MA: The MIT Press.

Alba, R., and M. S. Segundo (1995): "The Return to Education in Spain," *Economics of Education Review*, 14(2).

Alcalá, F., and P. J. Hernández (2005): "Las externalidades del capital Humano en las Empresas Españolas," Revista de Economía Aplicada (forthcoming).

Angrist, J. (1995): "The Economic Returns to Schooling in the West Bank and Gaza Strip," *American Economic Review*, (85), 1065–1087.

Barceinas, F., J. Oliver, J. Raymond, and J. Roig (2000): "Los Rendimientos de la Educación en España," *Papeles de Economía Española*, (86).

Black, D., and V. Henderson (1999): "A Theory of Urban Growth," *Journal of Political Economy*, 107(2), 252–284.

Card, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, pp. 1801–1864. Elsevier, Amsterdam.

Ciccone, A., and G. Peri (2006): "Identifying Human-Capital Externalities: Theory with Applications," *Review of Economic Studies*, 73(2), 381–412.

De la Fuente, A. (2003): "Human Capital in a Global Knowledge-based Economy, part II: assesment at the EU Country Level," European Commission Report, Employment and Social Affairs.

De la Fuente, A., R. Domenech, and J. F. Jimeno (2003): "Human Capital as a Factor of Growth and Employment at the Regional Level. The Case of Spain," UFAE and IAE Working Papers 610.04.

García-Fontes, W., and M. A. Hidalgo (2008): "¿Es posible estimar las externalidades del capital humano? Evidencia para las regiones españolas," *Temas actuales de economía*, 2, 49–71, Instituo de Análisis Económico y Empresarial de Andalucía.

Hidalgo, M. A. (2009): "A Demand and Supply Analysis of the Spanish Education Premium in the 1980s and 1990s," Discussion paper, Forthcoming *Revista de Economía Aplicada*.

Hidalgo, M. A., J. O'Kean, and J. Rodríguez (2008): "Labor Demand and Information Technologies: Evidence for Spain, 1980-2005," Working paper 08.12, Universidad Pablo de Olavide, Sevilla, Spain.

Katz, L. F., and K. M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107(1), 35–78.

Lucas, R. J. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22(1), 3–42.

Moretti, E. (2004a): "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data," *Journal of Econometrics*, 121(1-2), 175–212.

——— (2004b): "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," *American Economic Review*, 94(3), 656–690.

Raymond, J. (2002): "Convergencia Real de las regiones españolas y Captal Humano," *Papeles de Economía Española*, 93, 109–121.

Rudd, J. (2000): "Empirical evidence on human capital spillovers," Finance and Economics Discussion Series 2000-46, Board of Governors of the Federal Reserve System (U.S.).

Table 1: Schooling groups equivalence. Individual data

	EPF 81 survey groups	EPF 91 survey groups	ECPF 01 survey groups
Illiterate-Primary	1,2,3	1,2,3,4	1,2
Basic secondary and basic vocational	4	5,7	3
Advanced secondary and advanced vocational	5,6	6,7,8	4,5
College (short cycle)	7	9	6
College (long cycle)	8	10	7

Note:

EPF survey: Encuesta de presupuestos familiares.

ECPF survey: Encuesta Continua de presupuestos familiares

Definition of the schooling groups:

Groups EPF 81: 1 - Illiterate, 2 - No degree, 3 - Primary,
4 - Basic secondary, 5 - Advanced secondary,
6 - Vocational training, 7 - College (short cycle), 8 - College (long cycle)

Groups EPF 91: 1 - Illiterate, 2 - No degree, 3 - Basic primary,
4 - Advanced primary, 5 - Basic secondary ,6 - Advanced secondary,
7 - Basic vocational training, 8 - Advanced vocational training,
9 - College (short cycle), 10 - College (long cycle)

Groups ECPF 01: 1 - Illiteracy or no degree, 2 - Primary,
3 - Basic secondary and basic vocational training,
4 - Advanced secondary and advanced vocational training,
5 - Other advanced secondary studies,
6 - College (short cycle), 7 - college (long cycle)

Table 2: Average wages for Spanish regions (in pesetas, 1980-2001)

