The Distribution of Human Capital and Regional Productivity

Charisse Tubianosa Joson
MSc in Economics, Universitat de Barcelona

Advisor: Enrique López-Bazo

Abstract

This paper takes a novel approach in estimating the effect of human capital on productivity using a panel data for the NUTS 3 Spanish regions between the years 1996 and 2008. The hypothesis is that, besides the average level of education, how education is distributed in the population also matters to productivity level and growth. We use available data on the share of the population in each level of education to compute measures of inequality and skewness of the distribution in each region. Our findings suggest that inequality in educational attainment and, in particular, the degree of skewness of its distribution matter in explaining the variation in regional productivity.

JEL classification: O11, O15, O18, O47
Keywords: human capital, education, productivity, distributional impact, inequality, skewness, average educational attainment, Gini

Acknowledgements: I am indebted to Enrique López-Bazo for his incredible patience, for sharing his expertise and especially his time, for the invaluable feedback and suggestions—all of which have guided me and have kept me relatively calm throughout this process. I would also like to thank Raul Ramos for going out of his way to help me with my concerns on inequality computations, and Vicente Royuela for enlightening me on various STATA issues. A load of thanks also goes to Jordi Roca who is ever so prompt to his replies to my never-ending list of questions. I am very grateful to the chicas: Franzi, Camilla, Ha, and Neritza for regularly reminding of the world outside, to Jannick and Stefano, for always replying to my queries, to Roger for being an external reader and sharing his insights, and to John who has been a wall to lean on from the beginning of this endeavor. I am especially grateful to my family's support and for being my cheerleaders, strongly felt even across the miles, and to Pilar and Ramon who have watched over me and made me feel like family. A special mention goes to Oriol, my anchor, for his infinite patience, encouragement, care, and support. Finally, many thanks also to the UB Economics Faculty for the knowledge and the skills they have imparted to me and to the batch.

I. Introduction

Many studies have previously shown the importance of human capital in explaining growth. Investment in human capital, such as education and nutrition for instance, matters for a person's well-being and ability to improve productivity. For the most part, these studies have only considered the average years of schooling, and only a few have given attention to the amount of inequality in educational attainment of the population as possibly having an effect on growth. In this paper, we emphasize that measuring the impact of education in the aggregate level of productivity requires taking its distributional facets into consideration. Furthermore, this issue brings to bear significant implications for public policy. The revised Lisbon Strategy, with the goal of improving competitiveness in the EU, has identified investments in human capital as one of its important features. With a broader scope, human capital development featured prominently in the United Nations' Millenium Development Goals, so much so that many developing economies set aside considerable public funds for education and health care. Formal education has been the embodiment of human capital. Spain in particular has made significant inroads in providing more education to its population in the past two decades.[1] In line with this, it would be interesting to look into the impact of human capital at the subnational level, and what the results may imply for regional policy.

Ever since Becker (1962) highlighted the importance of human capital in the form of education and training in determining income growth, the literature has been rife with contributions to test and measure this relationship using average educational attainment as a measure for human capital (Benhabib and Spiegal, 1994). In recent years, a number of papers have looked beyond the average level of human capital accumulation and have also focused on inequality in educational attainment, similar in analysis to how inequality in physical capital may adversely affect income, productivity, and growth. For instance, De Gregorio and Lee (2002), using Barro and Lee's (1996) educational attainment data set, find cross-country evidence on the negative correlation between income inequality and education inequality. Other authors have also tested similar concerns at the regional level (Becker and Chiswick, 1966; Rodríguez-Pose and Tselios, 2009; Gille, 2015 are some examples), and have found that more unequal distribution of educational attainment in a population is associated positively with income inequality, and negatively with income per capita.

There are other characteristics of the distribution to be explored in analysing human capital's relationship with productivity and growth. The degree of skewness may also convey valuable information. We speculate for instance how a positive skew, meaning the median of the distribution is located below the average level, would affect productivity. However, to the author's knowledge, there are currently no studies looking specifically into how skewness of the distribution relates with productivity. Thus, the aim of this paper is to contribute to this gap by providing evidence of how two aspects of human capital distribution, namely inequality and the degree of skewness, affect

[1] A number of reforms, or education acts, were implemented since 1990, providing the framework for education at the primary, secondary, non-university higher education, and at the university level (OECD, 2006).

productivity growth, in addition to measuring the effect of the average level of human capital. As indicated in the literature, ceteris paribus, higher inequality in education is associated with lower productivity and growth. But what outcome do we expect from the degree of skewness? Studies that have looked at composition have found that economies or regions with a higher proportion of skilled labor, as well as having a higher proportion of the population with tertiary level education, are associated with higher productivity (e.g., Krueger and Lindahl, 2001; Fleisher, Li and Min, 2010). Based on this, we formulate a hypothesis that, given a level of inequality, a positive skew would have a negative effect on productivity. Conversely, if the median of the distribution of education in a region is higher than the average, or more have a level of education higher than the average—a negative skew—would have a positive effect on productivity.

More formally, this paper tests the following hypotheses using an aggregate production function: (1) a higher level of educational attainment has a positive impact on productivity; (2) a higher level of education inequality has a negative effect; and finally (3) the degree of skewness is inversely related to productivity. We construct a panel data for 50 Spanish provinces over the period 1996 to 2008, at four-year intervals to estimate the impact of these measures of education on regional productivity.

Using static and dynamic panel models, our results confirm findings in previous studies that indicate that the average years of schooling of the population of the region has a significant impact on productivity. More importantly, we provide evidence that distributional characteristics do matter. In particular, our results confirm that the degree of skewness of the distribution is indeed inversely proportional to the productivity level and growth. We find that a one standard deviation increase in skewness in a region reduces GDP per worker by 2.4%. To be clear, an increasingly positive skew, or having more of the population with years of schooling below the average, contributes negatively to productivity in the Spanish provinces.

Our results further show that the estimated coefficient is highly significant in both the static and dynamic models. However, the magnitude of the coefficient estimate is smaller when relating to productivity growth versus productivity in levels. In specifications where we combine all three human capital variables, the measure of inequality loses significance. This is not a surprising result given that all three measures are highly correlated and thus have a tendency to absorb, or deflect, each other's effects. But the results for the degree of skewness are robust even when we add our set of control variables that includes the initial level of productivity, public capital, density, sectoral composition of employment, and the employment rate. We also note that the estimated coefficient measuring the return to private physical capital provides a reasonable value particularly in the static models. Though test statistics support the empirical results from the dynamic models, these need to be taken with caution given the moderate number of observations in the time dimension that we have in our data set.

This paper opens new avenues for analysing how human capital impacts on productivity. In addition, these preliminary findings have important implications for the formulation of regional

policy. Given the primacy of the role given to human capital in regional public policy in contributing to sustaining growth, increasing competitiveness, and closing gaps in regional disparities, our results suggest that taking a closer look at how education is distributed is warranted. Tailoring programs based on distributional characteristics of regions may address their development needs more effectively.

The remainder of the paper is organized as follows: section II provides a brief overview of the theoretical and empirical underpinnings of this study, section III discusses the empirical strategy, section IV describes the data and variables to test the hypotheses, section V shows the empirical results, and section VI concludes.

II. Literature Review

a. Theoretical Framework

The literature is vast and the general point of departure is the inclusion of human capital as an input in an aggregate production function. This is based on the concept that knowledge and skills, whether these are innate or acquired, have a direct contribution to increasing incomes, in terms of private returns, and raising productivity (Becker, 1962; De la Fuente, 2011) thereby allowing an economy to develop further by adopting to and imitating new technologies, and innovating (Uzawa, 1965; Aghion and Howitt, 1985; Lucas, 1988; Romer, 1990). In this fashion, human capital is seen as being able to mitigate diminishing marginal returns that arise from accumulating physical capital, including land, and foster sustained growth in the long-term (Nelson and Phelps, 1966; Romer, 1990; Becker, 2008).

In choosing an indicator for human capital, the most common has been based on education. Education (and training) has been touted as one of the most important investments in human capital (Becker, 1962, 2008). Human capital has been characterized as having two aspects, the part that is innate and given, and the part that may be developed over time. Education is one important way to link these two, allowing individuals to compensate for the gaps from the initial conditions (Sauer and Zagler, 2014). The importance of education supports the view that human capital is essential for adopting and adapting to evolving technologies, and in turn help in shaping its development. This capacity heavily depends on the level of education of the active members of the population (Griffith, Redding, and Van Reenen, 2003).

The literature has also emphasized the importance of accounting for the initial stock of human capital, and that secondary and post-secondary education matter more than primary (Krueger and Lindahl, 2001). According to Barro (2001), a higher initial stock of human capital reflects a higher ratio of human capital to physical capital and that this may lead to higher growth, for one, because higher stock of human capital facilitates adoption of more advanced technologies, and for another, because adjustments to human capital have a longer gestation period than adjusting the stock of physical capital. Given this, when an economy or a region starts off with a higher stock of human capital, vis-à-vis physical capital stock, it would then have the capacity to grow faster by increasing the quantity of physical capital.

However, accounting for the stock of human capital by means of the average years of schooling of the population may not be sufficient in describing its effect in determining output. The distribution of human capital, here education, also matters. Reference to the equitable, or inequitable, distribution of education is likened to land and machinery as capital. According to Thomas et al. (2003): "*An equitable distribution of human capital (such as basic literacy and health or nutrition) constitutes a precondition for individual productivity and the ability to rise above poverty.*" But this does not mean that the measure of education inequality is a substitute for average educational attainment as an indicator. Rather, variables related to the distribution of education should act as complements to the average level of educational attainment, providing a clearer view of how human capital impacts on output and productivity.

In general, why is the distribution important? Lopez et al. (1998) and Thomas et al. (2003) also explain this succinctly in terms of welfare consideration and for production. Classically, when an asset (i.e. physical capital) is traded through free, competitive markets, its marginal product is equalized. If that is the case, Thomas et al. (2003) state that: "the asset's contribution to output should not be affected by its distribution across firms and individuals. However, if the asset is not exchanged in such a manner, then the marginal product is not equalized, and an aggregation problem will arise. This has a huge implication on the use of the aggregate production function to estimate the effect of human capital. Education has been described as only partially tradable, so the average level of educational attainment is not sufficient to reflect its contribution to production." Thus it will also depend on how education is distributed across the population. On top of this, education externalities and credit market constraints aggravate the effects of significant human capital distribution inequalities on growth and development. In estimating the effect of education, Sauer and Zagler (2014) point out that if the distributional characteristics are not accounted for, a negative correlation between educational inequality and average attainment results in a downward bias on the estimated effect of the latter.

