

22
23

Mind the Gap: Education Inequality at the Regional Level in Portugal, 1986-2005*

João Gabriel Fidalgo / Marta Simões / Adelaide Duarte FEUC / FEUC, GEMF / FEUC, GEMF

resumo

résumé / abstract

Portugal é um dos países mais desiguais em termos de rendimento entre os países desenvolvidos. De 1980 a 2005, a desigualdade de rendimento manteve valores elevados, resultado sobretudo do aumento da desigualdade dos rendimentos do trabalho. Dada a estreita ligação entre educação e rendimentos do trabalho, é fundamental analisar a distribuição de educação. O objectivo principal deste trabalho é avaliar empiricamente a desigualdade de educação em Portugal a um nível regional entre 1986 e 2005. Os resultados apontam para um aumento do nível médio de educação da força de trabalho em simultâneo com uma diminuição da desigualdade de educação, ao nível nacional. Contudo, várias regiões, as inicialmente mais pobres em educação, apresentam um aumento da desigualdade. Comprovamos também empiricamente a existência de uma curva de Kuznets de educação.

Portugal figure parmi les pays développés comme un des pays des plus inégaux en termes de revenu. Durant 1980-2005, l'inégalité du revenu maintient des valeurs élevées alimentée surtout pas une croissance de l'inégalité des gains. Du fait de l'existence d'une relation étroite entre l'éducation et les gains, l'étude de la distribution de l'éducation est importante. D'abord nous mesurons l'inégalité de l'éducation au niveau régional, au Portugal, entre 1986 et 2005. Nos résultats confirment une augmentation du niveau moyen de l'éducation de la force de travail accompagnée par une diminution de l'inégalité pour l'ensemble du pays. Nonobstant, certaines régions exhibent une augmentation de l'inégalité. Finalement, nous obtenons des résultats empiriques qui supportent l'hypothèse de la courbe de Kuznets pour l'éducation.

Portugal stands as one of the most unequal countries in terms of income among the developed countries. Over the period 1980-2005, income inequality kept high, fostered mainly by a monotonic increase in earnings inequality. Given the close link between education and earnings, it is of major importance to study the distribution of education. This paper examines the distribution of education at the regional level in Portugal between 1986 and 2005. Our results indicate that education inequality decreased for the whole country as the average education level of the workforce rose, over the sample years. This finding does not apply at the regional level however, with several districts initially poor in terms of education exhibiting an increase in education inequality. The evidence also supports the existence of a Kuznets curve of education: as the average level of education rises, education inequality first increases, and, after reaching a peak at 5.13 years of schooling, starts declining.

* The authors would like to thank the editor of the journal for valuable comments and suggestions.

1. Introdução



Income inequality has been a permanent theme on economic research, whereas human capital inequality is a very recent topic of study. The importance of human capital and its distribution only emerged after the surge of endogenous growth theories during the 1980s. One can ask about the relevance of this topic *per se* for Economics, which then leads to the question, *Is inequality really a bad thing?* According to some theories, inequality can be detrimental to economic growth and reduce individual welfare. For countries such as Portugal, that presents persistently high income inequality levels, the study of inequality is thus of great importance.

The stock of human capital is a fundamental determinant of a country's growth performance as we are rapidly moving towards a knowledge-based economy. Endogenous growth theories brought to the front line the importance of this production factor not only as a major inductor of innovation but also as an essential tool for absorption and adaptation to new knowledge and technologies. Furthermore, at the individual level, the stock of human capital is a major determinant of individual employability and earnings. In turn, from an intergenerational perspective, the educational level of the parents decisively influences the educational level achieved by their offspring. The level and the distribution of human capital are thus important determinants of the growth performance and the level of social cohesion of a country and should be the focus of research.