	1981	1991	2001
Andaluca	620.771	1.374.273	2.006.496
Aragn	726.164	1.552.504	2.404.515
Asturias	770.411	1.615.257	2.505.721
Baleares	670.676	1.460.103	2.191.425
Canarias	644.695	1.387.754	2.020.738
Cantabria	711.886	1.596.316	2.242.165
Castilla y Leon	690.036	1.567.936	2.155.668
Castilla-La Mancha	626.355	1.419.734	2.010.670
Catalua	779.218	1.709.741	2.393.477
Comunidad Valenciana	675.078	1.389.134	2.207.454
Extremadura	541.652	1.320.384	1.760.094
Galicia	683.013	1.452.852	2.153.270
Madrid	815.915	1.657.278	2.468.719
Murcia	597.704	1.388.218	1.942.676
Navarra	801.067	1.777.856	2.683.833
Pas Vasco	803.115	1.758.933	2.381.716
La Rioja	697.790	1.531.849	2.183.584
Espaa	694.516	1.514.939	2.232.096

Table 3: Average wages by schooling groups (in pesetas, 1980-2001)

		1981	1991	2001
Primary or less	mean	612.622	1.266.508	1.932.901
	stdv	278.190	584.631	652.061
Secondary	mean	821.892	1.721.526	2.395.398
	stdv	324.575	1.182.154	886.831
College	mean	1.061.663	2.323.486	2.842.428
	stdv	371.898	938.236	1.027.295
All	mean	694.516	1.514.939	2.232.096
	stdv	330.034	878.318	884.291

Table 4: First stage regression

Dependent variable:
Log change in fraction of workers with secondary or college education.

Age Interval	Coefficient
0-19	8.92 (1.20)
45-70	9.78 (2.19)
Constant	-2.91 (0.57)
Adjusted-R^2	0.69
F	36.78
Prob>F	0.00
n	34

Note: This regression estimates the relationship between our regional human capital intensity proxy (LED) and the instruments. Standard errors in parentheses.

Table 5: Mincer Regressions

	1981		1991		2001	
Aragon	0.120		0.161		0.106	
	(0.019)		(0.022)		(0.030)	
Asturias	0.145		0.167		0.133	
	(0.024)		(0.034)		(0.028)	
Baleares	0.036	**	0.110		0.100	
	(0.031)		(0.031)		(0.036)	
Canarias	0.030	**	-0.004	**	-0.028	**
	(0.022)		(0.025)		(0.030)	
Cantabria	0.079	*	0.128		0.021	**
	(0.027)		(0.037)		(0.038)	
Castilla y Leon	0.058		0.109		0.001	**
	(0.014)		(0.016)		(0.025)	
Castilla-La Mancha	0.015	**	0.084		0.008	**
	(0.017)		(0.019)		(0.029)	
Catalunya	0.192		0.181		0.114	
	(0.015)		(0.019)		(0.021)	
Com. Valenciana	0.098		0.078		0.079	
	(0.016)		(0.018)		(0.022)	
Extremadura	-0.167		-0.041	**	-0.098	
	(0.022)		(0.027)		(0.031)	
Galicia	0.024	**	0.064		-0.006	**
	(0.018)		(0.019)		(0.027)	
Madrid	0.184		0.129		0.101	
	(0.017)		(0.023)		(0.021)	
Murcia	0.011	**	0.042	**	-0.047	**
	(0.030)		(0.030)		(0.031)	
Navarra	0.244		0.263		0.211	
	(0.032)		(0.034)		(0.035)	
Pais Vasco	0.238		0.217		0.100	
	(0.018)		(0.019)		(0.027)	
Rioja (La)	0.111		0.029	**	0.035	**
	(0.033)		(0.034)		(0.035)	

Continued on next page

Table 5–Continued from previous page

	1981	1991	2001
Female dummy	-0.210	-0.253	-0.168
	(0.017)	(0.016)	(0.016)
Works in agriculture dummy	-0.450	-0.435	-0.212
	(0.014)	(0.019)	(0.028)
Married	0.361	0.086	0.033**
	(0.026)	(0.025)	(0.033)
Years of schooling	0.074	0.063	0.056
	(0.001)	(0.001)	(0.001)
Experience	0.019	0.034	0.029
	(0.001)	(0.001)	(0.002)
Experience square	0.000	-0.001	0.000
	(0.000)	(0.000)	(0.000)
Constant	11.969	12.908	13.472
	(0.039)	(0.040)	(0.045)
n	11402	8838	3751
Adjusted R^2	0.32	0.31	0.28

Standard errors in parentheses.