If the inequality of education, commonly measured as a Gini index, has not been given much attention, the degree of skewness of its distribution, to our knowledge, has received none. Whereas the Gini measures the dispersion of the distribution, the skewness looks at the degree of asymmetry of the distribution. This is important to note because two provinces may share the same value of the Gini, or any other measure of dispersion, and the shape of the distribution may very well vary. A distribution is positively skewed if the longer tail is to the right, or in the positive direction, and the bulk of the data is found on the left. The opposite is true for a negatively skewed distribution. In the context of this paper, a positive skew of the education variable means that the bulk of the population in the region has accumulated a relatively lower level of years of schooling relative to the average, whereas a small fraction shows much higher levels. Conversely, a negative skew means that a bigger proportion of the population has acquired more years of schooling compared with the average, with only a small fraction showing lower levels of education. In a way, this may be used to explain why a larger

proportion of the population with higher level of schooling (post-secondary for instance) is expected to have a positive effect on output and growth, especially when relating to the arguments mentioned above in terms of building capacity, adapting to technology, and innovating.

b. Empirical Evidence

The evidence on the importance of human capital in determining growth and productivity has been emphasized, refuted, exonerated, and excoriated in turn. Benhabib and Spiegal (1994) were among the first to cast a doubt on the strength of the relationship between human capital and growth. Prichett (2001) claimed that human capital investment in itself has not improved well-being in less developed countries. But these are interspersed with findings that have supported the role of human capital, particularly education. Some examples include Jenkins (1995) who found education to have a positive and significant effect on labour productivity growth in the UK. Black and Lynch (1996) have also provided evidence that the impact of education in the non-manufacturing sector is particularly strong. According to Moretti (2004), human capital increases aggregate productivity, not just individual productivity, and more recently Vogel (2013) has contributed with findings that human capital has a positive impact on productivity growth through innovation channels.

Succeeding authors have sought to explain the observed weakness of the effect of human capital in earlier studies (i.e., Temple, 2001; De la Fuente and Domenech, 2006; Cohen and Soto, 2007; and De la Fuente, 2011; Barro and Lee, 2013). The main culprit has been data quality since education data is prone to measurement errors. This is more evident in cross-country studies where educational standards and the method of collection vary. More recent findings have managed to show that the significance of human capital may be attributed to having better data, and better tools for econometric analysis.

Beyond average years of schooling, other studies have also put more attention to the distributional aspects of educational attainment and how they affect economic development, using either the standard deviation of the enrollment rates, the dispersion in the years of schooling, or the composition of human capital (for example, proportion of the population that has completed tertiary education, or classifying skilled and non-skilled based on employees' years of educational attainment in). Similar to income and wealth, the distribution of education across society may work as a boost or a drag to development. For instance, Birdsall and Londoño (1997) found a negative correlation between education dispersion and income growth. In looking at composition, skilled labour tends to increase growth and productivity (Vandenbussche and Aghion, 2006; Fleisher et al. 2010). Measures of dispersion of educational attainment have also been used with other indicators for development. These studies have found a positive relationship between education inequality and income inequality, and a negative one between education inequality and income level in cross-country (Becker and Chiswick, 1966; De Gregorio and Lee, 2002) and regional contexts (Rodríguez-Pose and Tselios, 2009; Gille, 2015). In fact, Castello and Domenech (2002) find evidence that measures of human capital inequality are more robust in explaining growth than income inequality.

The degree of skewness of the distribution of education may also have something to say on the matter. Studies that have looked at composition provide clues to this argument. However, to the author's knowledge, there are no studies specifically linking this characteristic of the distribution of education with productivity. The concept thus far seems to have only been applied in relation to income distribution, where, for instance, the shape, whether a positive or a negative skew may determine redistribution efforts of governments (Dalgaard, Hansen and Larsen, 2003).

III. Empirical Model

The most common point of departure for human capital and growth models is based on the textbook Solow model using a Cobb-Douglas production function, then introducing human capital as an additional input (e.g., Mankiw, Romer, and Weil, 1992):

$$(1) \quad Y_{r,t} = A_{r,t} K_{r,t}^{\beta_1} H_{r,t}^{\beta_2} L_{r,t}^{\beta_3}$$

where, Y_{rt} is the aggregate output of region r at year t, A_{rt} is the total factor productivity (TFP), an index of technical efficiency, K_{rt} is the stock of physical capital, and H_{rt} represents the variable related to human capital stock, which depends on a measure of the educational attainment of the population. L_{rt} is the level of employment. From this equation, we can define the relationship between labor productivity, the key human capital variables, and physical capital. Under the assumption of constant returns to scale, and $\beta_3 = 1 - \beta_1 - \beta_2$, the production function may be expressed in logarithms, as follows:

$$(2) \qquad \log(Y/L)_{r,t} = A_{r,t} + \beta_1 \log(K/L)_{r,t} + \beta_2 H/L_{r,t}$$

$$with\ H/L = e^{educ}$$

and,

$$(3)\ A_{r,t} = \theta_r + \delta X_{r,t} + \epsilon_{r,t}.$$

where $X_{r,t}$ is a set of control variables that consists of public capital (Aschauer, 1989), measured as the stock of public capital per surface area, log of population density, employment shares in agriculture and industry to account for sectoral composition (Kaldawei and Walz, 2001), and employment rate as a measure of resource use (Belorgey, Lecat, and Maury, 2004; Crescenzi and Rodríguez-Pose, 2012). θ_r is the region fixed effect and $\epsilon_{r,t}$ is an idiosyncratic error term. Education inequality and the degree of skewness fall under $A_{r,t}$, that is to say these distributional characteristics have an effect on productivity through efficiency.

Alternatively, growth in productivity, can be expressed as follows:

$$(4) \qquad \gamma_{r,t} = logy_{r,t} - logy_{r,t-\tau} = \mathring{A}_{r,t} + \beta_1(logk_{r,t} - logk_{r,t-\tau}) + \beta_2(H_{r,t}^{educ} - H_{r,t-\tau}^{educ})$$

where $y_{r,t} = (Y/L)_{r,t}$ and $k_{r,t} = (K/L)_{r,t}$. t-τ refers to value of variable in the preceding period, and $\mathring{A}_{r,t}$ denotes TFP growth that is assumed to depend on the variables described above with the addition of the initial output per worker to account for convergence in productivity. Education inequality, the degree of skewness, and the control variables refer to initial values at t-τ .

Given that the main focus of the paper is in determining the impact of the distributional characteristics of education, it is important to define precisely the measures of inequality and skewness of education. The Gini index is the most common measure of inequality. In this paper, education inequality is also based on a Gini coefficient following the methodology from Thomas et al. (2003) computed based on the proportion of school attainment data in each region. Previous indicators that have been used to measure education inequality were based on enrollment or education finance. But the authors note that these may be inadequate since they do not reflect the cumulated educational outcome.

Education inequality based on the Gini index for each region *r* is computed as:

$$(5)\ Gini_r^{HC} = \frac{1}{\mu_r}\sum_{i=2}^{n}\sum_{j-1}^{i-1} p_{ir}\ |s_{ir} - s_{jr}|p_{jr}$$

where:

μ_r *- average years of schooling for the concerned population*

p_{ir} *and* p_{jr} *- proportions of the population with levels of schooling, for i and j, respectively*

s_{ir} *and* s_{jr} *- years of schooling at the educational attainment levels, for i and j, respectively*

n- number of levels/categories in attainment data

Thomas et al. (2003) note that unlike computing the Gini for income, a continuous variable, the education Gini, based on years of schooling, is a discrete variable which has a lower boundary of zero and an upper boundary which, for the case of Spain, is about 17 years. As such, the education Lorenz curve looks more like a crooked/bent line with distinct points at each educational level/category. The Gini is akin to a ratio which may take a value between 0 and 1. A Gini coefficient that is close to zero signifies that educational attainment is evenly distributed across the population. A coefficient closer to 1, however, indicates a very unequal distribution.

Dispersion of a distribution is not the only thing we are looking at. Given the position (average education) and the amount of dispersion or inequality (Gini index), we are also interested in a further characteristic of the distribution. In particular, our interest lies on how skewed is the distribution of education and if the bulk of the population concentrates to the right or to the left of the average level of education. The standard procedure is to compute the coefficient of skewness as the standardized third moment of the variable of interest:

$$(6)\ skewness_r = {m_{3r}}\Big/{m_{2r}^{3/2}}$$

where $m_{3r} = \sum_{i\epsilon r}(x_{ir} - \mu_r)^3/n_r$ and $m_{2r} = \sum_{i\epsilon r}(x_{ir} - \mu_r)^2/n$

Similar to the Gini, skewness has no unit reference, unlike the mean or average which follows the unit of the original variable (here, schooling in years). However, unlike the Gini, skewness can take negative values. The distribution may be described as highly skewed if skewness is greater than |1|; moderately skewed if the value is between |1| and |0.5|; and approximately symmetric if the skewness is between -0.5 and 0.5 (Bulmer 1979). To reiterate, a positive skew indicates that the mass of the sample is located in the left side of the distribution. In the context of an education distribution, this means that most of the population has attained levels of schooling that are below the average. If it is negative, then most of the population are situated to the right, and the bulk of the population has attained higher than average years of schooling.

With the construction of these measures of the characteristics of the distribution of human capital, the hypotheses in this thesis are tested using the specifications sketched in (2) and (3). For both types of analyses, we employ a static model with region fixed effects. This ensures that time-invariant elements of productivity that may cause bias are removed. The rationale is that if unobservable time-invariant variables are controlled for, then any variation in the outcome variable must be caused by factors other than these fixed characteristics (Stock and Watson, 2003). However, we do not control for time effects in testing the hypotheses. The idea behind controlling for time effects is to account for shocks, or say the business cycle that may also affect the outcome variable in a similar way in all regions. In our case, the set of control variables chosen will perform this function. As a panel data, the sectoral employment shares will account for structural changes that may have occurred over the study period. Changes in employment rate, or use of labor, also absorb effects related to the business cycle, or any economic shocks in the system. Furthermore, capital may adjust in response to changes in the business cycle given the analysis considers changes every four years in the magnitude of interest. Altogether, these variables serve as proxies for time effects in capturing exogenous changes in productivity common to all regions. The other reason for not including time effects is related with our regressors of interest. There is a common trend among regions with regard to changes in human capital variables over the period under analysis. In varying degrees, regions generally experience increasing levels of educational attainment, decreasing inequality, and perhaps a skew that tends to move to the left side of the distribution. Since our interest is in testing how these variables affect productivity, controlling for time effects will not allow us to identify the effect of the distribution of human capital.