The intergenerational transmission of inequality is a matter of much concern, since it lowers the equality of opportunity among the young and produces economic inefficiencies given the waste of skills from children with a low family background. The vicious cycle of inequality is a complex network of mechanisms. First, given the liquidity constraints, income inequality leads to human capital inequality. In other words, poor households will lack the financial resources to invest in their children's education. Additionally, parents' educational level strongly determines the educational attainment of children, which means that human capital inequality will tend to persist over generations. These mechanisms act therefore as intergenerational mechanisms of inequality reproduction. At the same time, there will be an intra-generational effect, since the distribution of human capital will determine the contemporaneous distribution of income, via wages and individual employability. And if inequality prevents poorer individuals and their children from investing in human capital, then inequality is detrimental to growth.

The main goal of our paper is to gauge education inequality at the regional level. Our analysis covers the eighteen Portuguese *distritos* and two insular regions, from 1986 to 2005. We also investigate if the evolution of education inequality at regional level fits the hypothesis of a Kuznets curve of education.

In the next section we provide a brief overview of the theoretical and empirical studies that focus on the interactions between inequality and economic growth. We also review the empirical studies that have explored the hypothesis of a Kuznets curve of education and summarize the evolution of income and earnings inequality in Portugal over the sample period. In section 3, we develop a pioneering analysis of education inequality in Portugal at the regional level. Section 4 concludes.

2. Literature review

a. Inequality and economic growth

No definite sign can *a priori* be anticipated for the relationship between inequality and economic growth, given the numerous links between the two variables. The main channels through which inequality impacts economic growth can be nevertheless summarized into four groups: (i) borrowing constraints and the investment in physical and human capital; (ii) fiscal policy; (iii) socio-political instability and macroeconomic volatility; and (iv) saving rates behaviour (Perotti, 1996; Aghion *et al.*, 1999; Barro, 2000).



The first channel considers that in the presence of credit constraints the access to certain investment opportunities (physical and human capital) will depend on the initial individual distribution of income and wealth and, as a result, a reduction in inequality would lead to an increase in investment and growth (see e.g. Galor & Zeira, 1993; Perotti, 1996). However, if setup costs of physical investments are sizeable with respect to the median income, a reduction in inequality will be detrimental. It is, however, expected that this channel will produce larger effects in developing countries, where credit constraints are more pervasive (Barro, 2000).

According to the second channel, the level of taxation in the economy is determined by the median voter, i.e. the agent with the median level of income. In more unequal economies, where the median voter is thus relatively poor, this agent will require deeper redistributive policies through voting (the political mechanism). Hence, given that these expenditures are in part financed by distortionary taxes, the more unequal a society is, the larger the distortionary effect over the economy, resulting in less investment and growth (Bertola, 1993; Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Aghion *et al.*, 1999).

Additionally, unequal societies tend to be more unstable, with frequent riots and other antisocial behaviours, which waste resources and time, leading to lower productivity and growth (see e.g. Alesina and Perotti, 1996; Perotti, 1996). There is also empirical support for claiming that higher inequality is linked to greater macroeconomic volatility, mainly resulting from socio-political instability (Aghion *et al.*, 1999) and higher sensitivity to adverse shocks. This third transmission mechanism is known as the socio-political instability channel.

The fourth channel predicts a positive impact of inequality on growth, as higher inequality implies higher savings rates and, consequently, higher investment and growth. This channel is based on the idea that the savings rate is a positive function of the level of income, which is the same as saying that rich people exhibit a higher savings rate (Barro, 2000).

Although the majority of the channels of influence described above point to a negative relationship between inequality and growth, especially for poor countries, given their higher socio-political instability, macroeconomic volatility, and larger credit constraints, empirical studies have provided mixed results. For instance, Perotti (1996) arrived at a negative relationship between inequality and growth, while Barro (2000) found a positive link between inequality and growth for rich countries although for the joint sample of rich and poor countries the sign was never statistically significant. But this conflicting evidence seems to result more from differences in methods than anything else, namely from: (i) different econometric specifications; (ii) the choice of income to assess inequality instead of wealth, the variable used in theoretical models; and (iii) different inequality measures¹.

b. The Kuznets Curve of education

The Kuznets curve (Kuznets, 1955) establishes an inverted-U relationship between economic development and inequality by relating output per capita to a measure of income inequality². Given the close link between the level of human capital/education and the level of income, some researchers tried to find a similar relationship between the level of human capital/education and inequality in its distribution³.