Statistically significant at the 1% level unless otherwise noted:

* Statistically significant at the 10% level

** Insignificant

Table 6: Mincerian approach estimation

Dependent variable: Change in regional effects

	1981-2001				1981-1991-2001			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LED	0.160	0.187	0.208	0.238	0.153	0.206	0.192	0.261
	(0.056)*	(0.071)*	(0.068)*	(0.083)*	(0.076)**	(0.089)*	(0.094)**	(0.110)*
1990s dummy	-	-	-	-	-0.372	-0.350	-0.355	-0.326
	-	-	-	-	(0.035)	(0.039)	(0.042)	(0.048)
Employment	-	-	0.171	0.169	-	-	0.080	0.098
	-	-	(0.149)***	(0.153)***	-	-	(0.104)***	(0.106)***
Capital	-	-	0.042	0.180	-	-	0.002	-0.032
	-	-	(0.144)***	(0.152)***	-	-	(0.158)***	(0.162)***
ICT Capital	-	-	-0.088	-0.145	-	-	-0.032	-0.040
	-	-	(0.131)***	(0.140)***	-	-	(0.092)***	(0.093)***
Constant	1,320	1,291	1,294	1,214	0.849	0.811	0.837	0.803
	(0.060)	(0.075)	(0.106)	(0.124)	(0.056)	(0.065)	(0.080)	(0.085)
Observations	17	17	17	17	34	34	34	34
R-squared	0.35	0.32	0.45	0.42	0.96	0.96	0.96	0.96

Note:

Standard errors in parentheses

LED: growth rate of the share of workers with secondary degrees or higher

1990s dummy: dummy equal to 1 for the 1991-2001 period

Employment: Log change of regional employment.

Capital: Log changes of regional physical capital.

ICT Capital: Log changes of regional ICT physical capital.

Models:

(1)-(4): pooled regressions for periods 1981-1991 and 1991-2001

(5)-(8): regressions for 1981-2001

(1),(3),(5),(7): OLS estimations

(2),(4),(6),(8): two-stage least-squares estimations.

Statistically significant at the 1% level unless otherwise noted:

* Statistically significant at the 5% level

** Statistically significant at the 10% level

*** Statistically insignificant

Table 7: First step wage regressions by schooling level

Independent variable: individual wages	Year	**Regression 1: Primary or less**		**Regression 2: Secondary**		**Regression 3: College**	
Andaluca	91	0.663		0.752		0.719	
	01	1.356		1.311		1.123	
Aragon	81	0.147		0.119		0.026	***
	91	0.904		0.878		0.667	
	01	1.482		1.361		1.293	
Asturias	81	0.193		0.073	***	0.122	
	91	0.932		0.804		0.733	
	01	1.563		1.332		1.261	
Baleares	81	0.042	***	-0.024	***	0.301	
	91	0.836		0.877		0.558	
	01	1.488		1.412		1.148	
Canarias	81	0.027	***	0.072	***	0.073	***
	91	0.671		0.776		0.715	
	01	1.358		1.159		1.217	
Cantabria	81	0.149		0.043	***	-0.158	**
	91	0.882		0.705		0.886	
	01	1.452		1.236		1.102	
Castilla y Leon	81	0.105		0.017	***	-0.053	***
	91	0.835		0.789		0.733	
	01	1.382		1.250		1.176	
Castilla-La Mancha	81	0.023	***	0.075	**	-0.108	**
	91	0.787		0.796		0.743	
	01	1.401		1.271		1.172	
Catalunya	81	0.228		0.183		0.128	
	91	0.902		0.855		0.829	
	01	1.476		1.456		1.218	
Com. Valenciana	81	0.122		0.077	**	0.106	**
	91	0.805		0.729		0.660	
	01	1.450		1.421		1.220	
Extremadura	81	-0.198		0.073	***	-0.087	***
	91	0.663		0.669		0.629	
	01	1.301		1.172		1.131	