Additionally, for estimating (3) with the inclusion of the level of productivity at the initial period ($logy_{r,t-\tau}$), we make use of a system generalized method of moments (GMM) estimator to obtain consistent estimates of this dynamic panel data model. System-GMM is used to address serial correlation introduced by the inclusion of the lag of the dependent variable as an additional explanatory variable, which would then be correlated with the residuals.

IV. Data, Variables and Descriptive Analysis

We have constructed a panel data for the 50 Spanish provinces (excluding Ceuta and Melilla) between the years 1996 to 2008 at four-year intervals to account for the human capital variables changing slowly over time. Based on the framework provided in the preceding section, the selected variables are described in Table 1. We combined GDP in Purchasing Power Standards and the total number of employed persons from Eurostat to obtain productivity per worker. As the earliest data on GDP levels in Eurostat are only available until the year 2000, we have filled the gap for the earlier years by applying growth rates based on the regional GDP series from the Instituto National de Estadistica Spanish Regional Accounts (INE-SRA). For the human capital variables, we used the data available from The Valencian Institute of Economic Research (IVIE), which provides the proportion of the economically active population at each of the 8 education categories of Spain for every region, and we converted these educational attainment levels into the corresponding years of education to compute the weighted average level of educational attainment, the Gini index, and the degree of skewness. The table showing the equivalent years used for each classification is in Appendix 1.

Data availability constrains the aspiration to test the hypotheses over a longer time period. Given the time constraint to complete this study, harmonizing data from a variety of sources to achieve a balanced panel was only readily available for the time period from 1996 to 2008. In addition, we cap the analysis at 2008, prior to the manifestation of the effects of the Global Financial Crisis (GFC) of 2008-2009. Spain in particular suffered the brunt of the GFC after 2008, and the difficulties mounted following the European Debt Crisis of 2012-2013. Though this subsequent period would have made for an additionally interesting study, at the moment data availability that would provide the appropriate analysis to account for the considerable structural changes still pose a challenge. For these reasons, we limit our investigation to the time before the effects of these events would have been reflected.

Table 2 shows the descriptive statistics of all the variables in the analysis. At first glance, we see that the representative (average) province in Spain grew at an average of 2.7% per year. Nonetheless, remarkable differences across the provinces are also observed, going from a high of 7% to a decline of 1.4% at the other extreme. The same wide disparity can be seen for GDP per worker in levels. Maps in Figures 1 and 2 also confirm these regional disparities in productivity growth and in levels.

Following our framework in arguing that part of these regional productivity gaps may be due to differences in human capital accumulation and distribution, we also observe interesting patterns in these variables. Within the period of analysis, the average Spanish province had about 10.7 years of schooling, but the gap between the highest and the lowest educational attainment is quite wide at about 3 years. This is important because it shows that there are provinces with populations who, on average, do not even complete secondary level of education, and this may have significant implications on why productivity varies in our regions. Looking at the distribution characteristics, we also see that the Gini index shows wide variability among the provinces. In addition to this, the degree of skewness reveals

rather extreme values going from very negative, meaning a larger proportion of the population has completed years of schooling higher than the average, to very positive, that a larger proportion has educational attainment below the average (better illustrated in Appendix 2 to compare positive and negative skewness between two provinces). We see these differences also confirmed in the maps in Figures 3 and 4. The chosen control variables also show substantial differences across provinces, suggesting that these may likewise contribute in explaining the observed regional disparities in productivity.

In the next step, we check how the variability in our measures of productivity relates with the three human capital variables. Figure 5 provides some initial support to our arguments, though they are by no means providing causal implications. GDP per worker is positively correlated with the average years of schooling. Also, provinces with higher levels of education inequality are associated with lower levels of GDP per worker. Finally, provinces with distributions that have a positive skew are correlated with lower levels of productivity. On the other hand, for productivity growth, the relationships seem to follow the same patterns, however they are less clear.

We also observe the collinearity between our human capital variables. Figure 6 illustrates the direction of these relationships. We keep in mind these high levels of correlation when interpreting results of the regressions, especially when we combine them in our specifications.

V. Empirical results

We report the results for the estimation of equation (2) in Table 3. Given that the paper focuses on the impact of the human capital variables (average years of schooling, the Gini index, and the degree of skewness), we present only the results for these variables, and for the measure of physical capital to capture the basic elements of the production function. Each column shows the estimates for the different combinations of the human capital variables in the static model. Panel (b) shows the estimates when including the set of control variables,[2] and the last row indicates the p-values of the F-test to determine their joint significance.

Before commenting on the results regarding the coefficients of the human capital measures, it is worth mentioning that the estimates of the return to physical capital are quite reasonable. As expected, the values are close to the share of capital in total income (between 30-40%). This indicates that the estimates based on the fixed-effects estimator of the model in levels may well be providing accurate estimates of the parameters of the production function.

We note that taken separately (columns (1) to (3) and (7) to (9)), each of the human capital variables are highly significant and exhibit the expected signs, even with the addition of controls. However, when taken together in column (12), the addition of the control variables causes a sharp decrease in the significance of the coefficients of the measures of education. Nonetheless, the effect of skewness is still significant at 5%, confirming the robustness of its effect on the level of productivity.

[2] Detailed estimation results may be made available upon request.

Specifically, this result implies that a one standard deviation increase in skewness in one province is associated with a 2.4% decrease in productivity. To stress the importance of this result, it may suggest that provinces with high levels of (positive) skewness would benefit more from programs that support securing tertiary level education, and this may contribute to catching up with productivity levels of other provinces, though likely with some lag before the effect can be seen.

Table 4 presents the results for estimating equation (3), for productivity growth, and shows similar patterns when the human capital variables are used in isolation. In the specification that includes the set of controls (column 12), the coefficient of skewness is significant and negative. This suggests that after the control variables have accounted for the variability in growth, what remains is strongly associated with the initial degree of skewness of the distribution of education, and marginally with the initial level of inequality. Thus, all else remaining the same, a one standard deviation increase in the degree of skewness in the previous period is related with a reduction in productivity growth by about 0.48%. Meanwhile, a one standard deviation increase in the level of inequality in the initial period in one province lowers productivity growth by 0.33%. This may imply that high levels of education inequality combined with positive degrees of skewness of its distribution may work against efforts to boost productivity growth.

As a robustness check, we replace the variable for the growth in average years of schooling with the initial average schooling level[3] and test its effect on subsequent growth (details of the results are in Appendix 3). Column (6) shows that only the initial level of Gini is significant and the coefficient estimate is negative. However, when controls are included (column (12)), none of the human capital variables turn up significant. It should be noted that some of the control variables highly correlate with the measures of human capital. This may explain these results, as well as the change in the level of the significance of the coefficients of the human capital variables in our various estimates.

In estimating equation (3) with the inclusion of the initial level of the GDP per worker, we implement the system-GMM estimator to this dynamic fixed-effects panel data model (Arellano and Bover, 1995; Blundell and Bond, 1998) to address the inconsistency resulting from the introduction of the lagged term. The results are shown in Table 5. One issue with the system-GMM is that it generates a large number of instruments that may result in biased estimates due to the presence of weak instruments. Roodman (2006) suggests a correction by controlling the number of instruments through setting lag limits for instruments and/or using a procedure that consists of collapsing the available instruments. As there are no specific guidelines on what defines too many instruments, we estimate with the default use of lagged instruments and the collapse option. We note also that the dynamic panel estimators do not account for correlation between individuals. This becomes an issue if there are shocks that affect all provinces simultaneously. For this reason, we include period dummies in these estimates (Roodman, 2006).

[3] Similar to Krueger and Lindahl (2001)

The estimates shown in Table 5 supports the effect of skewness on productivity growth even when controlling for the initial level of GDP per worker. Given the correlation between this variable and the measures of education (as revealed by the previous results for productivity levels), it would not have been surprising for the lagged GDP per worker to absorb the effects of the human capital variables. However, this seems not to be the case and skewness still has an effect. This implies that, ceteris paribus, a province whose degree of skewness is one standard deviation higher than another in the previous period is associated with a decrease in productivity growth by about 0.39% in the subsequent period. The growth in average years of schooling is also significant, with a negative coefficient estimate. This is not an unexpected result in the literature (such as Ramos, Surinach and Artis, 2010; De la Fuente, 2011) as it may indicate that over-education may be playing a role. It may also reflect the intermediate effect of policies that have been put in place to stimulate investment in higher levels of education in less dynamic regions.

Though skewness seems to play a significant role both in the levels and in the growth of productivity, it is observed that the magnitude of the coefficient estimate is much lower when related with productivity growth. We also see that the estimate for initial GDP per worker is highly significant, and negative, consistent with the predictions of convergence theories. On the other hand, Gini index does not show any significant effects on productivity growth, whereas average years of schooling shows a significant effect only when including the additional controls (column (12)).

The test statistics for serial correlation (Arellano-Bond test) confirms first-order serial correlation, and fails to reject at the higher order. In addition, Sargan's and Hansen's tests also indicate that the instrument set is appropriate. However, the interpretation of the results in Table 5 must be taken with a degree of caution. These dynamic models are designed for a large number of individual observations and small time periods. With only 50 provinces and T=4, our data dimensions would be considered moderate. We note for instance that the coefficient of the return to physical capital is much lower than what is seen as an acceptable level for a developed economy.

In summary, our estimates provide strong evidence supporting our hypothesis that the distribution of human capital contributes to explaining variability in regional productivity. Though we observe a high degree of correlation between the human capital variables that could have muted their impact when taken together, we find that skewness of the distribution in particular shows a quite robust significant effect. Increasing positive levels of skewness tends to be associated with lower productivity. This result is robust for both the level of productivity and its growth, and is also robust with the inclusion of the set of control variables.