1 Dominics, Groot and Florax (2006) apply meta-analysis to 22 empirical studies on inequality and growth and find evidence that the variation in the inequality coefficient estimates are systematically associated with differences in estimation methods, sample coverage and data quality.

2 The transition from an agricultural-rural economy, exhibiting low inequality, to an industrial-urban economy is considered to be responsible for the initial increase in inequality. The massive industrialization and the equalization of returns across sectors will eventually lead to a decrease in inequality.

3 In this paper inequality is computed relative to the distribution of years of formal education of workers, similar to what is done in Birdsall and Londoño (1997) and Castelló and Doménech (2002), that build inequality measures based on data on average years of schooling from Barro (1993; 1996; 2001) and Nehru, Swanson and Dubey (1995).

Initial studies by Ram (1990) and Londoño (1990) were favourable to the idea of a Kuznets curve of education. Ram (1990), using cross-section data for 94 countries, obtained an inverted-U relationship between average years of schooling and the standard deviation of education, reaching a peak at 6.8 years of schooling. Londoño (1990) applying a similar methodology found the same inverted-U shaped curve. More recently, De Gregorio and Lee (2002), using panel data covering an interesting number of countries over a period of three decades (1960-1990), also found a Kuznets curve of education, using the standard deviation of years of schooling, again an absolute measure of inequality.

More recent studies (Thomas, Wang and Fan, 2001; Checchi, 2001; Castelló and Doménech, 2002; Lim and Tang (2008)) argue that the Kuznets curve of education is a direct consequence of using absolute measures of inequality, such as the standard deviation. Using relative measures of inequality, such as the Gini coefficient, leads to a negative relationship between education inequality and average years of schooling instead, which is a consequence of the asymptotical behaviour of the two types of measures (see Thomas, Wang and Fan, 2001). Some studies have nevertheless been able to find an inverted-U relationship between average years of schooling and the education Gini. Lin (2007), in an intra-country study for Taiwan between 1976 and 2003, found indeed support in favour of the hypothesis of a Kuznets curve of education, using the Gini index.

Another recent related debate concerns the specification of human capital to use when the objective is to measure inequality in its distribution and analyse the respective relationship with the average level. For instance, Lim and Tang (2008) consider that conclusions based on education/years of schooling inequality might be misleading and adopt a Mincerian specification of human capital as the conceptually most appropriate. They argue that the Kuznets curve of (Mincerian) human capital, instead of the Kuznets curve of education, is the 'natural explanation' for the Kuznets curve of income. Using the Mincerian specification for human capital (instead of years of schooling) and the corresponding inequality measure, the authors confirm the hypothesis of a Kuznets curve of human capital using the Gini index, i.e. a relative measure of inequality, for a sample of 99 countries over the period 1960-2000.

c. The evolution of income and earnings inequality in Portugal since the 1980s

The distribution of education is an important explanatory factor of the distribution of income or earnings as pointed out by several studies.

Gouveia and Tavares (1995) found a decrease in income inequality in Portugal during the 1980s, based on estimations from the Household Budget Survey (IOF). The authors suggested several reasons for this behaviour: the reduction in inequality of the distribution of years of schooling; the Portuguese comparative advantage in labour-intensive industries, with trade liberalization during this period resulting in higher demand for low-skilled workers; and the increasing equalising effect of redistributive policies. In contrast, Cardoso (1998), investigating the evolution of earnings inequality between 1980 and 1992 found an increase in earnings inequality. According to this author, the Portuguese earnings distribution has a compressed bottom and a stretched top, meaning that the problem of earnings inequality was that the upper wages were very high given the overall distribution. These, apparently, contradictory findings can perhaps be reconciled if one takes into account that during this period there was an increase in earnings inequality (and capital income), but that this evolution was more than offset by the equalising effect of direct taxes (Rodrigues, 1994).