Continued on next page

Table 7 –continued from previous page

Independent variable: individual wages	year	Regression 1 Primary or less		Regression 2 Secondary		Regression 3 College	
Galicia	81	0.016	***	0.090	*	0.054	***
	91	0.768		0.778		0.734	
	01	1.355		1.295		1.155	
Madrid	81	0.218		0.186		0.124	
	91	0.873		0.811		0.715	
	01	1.493		1.357		1.241	
Murcia	81	0.050	***	-0.074	***	-0.172	**
	91	0.743		0.746		0.732	
	01	1.355		1.232		1.056	
Navarra	81	0.286		0.241		-0.042	***
	91	1.026		0.913		0.851	
	01	1.628		1.502		1.284	
Pas Vasco	81	0.270		0.212		0.189	
	91	0.968		0.844		0.802	
	01	1.527		1.366		1.156	
Rioja (La)	81	0.166		0.106	***	-0.119	***
	91	0.784		0.577		0.703	
	01	1.391		1.330		1.268	
Female dummy		-0.278		-0.179		-0.125	
Works in agriculture dummy		-0.407		-0.311		0.027	***
Married		0.168		0.231		0.229	
Experience		0.030		0.027		0.025	
Experience square		0.000		0.000		0.000	
Constant		12.559		12.932		13.364	
Adjusted R^2		0.60		0.51		0.57	
n		15,930		4,496		3,274	

Note: Standard errors are not reported.

Statistically significant at the 1% level unless otherwise noted:

* Statistically significant at the 5% level

** Statistically significant at the 10% level

*** Insignificant

Table 8: Constant Composition approach estimation.

Dependent variable: log changes in weighted average wages (weights are for year 1990 for 1981-2001 estimation and middle year for 1981-1991-2001 estimation).

	1981-2001				1981-1991-2001			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
LED	0.134	0.167	0.163	0.214	0.154	0.192	0.183	0.235
	[0.056]*	[0.069]*	[0.071]*	[0.079]*	[0.081]**	[0.094]*	[0.101]**	[0.118]**
1990s dummy	-	-	-	-	-0.000	0.016	0.012	0.033
	-	-	-	-	[0.037]***	[0.042]***	[0.045]***	[0.051]***
Employment	-	-	0.121	0.134	-	-	0.021	0.034
	-	-	[0.155]***	[0.147]***	-	-	[0.112]***	[0.114]***
Capital	-	-	0.093	0.214	-	-	-0.023	-0.048
	-	-	[0.149]***	[0.145]***	-	-	[0.170]***	[0.173]***
ICT Capital	-	-	-0.108	-0.169	-	-	-0.037	-0.043
	-	-	[0.136]***	[0.135]***	-	-	[0.098]***	[0.099]***
Constant	1,127	1,092	1,105	1,017	0.553	0.525	0.565	0.540
	[0.060]	[0.073]	[0.110]	[0.119]	[0.060]	[0.069]	[0.086]	[0.091]
Observations	17	17	17	17	34	34	34	34
R-squared	0.27	0.28	0.34	0.41	0.37	0.36	0.38	0.37

Note:

Standard errors in parentheses.
LED: growth rate of share of workers with secondary degrees or higher.
1990s dummy: dummy equal to 1 for the 1991-2001 period.
Employment: Log change of regional employment.
Capital: Log changes of regional physical capital.
Models:
(1)–(4): pooled regressions for 1981-1991, and 1991-2001 periods.
(5)–(8): regressions for the 1981-2001 period.
(1),(3),(5),(7): OLS estimations.
(2),(4),(6),(8): two-stage least-squares regressions.
Statistically significant at the 1% level unless otherwise noted:
* Statistically significant at the 5% level
** Statistically significant at the 10% level
*** Insignificant

Table 9: Bias due to using Mincerian Approach instead of the Constant Composition Approach.

	Mincerian		Constant Composition				
Regression	Coefficient	Stand. error	Coefficient	Stand. error	t	P-value	Bias
(1)	0.16	0.06	0.13	0.06	1.914	0.0300	19.40
(2)	0.19	0.07	0.17	0.07	1.177	0.1215	11.98
(3)	0.21	0.07	0.16	0.07	2.669	0.0048	27.61
(4)	0.24	0.08	0.21	0.08	1.221	0.1132	11.21
(5)	0.15	0.08	0.15	0.08	-0.050	0.5290	-0.65
(6)	0.21	0.09	0.19	0.09	0.631	0.2652	7.29
(7)	0.19	0.09	0.18	0.10	0.383	0.3525	4.92
(8)	0.26	0.11	0.24	0.12	0.940	0.1754	11.06

The regressions are those specified in tables 6 and 8