VI. Conclusion

Our study confirms that the variations observed in the distribution of human capital variables are significant in explaining differences in productivity across Spanish provinces. In particular, the degree of skewness, whether the concentration of the population in a region has attained education levels higher or lower than the mean, has an impact on productivity. A likely channel may be through its

effect in enabling regions to adapt to evolving technologies and to innovate if more of their populations have higher levels of education.

These results may indicate that formulating policy on education to improve competitiveness and help lagging regions catch up merits also looking into the distribution characteristics. One-size-fits-all formulas for development policies are rarely recommended. This study is another illustration why more care must be taken. Distributional characteristics can provide clues on particular needs of regions. For example, a province with a high degree of positive skewness may benefit from policies promoting high secondary and tertiary levels of education, including addressing related credit constraints.

Though this paper is not without caveats, bear in mind that this study constitutes a preliminary look into deeper aspects of the distribution of education. For instance, the results may be criticized for its likely omission of other relevant variables that affect productivity. But as set out in the beginning, this exercise has been implemented with a specific focus on human capital in a basic aggregate production function—which has met considerable success—and not to capture all the factors that explain productivity.

Nevertheless, we do not end here. Our interest goes beyond the borders of Spain with the intention of applying the models for the EU regions. Data from Eurostat provides educational attainment at the NUTS 2 level, though only for 3 broad education categories. We are interested in testing our hypothesis using these categories by first regrouping the 8 categories we have used here to match Eurostat, and re-estimating our specifications using the new groups on the Spanish provinces. Following these results, we proceed with estimating for the EU. Another avenue to be explored is to look into spatial relationships for the human capital variables. But these are just some of the future steps that may be taken and are by no means exhaustive. The bottom line is that our results find support for the importance of the distribution of education on development outcomes, and that this is a promising area for further research.

References

Aghion, Philippe, Peter Howitt, Maxine Brant-Collett, and Cecilia García-Peñalosa. *Endogenous growth theory*. MIT press, 1998.

Arellano, Manuel, and Olympia Bover. "Another look at the instrumental variable estimation of error-components models." *Journal of econometrics* 68, no. 1 (1995): 29-51.

Aschauer, David Alan. "Is public expenditure productive?" *Journal of Monetary Economics* 23, no. 2 (1989): 177-200. doi:10.1016/0304-3932(89)90047-0.

Bancaja Foundation and Ivie (The Valencian Institute of Economic Research). Capital Humano en España y su distribución provincial. January 2014. Database available online: http://www.ivie.es/en/banco/caphum/series.php

Barro, Robert J., and Jong-Wha Lee. "International comparisons of educational attainment." *Journal of monetary economics* 32, no. 3 (1993): 363-394.

Barro, Robert J., and Jong Wha Lee. "International measures of schooling years and schooling quality." *The American Economic Review* 86, no. 2 (1996): 218-223.

Barro, Robert J., and Jong-Wha Lee. "International data on educational attainment: updates and implications." *oxford Economic papers* 53.3 (2001): 541-563.

Barro, Robert J. "Human capital and growth." *The American Economic Review* 91, no. 2 (2001)., 12-17.

Barro, Robert J., and Jong-Wha Lee. "A new data set of educational attainment in the World, 1950-2010." *Journal of Development Economics* 104 (2013): 184-198.

BBVA Foundation and Ivie (Instituto Valenciano de Investigaciones Económicas). El *stock* y los servicios del capital en España y su distribución territorial y sectorial (1964-2014). Enero de 2017. Base de datos disponible en Internet: http://www.fbbva.es/TLFU/microsites/stock09/fbbva_stock08_index.html

Becker, Gary S. "Investment in human capital: A theoretical analysis." *Journal of political economy* 70, no. 5, Part 2 (1962): 9-49.

Becker, Gary S., and Barry R. Chiswick. "Education and the distribution of earnings." *The American Economic Review* 56, no. 1/2 (1966): 358-369.

Becker, Gary Stanley. *Human capital: a theoretical and empiricarsl analysis, with special reference to education*. Chicago, Ill.: The Univ. of Chicago Press, 2008.

Benhabib, Jess, and Mark M. Spiegel. "The role of human capital in economic development evidence from aggregate cross-country data." *Journal of Monetary Economics* 34, no. 2 (1994): 143-73. doi:10.1016/0304-3932(94)90047-7.

Belorgey, Nicolas, Rémy Lecat, and Tristan-Pierre Maury. "Determinants of productivity per employee: An empirical estimation using panel data." *Economics Letters* 91, no. 2 (2006): 153-157.

Birdsall, Nancy, and Luis Londoño. "Asset inequality matters: an assessment of the World Bank's approach to poverty reduction." *American Economic Review* 87, no. 2 (1997): 32-37.

Black, Sandra E., and Lisa M. Lynch. "Human-capital investments and productivity." *The American Economic Review* 86, no. 2 (1996): 263-267.

Blundell, Richard, and Stephen Bond. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of econometrics* 87, no. 1 (1998): 115-143.

Bulmer, Michael George. *Principles of statistics*. Courier Corporation, 1979.

Caixa Galicia Foundation and Ivie (The Valencian Institute of Economic Research). Consumo de los hogares y distribución de la renta en España (1973 - 2005): Una perspectiva regional. 2009. Database available online: http://www.ivie.es/

Castello, Amparo, and Rafael Domenech. "Human capital inequality and economic growth: some new evidence." *The Economic Journal* 112, no. 478 (2002). doi:10.1111/1468-0297.00024.

Cohen, Daniel, and Marcelo Soto. "Growth and human capital: good data, good results." *Journal of Economic Growth* 12, no. 1 (2007): 51-76. doi:10.1007/s10887-007-9011-5.

Crescenzi, Riccardo, and Andrés Rodríguez-Pose. "Infrastructure and regional growth in the European Union." *Papers in Regional Science* 91, no. 3 (2012): 487-513. doi:10.1111/j.1435-5957.2012.00439.x.

Dalgaard, Carl-Johan., Henrik Hansen, and Theo Larsen. *Income skewness, redistribution and growth: a reconciliation* (No. 2003-14). EPRU Working Paper Series.

De Gregorio, Jose, and Jong-Wha Lee. "Education and income inequality: new evidence from cross-country data." *Review of Income and Wealth* 48, no. 3 (2002): 395-416. doi:10.1111/1475-4991.00060.

De la Escosura, Leandro Prados, and Joan R. Rosés. "Human capital and economic growth in Spain, 1850–2000." *Explorations in Economic History* 47, no. 4 (2010): 520-532.

De la Fuente, Angel, and Rafael Doménech. "Human capital in growth regressions: how much difference does data quality make?" *Journal of the European Economic Association* 4, no. 1 (2006): 1-36. doi:10.1162/jeea.2006.4.1.1.

De la Fuente, Angel. "Human capital and productivity." *Nordic Economic Policy Review* 2, (2011): 103-132.

Fleisher, Belton, Haizheng Li, and Min Qiang Zhao. "Human capital, economic growth, and regional inequality in China." *Journal of Development Economics* 92, no. 2 (2010): 215-31. doi:10.1016/j.jdeveco.2009.01.010.

Gille, Véronique. "Distribution of human capital and income: An empirical study on Indian States." *Journal of Macroeconomics* 43 (2015): 239-56. doi:10.1016/j.jmacro.2014.11.003.

Griffith, Rachel, Stephen Redding, and John Van Reenen. "R&D and absorptive capacity: theory and empirical evidence." *The Scandinavian Journal of Economics* 105, no. 1 (2003): 99-118.

Jenkins, Stephen P. "Did the middle class shrink during the 1980s? UK evidence from kernel density estimates." *Economics letters* 49, no. 4 (1995): 407-413.

Kaldewei C. and Walz U. (2001), "The determinants of regional growth in Europe: an empirical investigation of regional growth models", Working Paper, Department of Economics, University of Tuebingen.

Krueger, Alan, and Mikael Lindahl. "Education for growth: why and for whom?" *Journal of Economic Literature* 39, no. 4 (2001):1101-1136.

Lopez, Ramon, Vinod Thomas, and Yan Wang. "Addressing the education puzzle: the distribution of education and economic reform." *Policy Research Working Papers*, 1999. doi:10.1596/1813-9450-2031.

Lucas, Robert E. "On the mechanics of economic development." *Journal of Monetary Economics* 22, no. 1 (1988): 3-42. doi:10.1016/0304-3932(88)90168-7.

Moretti, Enrico. "Workers education, spillovers, and productivity: evidence from plant-level production functions." *American Economic Review* 94, no. 3 (2004): 656-90. doi:10.1257/0002828041464623.

Mankiw, Gregory, David Romer, and David Weil. "A Contribution to the empirics of economic growth." *Quarterly Journal of Economics*, 107, no.2 (1992): 407-437.

Nelson, Richard R., and Edmund S. Phelps. "Investment in humans, technological diffusion, and economic growth." *The American economic review* 56, no. 1/2 (1966): 69-75.

OECD. *Equity in Education: Spain Country Note*. Report. OECD. 2006.

Pritchett, Lant. "Where has all the education gone?" *The World Bank Economic Review* 15, no. 3 (2001): 367-91. doi:10.1093/wber/15.3.367.

Ramos, Raul, Jordi Suriñach, and Manuel Artís. "Human capital spillovers, productivity and regional convergence in Spain." *Papers in Regional Science* 89, no. 2 (2010): 435-47. doi:10.1111/j.1435-5957.2010.00296.x.

Rodríguez-Pose, Andrés, and Vassilis Tselios. "Education and income inequality in the regions of the European Union." *Journal of Regional Science* 49, no. 3 (2009): 411-37. doi:10.1111/j.1467-9787.2008.00602.x.

Romer, Paul M. "Endogenous technological change." *Journal of political Economy* 98, no. 5, Part 2 (1990): S71-S102.

Roodman, David. "How to do xtabond2: An introduction to difference and system GMM in Stata." (2006).

Roodman, David. "A note on the theme of too many instruments." *Oxford Bulletin of Economics and statistics* 71, no. 1 (2009): 135-158.

Sauer, Petra, and Martin Zagler. "(In)equality in Education and Economic Development." *Review of Income and Wealth* 60 (2014). doi:10.1111/roiw.12142.

Stock, James, and Mark Watson. "Understanding changes in international business cycle dynamics." 2003. doi:10.3386/w9859.

Temple, Jonathan. "Generalizations that aren't? Evidence on education and growth." *European Economic Review* 45 (2001): 905-918.