The 1990s show a more mixed picture. In the first half of the 1990s, in particular, there was an inversion in the downward trend observed in the 1980s, with the high-income households benefitting from the highest income growth rates (Rodrigues, 2007). In the second half, in turn, there seems to be no clear trend. Thus, while Sologon and O'Donoghue (2009) report an unambiguous increase in earnings inequality based on earnings variance estimation, Budria (2007), using data from the European Community Household Panel, showed a decrease in earnings inequality between 1994 and 2001, with a statistically significant decrease in the Gini coefficient for earnings.

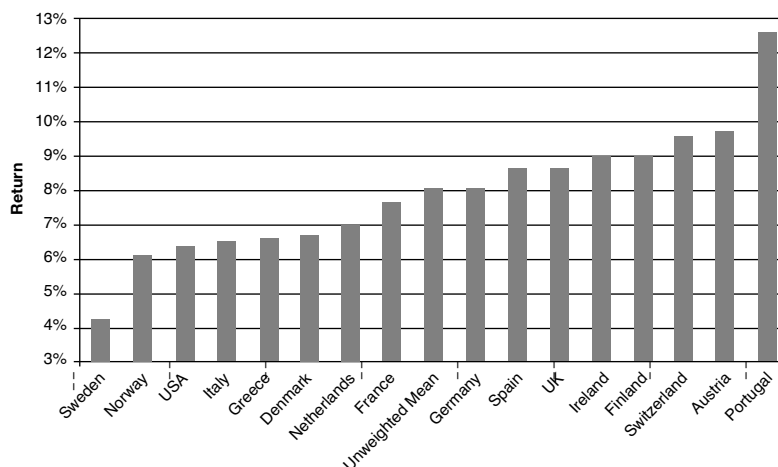




For the 2000-2005 period, the most recent period for which income inequality studies are available, Alves (2009), drawing from the 2005/06 Household Expenditure Survey (IDEF)⁴, identifies a slight increase in inequality for monetary and total income.

Returns to education seem to have played a major role in the increase of earnings inequality since the 1980s. Firstly, the high returns to education in Portugal (see Figure 1) and its increase in recent years have fostered between-group earnings inequality, since the gap between skilled and unskilled workers has increased. Secondly, the increase in within-group inequality can also be attributed to returns to education, given that these returns increase along the conditional distribution, i.e. the distribution controlling for observable characteristics (Martins and Pereira, 2004) with the most skilled workers experiencing the highest earnings increase. As an illustration, Carneiro (2008) found that education explains 40 to 50% of wages inequality in Portugal in 2004⁵.

Figure 1 – Returns to schooling in selected developed countries



Source: Martins and Pereira (2004). Based on OLS estimations.

TABLE 1 – Evolution of inequality for total income and monetary income, 1980-2005

	Total Income			Monetary Income		
	Gini Index	Atkinson ($\epsilon=0.5$)	Atkinson ($\epsilon=2$)	Gini Index	Atkinson ($\epsilon=0.5$)	Atkinson ($\epsilon=2$)
1980	0.3193	0.0830	0.2934	-	-	-
1989	0.3169	0.0818	0.2886	0.3285	0.0880	0.3393
1995	0.3473	0.0979	0.3205	0.3576	0.1048	0.3520
2000	0.3481	0.0985	0.3140	0.3658	0.1088	0.3490
2005	0.3437	0.0976	0.3099	0.3729	0.1153	0.3617

Source: Rodrigues (1996), for 1980, and Rodrigues (2007), for 1989, 1995 and 2000. The values for 2005 were directly provided by Nuno Alves to whom we are very grateful. All the estimations were drawn from the Household Budget Survey.

4 This survey, also carried out by Statistics of Portugal (INE), was the successor of the IOF.

5 Martins and Pereira (2004) argue that the factors that explain the different returns to schooling within the same educational group (such as over-education, natural ability, school quality or the field of study) produce larger effects on the group of highly-qualified workers, and therefore we should expect greater earnings dispersion within this group than within other educational groups. These authors try to overcome the widespread idea that wage inequality would be minimized solely by promoting higher schooling levels.