Thomas, Vinod, Yan Wang, and Xibo Fan. "Measuring education inequality: Gini coefficients of education for 140 countries." *Journal of Education Planning and Administration* 17 (2003): 5-33.

Uzawa, Hirofumi. "Optimum technical change in an aggregative model of economic growth." *International economic review* 6, no. 1 (1965): 18-31.

Vandenberghe, V. "The productivity challenge. What to expect from better-quality labour and capital inputs?" *Applied Economics* 49, no. 40 (2017): 4013-025. doi:10.1080/00036846.2016.1273504.

Vandenbussche, Jérôme, Philippe Aghion, and Costas Meghir. "Growth, distance to frontier and composition of human capital." *Journal of economic growth* 11, no. 2 (2006): 97-127.

Vogel, Johanna. "The two faces of R&D and human capital: Evidence from Western European regions." *Papers in Regional Science* 94, no. 3 (2013): 525-51. doi:10.1111/pirs.12084.

World Bank. *Achieving education for all by 2015: Simulation Results for 47 Low-Income Countries*. Working paper no. 76. Human Development Network, Africa Region and Education Department, World Bank. 2002.

World Bank. *Making transition work for everyone: poverty and inequality in Europe and Central Asia*. World Bank. 2000.

Table 1 Description of variables

Explanatory variables	*Definition*	*Source*
Growth Y/L	Annual growth of regional GDP per worker averaged every 4 years, constant prices	*Own elaboration using Eurostat and INE-SRA*
Log Y/L	Log of regional GDP per employed person, constant prices	*Eurostat and INE-SRA*
Average years of schooling	Computed for the economically active population, measured at intervals of four years	*Bancaja Foundation and the Valencian Institute of Economic Research (IVIE)*
Gini of education	Computed measure of dispersion at intervals of 4 years	*Own elaboration using IVIE data*
Skewness of education	Computed measure of skewness at intervals of 4 years	*Own elaboration using IVIE data*
Log K/L	Stock of productive capital per worker in logarithmic form, constant prices	*BBVA Foundation and IVIE*
Control variables:		
Log of public capital stock per surface area	Stock of public investment in transportation infrastructure (including road, airports, ports) per kilometer square of provincial areas in logarithms, constant prices	*BBVA Foundation and IVIE*
Log of population density	Inhabitants per kilometer square in logarithms	*Eurostat*
Share of employment in Agriculture	Proportion of total employed persons engaged in agriculture	*Bancaja Foundation and IVIE*
Share of employment in Industry	Proportion of total employed persons engaged in industry	*Bancaja Foundation and IVIE*
Employment rate	Proportion of total employed to the economically active population	*Bancaja Foundation and IVIE*

Table 2 Summary statistics

VARIABLES	N	mean	sd	min	max
Growth Y/L	200	0.0268	0.0136	-0.0144	0.0710
Log Y/L	200	10.75	0.151	10.34	11.15
Average years of schooling	200	10.71	0.800	8.929	12.79
Growth ave years of schooling	200	0.0545	0.0331	-0.0290	0.145
Gini of education	200	0.180	0.0226	0.111	0.228
Skewness of education	200	0.00320	0.243	-0.643	0.636
Log K/L	200	11.01	0.462	10.51	13.92
Growth K/L	200	0.0157	0.0230	-0.0505	0.0814
Log of public capital per surface area	200	13.07	0.877	11.45	15.41
Log population density	200	4.156	1.087	2.175	6.670
Share of workers employed in agriculture	200	0.0976	0.0670	0.00699	0.473
Share of workers employed in the industry sector	200	0.170	0.0704	0.0429	0.369
Employment rate	200	0.860	0.0675	0.608	0.957

Table 3 Estimation of the effect of the measures of education, equation in levels

	Panel (a) Log levels of GDP per worker (without control variables)					
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES						
ave schooling	0.167***			0.208***	0.122***	0.146***
	(0.0112)			(0.0238)	(0.0128)	(0.0364)
gini		-4.456***		1.356*		0.640
		(0.530)		(0.801)		(0.965)
skewness			-0.491***		-0.169***	-0.151***
			(0.0537)		(0.0365)	(0.0495)
log K/L	0.320***	0.455***	0.442***	0.314***	0.319***	0.317***
	(0.0935)	(0.128)	(0.110)	(0.0940)	(0.0869)	(0.0873)
Constant	5.430***	6.536***	5.882***	4.815***	5.924***	5.583***
	(0.965)	(1.475)	(1.215)	(1.117)	(0.904)	(1.154)
Observations	200	200	200	200	200	200
R-squared	0.796	0.683	0.736	0.802	0.811	0.812

	Panel (b) Log levels of GDP per worker (with control variables)					
	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES						
ave schooling	0.0516***			0.0450	0.0336*	0.00108
	(0.0189)			(0.0288)	(0.0200)	(0.0356)
gini		-0.925**		-0.159		-0.683
		(0.414)		(0.612)		(0.763)
skewness			-0.118**		-0.0791**	-0.0975**
			(0.0451)		(0.0390)	(0.0425)
log K/L	0.302***	0.311***	0.318***	0.302***	0.304***	0.306***
	(0.0746)	(0.0750)	(0.0716)	(0.0747)	(0.0704)	(0.0687)
Constant	2.522**	2.804**	2.725**	2.594**	2.831**	3.214**
	(1.142)	(1.234)	(1.134)	(1.223)	(1.112)	(1.273)
Observations	200	200	200	200	200	200
R-squared	0.887	0.885	0.887	0.887	0.890	0.891
F-test control variables (p-value)	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 Estimation of the effect of the measures of education, equation in differences

Panel (a) Annualized growth of GDP per worker, 4-year intervals (without control variables)						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES						
growth ave schooling	-0.115***			-0.0325	-0.0605*	-0.0285
	(0.0253)			(0.0375)	(0.0306)	(0.0396)
gini (t-4)		-0.248***		-0.207***		-0.135
		(0.0519)		(0.0714)		(0.0877)
skewness (t-4)			-0.0202***		-0.0148**	-0.00884
			(0.00428)		(0.00555)	(0.00746)
growth K/L	0.269***	0.294***	0.309***	0.295***	0.310***	0.310***
	(0.0548)	(0.0542)	(0.0614)	(0.0542)	(0.0586)	(0.0564)
Constant	0.0288***	0.0701***	0.0252***	0.0638***	0.0276***	0.0509***
	(0.00148)	(0.00976)	(0.00104)	(0.0122)	(0.00156)	(0.0151)
Observations	200	200	200	200	200	200
R-squared	0.226	0.269	0.252	0.272	0.270	0.282

Panel (b) Annualized growth of GDP per worker, 4-year intervals (with control variables)						
	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES						
growth ave schooling	-0.0682*			-0.0192	-0.0468	-0.0152
	(0.0387)			(0.0432)	(0.0343)	(0.0429)
gini (t-4)		-0.231**		-0.206**		-0.145*
		(0.0937)		(0.0962)		(0.0824)
skewness (t-4)			-0.0253***		-0.0229***	-0.0197***
			(0.00638)		(0.00646)	(0.00680)
growth K/L	0.235***	0.260***	0.257***	0.258***	0.256***	0.269***
	(0.0567)	(0.0571)	(0.0621)	(0.0584)	(0.0613)	(0.0619)
Constant	-0.258**	-0.0799	-0.161	-0.0874	-0.130	-0.0271
	(0.128)	(0.161)	(0.130)	(0.163)	(0.133)	(0.155)
Observations	200	200	200	200	200	200
R-squared	0.320	0.343	0.352	0.344	0.361	0.372
F-test control variables (p-value)	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 System-GMM estimation of the dynamic model

Panel (a) Annualized growth of GDP per worker, 4-year intervals (without control variables)						
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
growth ave schooling	-0.0607*			-0.0368	-0.0557**	-0.0397
	(0.0303)			(0.0288)	(0.0236)	(0.0260)
gini (t-4)		-0.190*		-0.151		-0.0892
		(0.0984)		(0.0989)		(0.0569)
skewness (t-4)			-0.0256***		-0.0235**	-0.0216**
			(0.00900)		(0.00969)	(0.00922)
growth K/L	0.184***	0.201***	0.178***	0.196***	0.177***	0.185***
	(0.0475)	(0.0427)	(0.0450)	(0.0429)	(0.0457)	(0.0440)
log Y/L (t-4)	-0.114***	-0.111***	-0.121***	-0.112***	-0.121***	-0.119***
	(0.0234)	(0.0183)	(0.0206)	(0.0201)	(0.0231)	(0.0216)
Constant	1.216***	1.226***	1.300***	1.232***	1.306***	1.301***
	(0.246)	(0.201)	(0.220)	(0.221)	(0.247)	(0.230)
Observations	200	200	200	200	200	200
Instruments	10	10	10	11	11	12
AR2	0.244	0.378	0.888	0.255	0.471	0.466
Sargan	0.080	0.200	0.620	0.149	0.433	0.483
Hansen	0.177	0.299	0.601	0.234	0.350	0.400

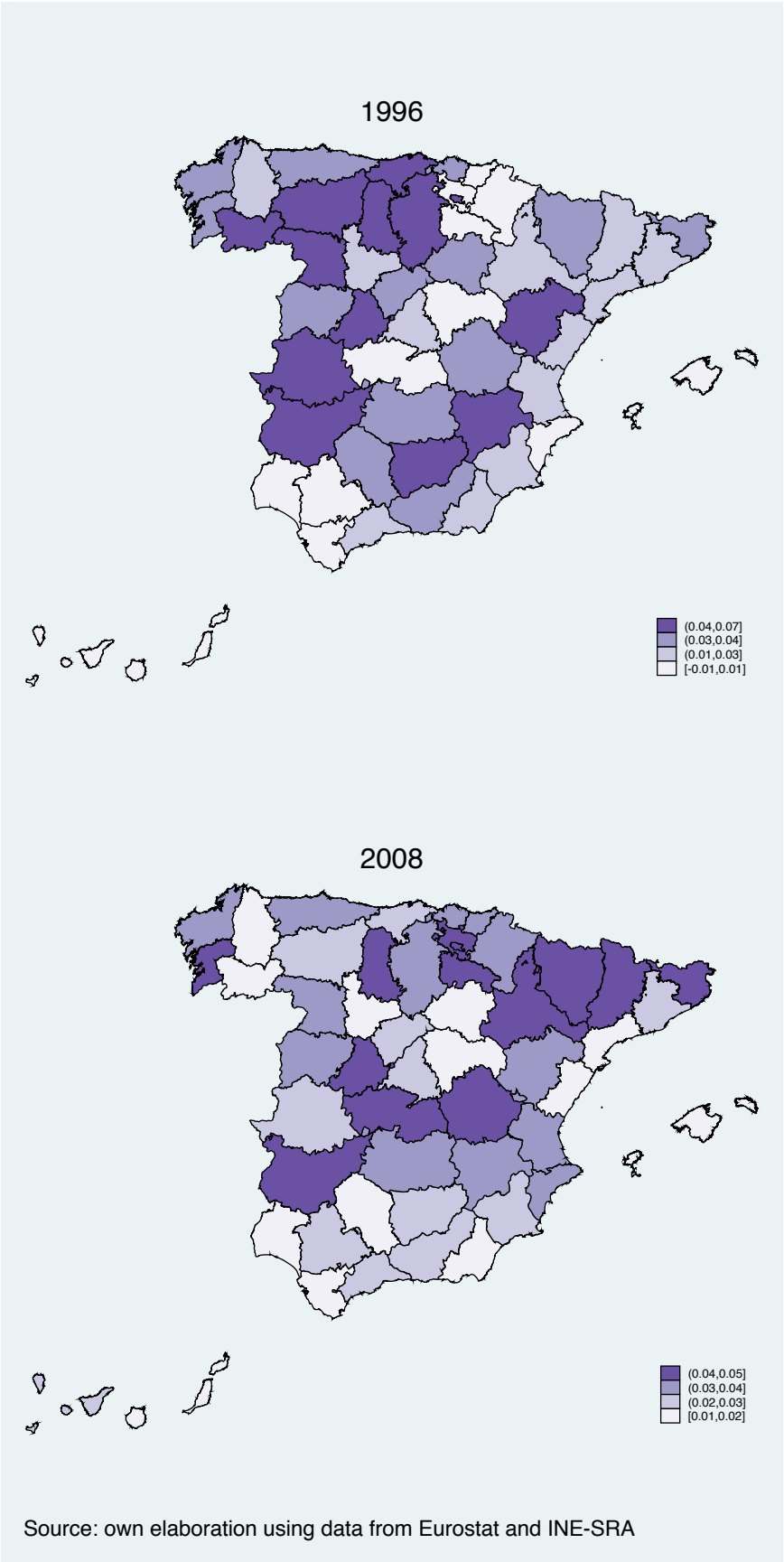
Panel (b) Annualized growth of GDP per worker, 4-year intervals (with control variables)						
VARIABLES	(7)	(8)	(9)	(10)	(11)	(12)
growth ave schooling	-0.0422*			-0.0493**	-0.0413*	-0.0473*
	(0.0248)			(0.0238)	(0.0232)	(0.0245)
gini (t-4)		-0.0280		0.0380		0.0336
		(0.0790)		(0.0915)		(0.0802)
skewness (t-4)			-0.0166**		-0.0155*	-0.0161*
			(0.00788)		(0.00842)	(0.00867)
growth K/L	0.172***	0.177***	0.185***	0.166**	0.184***	0.178***
	(0.0554)	(0.0572)	(0.0494)	(0.0634)	(0.0509)	(0.0577)
log Y/L (t-4)	-0.168***	-0.165***	-0.145***	-0.171***	-0.147***	-0.149***
	(0.0505)	(0.0479)	(0.0363)	(0.0552)	(0.0406)	(0.0443)
constant	1.537***	1.514***	1.341***	1.559***	1.366***	1.373***
	(0.464)	(0.432)	(0.341)	(0.496)	(0.382)	(0.404)
Observations	200	200	200	200	200	200
Instruments	15	15	15	16	16	17
AR2	0.562	0.843	0.862	0.602	0.531	0.555
Sargan	0.544	0.697	0.681	0.254	0.515	0.484
Hansen	0.400	0.592	0.597	0.175	0.384	0.375
F-test control variables (p-value)	0.0947	0.0998	0.0270	0.1166	0.0546	0.0763

All estimates include time fixed-fixed effects

Standard errors in parentheses

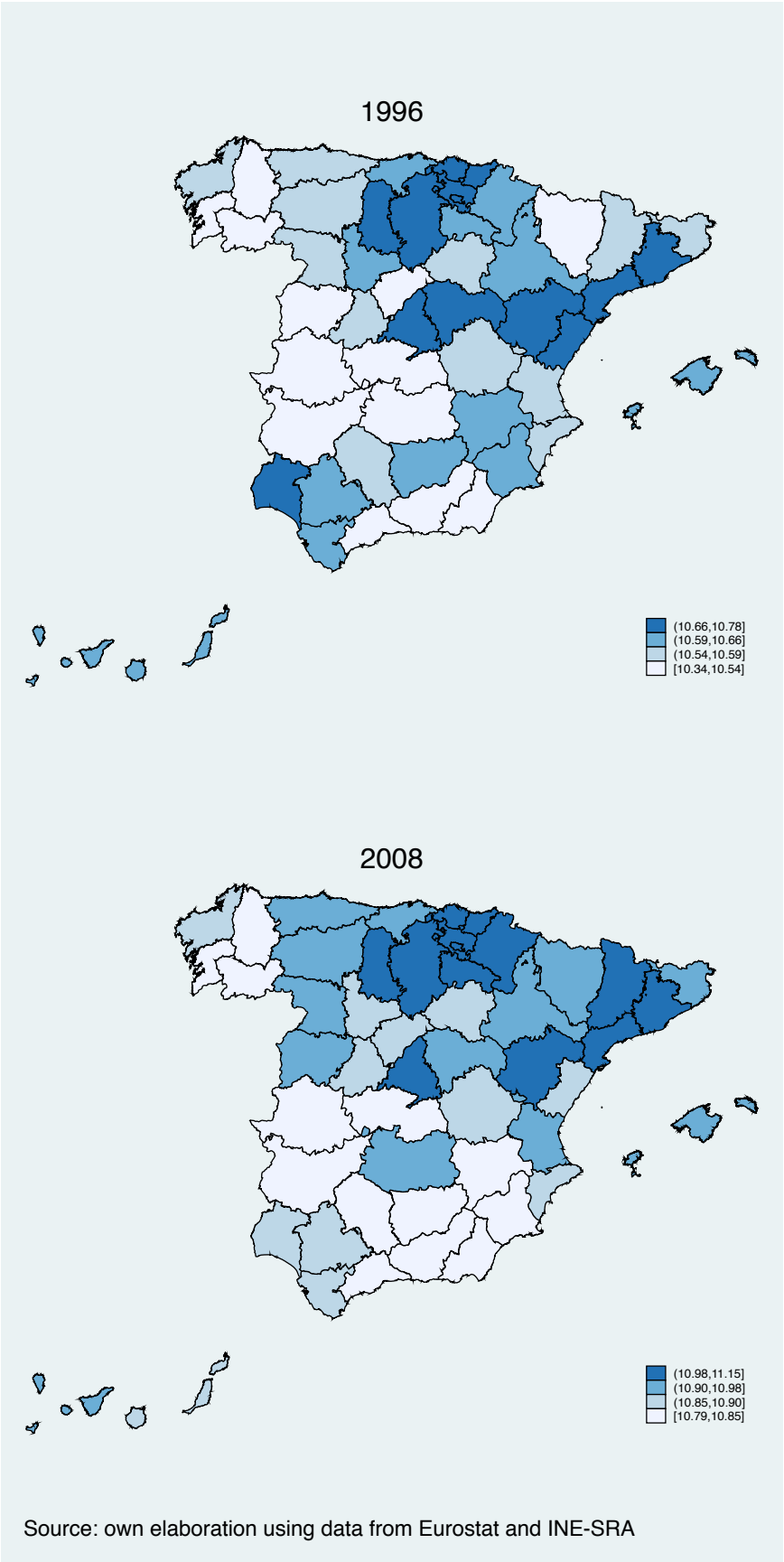
*** p<0.01, ** p<0.05, * p<0.1

Figure 1 Productivity growth in Spanish provinces



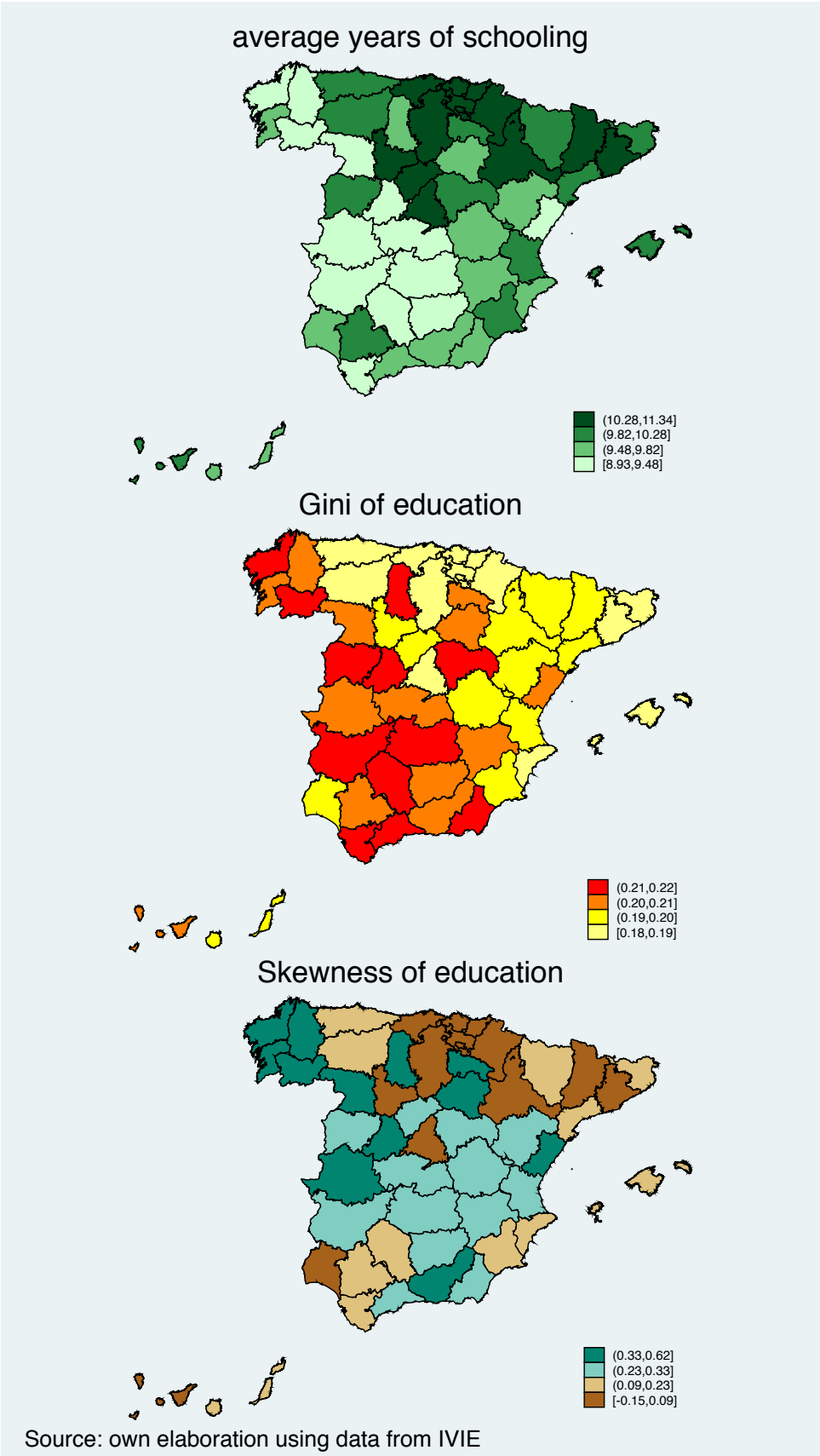
Source: own elaboration using data from Eurostat and INE-SRA

Figure 2 Productivity levels in Spanish provinces (GDP per worker)



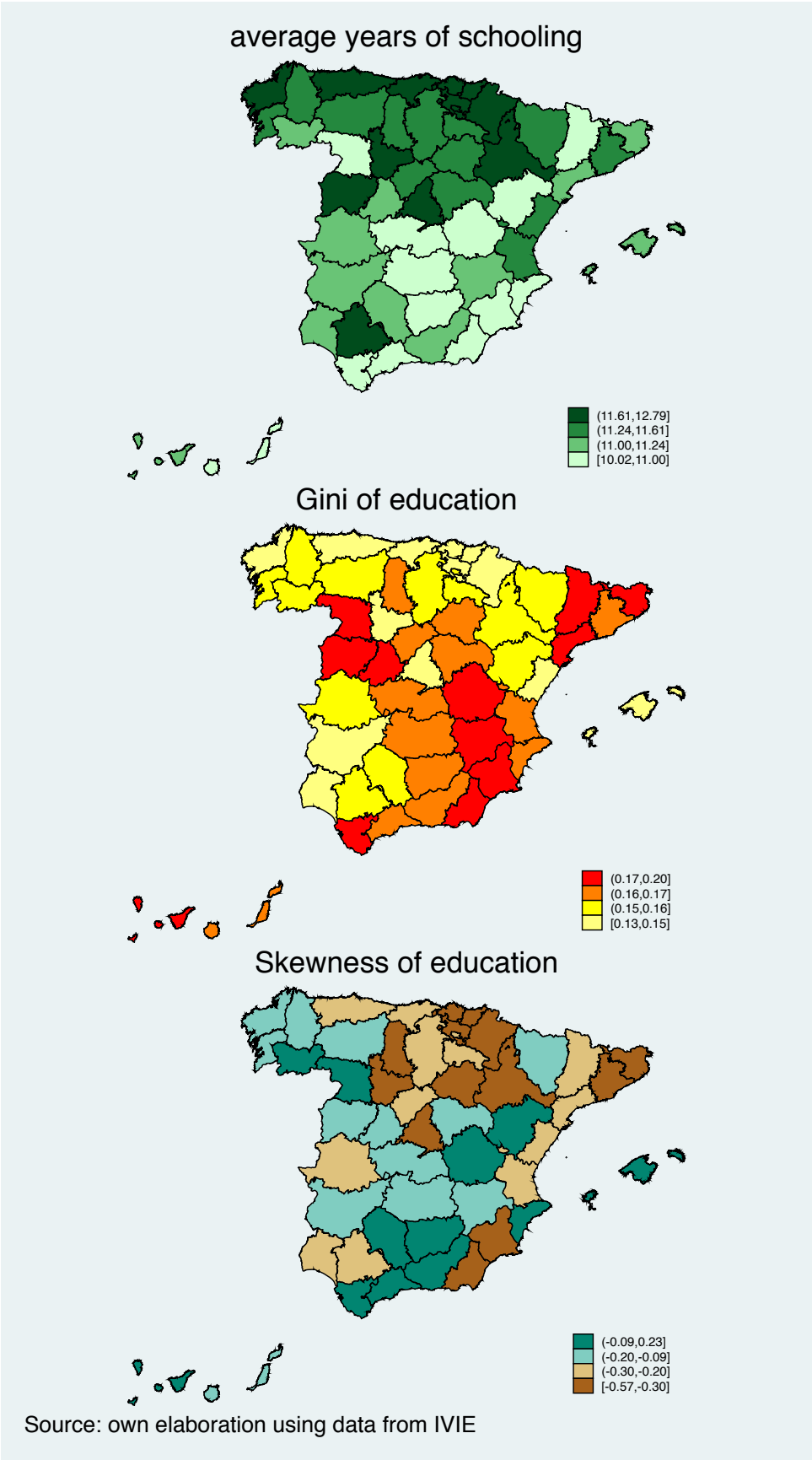
Source: own elaboration using data from Eurostat and INE-SRA

Figure 3 Measures of human capital in the Spanish provinces (1996)



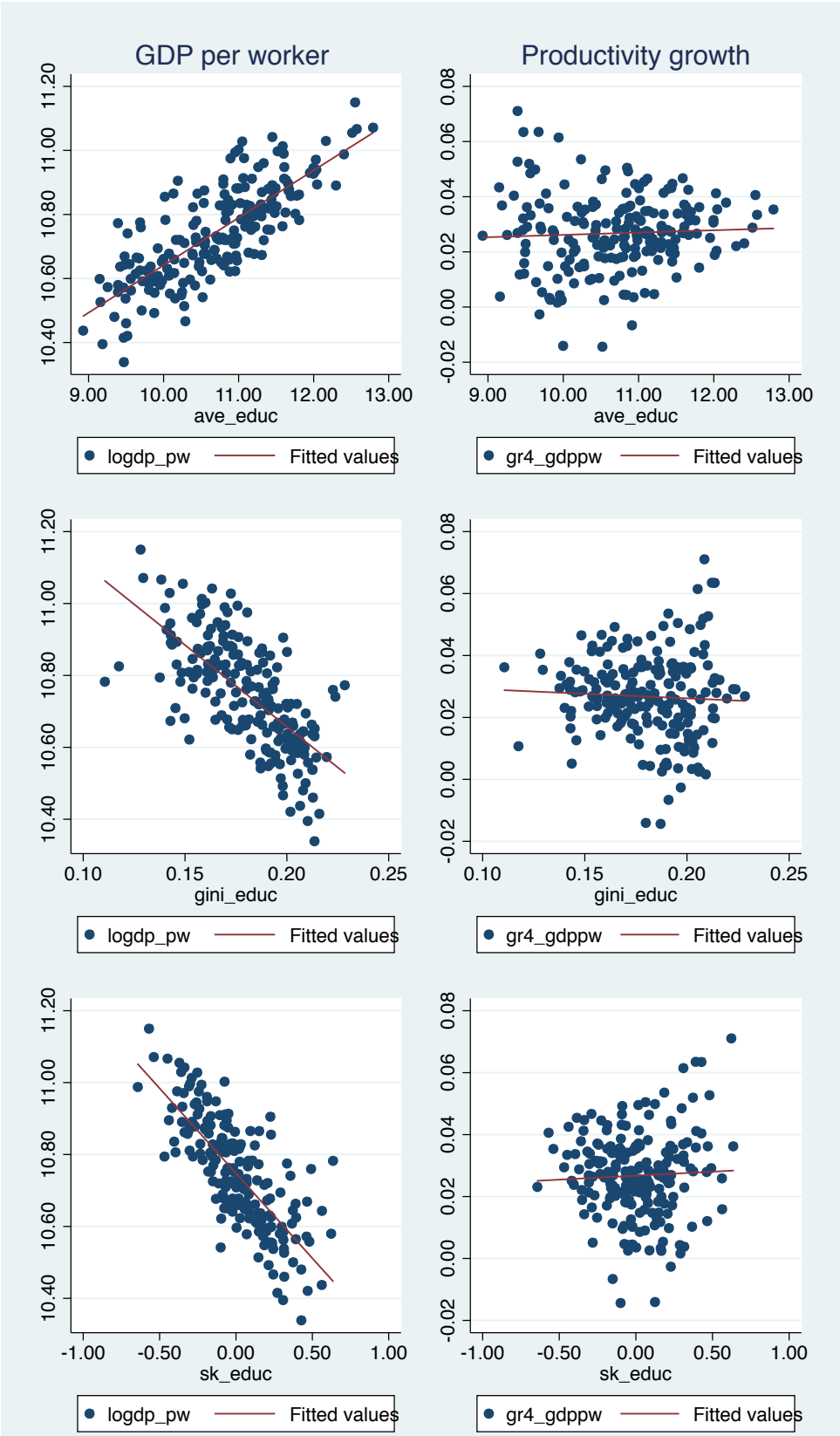
Source: own elaboration using data from IVIE

Figure 4 Measures of human capital in the Spanish provinces (2008)



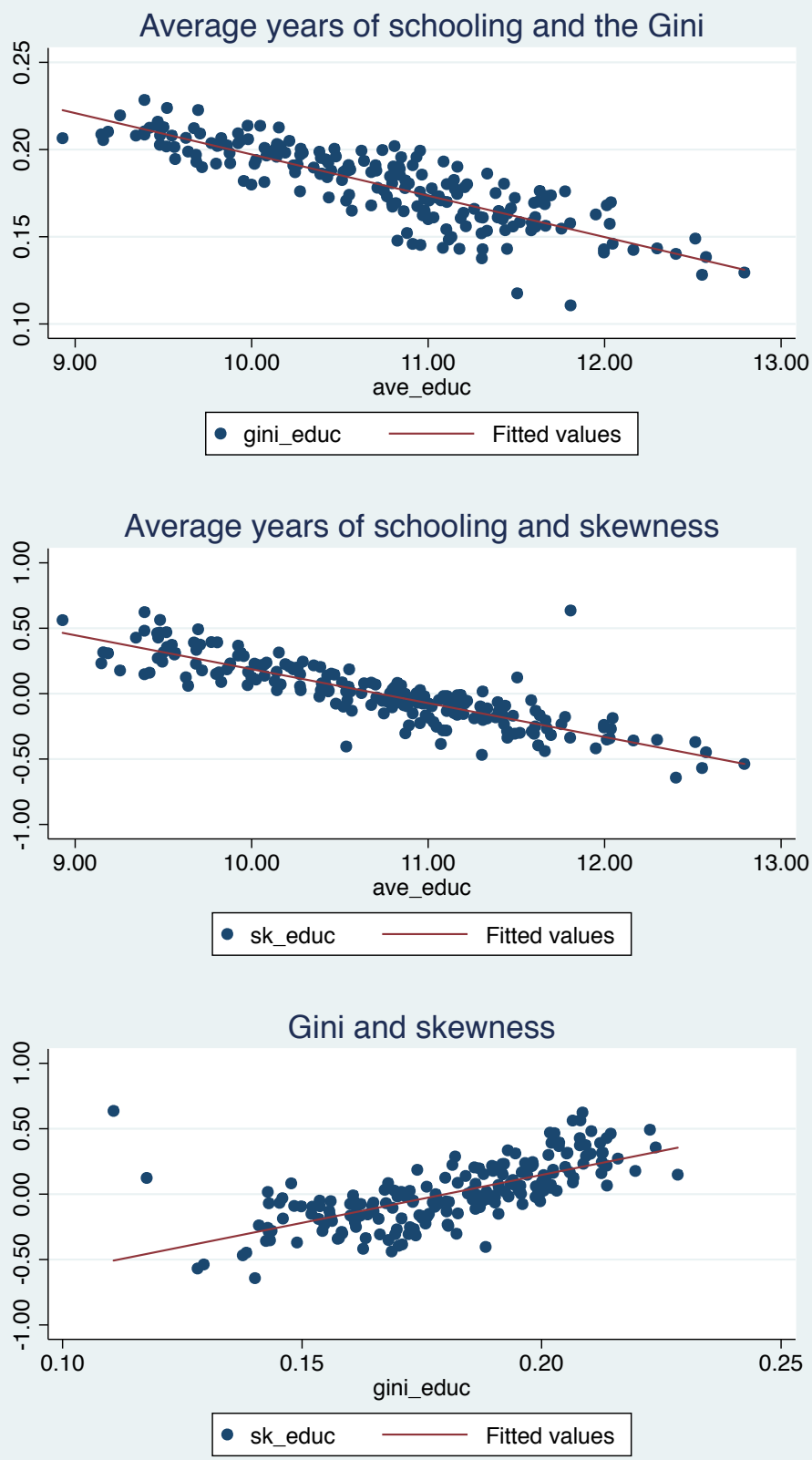
Source: own elaboration using data from IVIE

Figure 5 Productivity and human capital variables



Source: own elaboration using data from Eurostat, INE-SRA, and IVIE

Figure 6 Correlation between the human capital variables



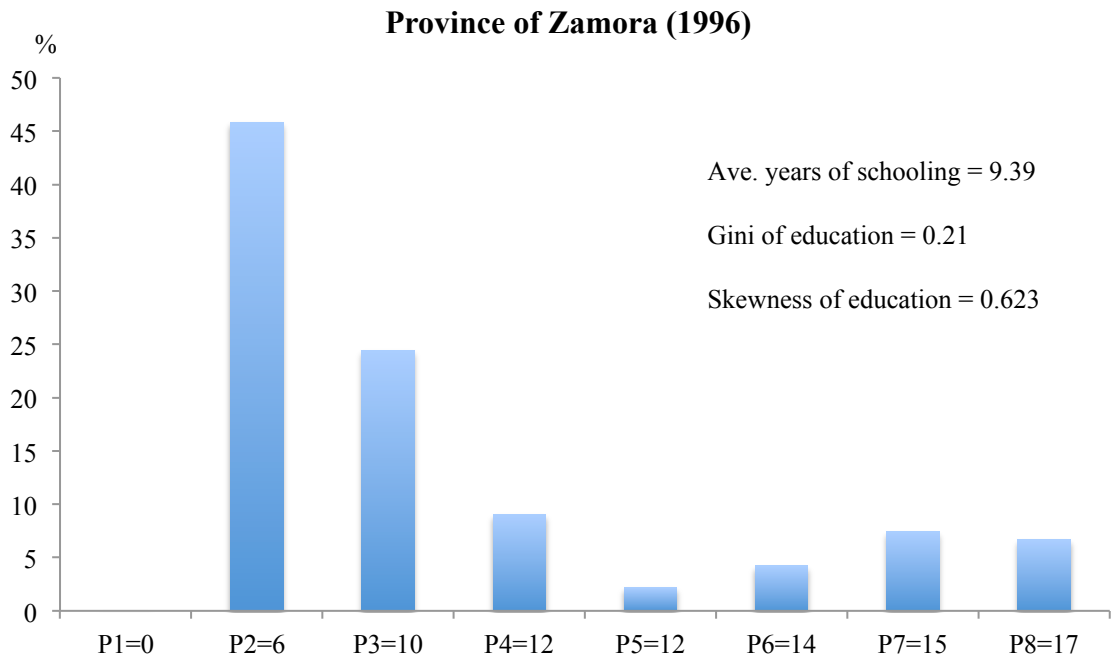
Source: own elaboration using data from IVIE

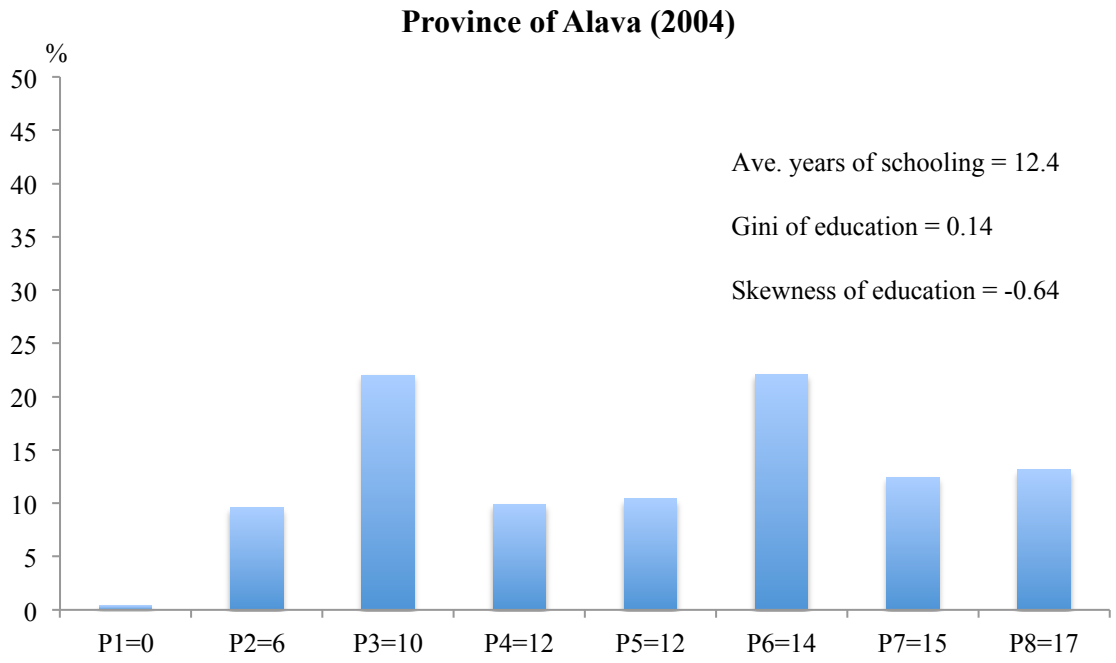
Appendix 1 Education categories in Spain

	No. of years *
Analfabetos	0
Sin estudios y primarios	6
Bachiller Elemental	10
Bachiller Superior	12
FP_I[a]	12
FP_II[b]	14
Anteriores al superior	15
Superiores	17

[a] Formación Profesional I/Ciclos Formativos de grado medio
[b] Formación Profesional II/Ciclos Formativos de grado superior
Source: IVIE

Appendix 2 Proportion of the population at each education level





Source: Own elaboration using data from IVIE
Note: P indicates the equivalent years of schooling completed at each level

Appendix 3 Estimation of the effect of the measures of education, equation in differences

Panel (a) Annualized growth of GDP per worker, 4-year intervals (without control variables)

VARIABLES	(1)	(4)	(5)	(6)
ave schooling (t-4)	0.00597***	0.00200	0.00377*	-0.00282
	(0.00110)	(0.00287)	(0.00209)	(0.00371)
gini (t-4)		-0.177		-0.223**
		(0.126)		(0.104)
skewness (t-4)			-0.00837	-0.0144
			(0.00884)	(0.0100)
growth K/L	0.316***	0.305***	0.317***	0.304***
	(0.0567)	(0.0524)	(0.0583)	(0.0539)
constant	-0.0388***	0.0358	-0.0152	0.0961*
	(0.0114)	(0.0525)	(0.0223)	(0.0547)
Observations	200	200	200	200
R-squared	0.258	0.272	0.261	0.282

Panel (b) Annualized growth of GDP per worker, 4-year intervals (with control variables)

VARIABLES	(7)	(10)	(11)	(12)
ave schooling (t-4)	0.0112***	0.00967**	0.00769***	0.00430
	(0.00249)	(0.00445)	(0.00272)	(0.00581)
gini (t-4)		-0.0494		-0.0988
		(0.141)		(0.132)
skewness (t-4)			-0.0248***	-0.0154
			(0.00878)	(0.00935)
growth K/L	0.287***	0.285***	0.331***	0.280***
	(0.0573)	(0.0566)	(0.0643)	(0.0578)
constant	-0.139	-0.112	0.114	-0.0481
	(0.131)	(0.167)	(0.165)	(0.163)
Observations	200	200	150	200
R-squared	0.362	0.363	0.368	0.374
F-test control variables (p-value)	0.000	0.000	0.022	0.000

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1