In Table 1 we present some key inequality measures for the distributions of total and monetary income taken from selected studies (Rodrigues (1996, 2007)). As it can be seen, all measures show a downward trend over the 1980s, followed by an increase during the 1990s. In the early 2000s, the figures for the distribution of total income show a very slight decrease, while the indexes for the distribution of monetary income show a clear increase.



3. Education inequality at the regional level in Portugal

3.1. The data

Our dataset is drawn from *Quadros de Pessoal* (QP), an annual census survey conducted by the Portuguese Ministry of Labour that covers all firms with wage earners⁶. The dataset collects information on the characteristics of workers (e.g. age, gender and the highest completed level of education) and firms (e.g. sales, number of employees, industry and region). QP also records information on workers' monthly earnings, which include basic earnings, regular subsidies, irregular subsidies and overtime pay, and on hours of work. All the information refers to October of each year. Public administration workers and those serving in the armed forces are absent from QP. For this study, we selected three data points, 1986, 1996, and 2005, with 1,722,455, 1,955,792, and 2,656,124 workers, respectively.

3.2. Methodology and concepts

Due to data limitations, the majority of empirical studies on the relationship between inequality and growth use inequality measures for the income distribution, although the theoretical literature focus on the link between the distribution of wealth and growth. Since human capital (or schooling) as been shown to be an even more important determinant of the distribution of wealth, the measures analysed should focus on human capital or education inequality rather than on income inequality.

Our paper focus is therefore on education inequality. The variable that will be used to measure the level of education of the labour force is years of schooling of workers. The geographical focus is the eighteen Portuguese *Distritos*, plus Madeira and Açores. *Distritos* are a territorial and administrative division of the Portuguese mainland.

All inequality measures mirror different subjective views about the distribution under analysis (Fields, 2001). However, and despite important differences across the different measures, the following four properties are, in general, required: (i) anonymity, which assumes that two distributions are equally unequal if one is obtained from a permutation of the other; (ii) income homogeneity, which implies that inequality is independent from the scale in which income is measured; (iii) population homogeneity, meaning that if a distribution is a replication of another, then they are equally unequal; (iv) the Pigou-Dalton transfer principle, which imposes that the result of an income transfer from a relatively rich person to a relatively poor one (without changing their position in the distribution and the mean of the distribution) is a reduction in inequality.

We select the following measures of inequality: the Gini index (G), the Atkinson index (A) and the Theil's first measure (T). These inequality indices can be extrapolated in order to analyse education inequality simply by substituting years of schooling for income in the definitions of the selected indices. (The corresponding definitions are given in Appendix A.)

Each measure reacts differently to changes in the shape of the distribution: the Gini index is sensitive to changes in the middle of the distribution⁷; Theil's first measure reacts mainly to

⁶ *Quadros de Pessoal* database is provided by Gabinete de Estratégia e Planeamento, Ministério do Trabalho e da Solidariedade Social (GEP – MTSS).

⁷ In the case of education inequality, the Gini index also becomes very sensitive to the share of population uneducated (with zero years of schooling). In the presence of uneducated individuals, the Lorenz curve becomes truncated along the horizontal axis which influences a great deal the value of the education Gini (see



changes at the bottom end of the distribution; and, the sensitivity of the Atkinson index to different parts of the distribution depends on the parameter that reflects inequality aversion, ϵ^8 .

A common way to obtain a visual picture of inequality is to plot the Lorenz curve. In our study, this curve depicts the cumulative percentage of education acquired by each cumulative percentage of population (from the lowest to the highest in terms of years of schooling). The closer the Lorenz curve is to the 45 degree line (the perfectly equal distribution line), the more equal is the distribution under consideration. A subsequent concept is Lorenz-dominance, which allows one to say that if a Lorenz curve, say for distribution X, lies for some point above and never below another Lorenz Curve, associated to a distribution Y, then X Lorenz-dominates Y (or $L_X > L_Y$). The final result is that X is more equal than Y, according to the Lorenz criterion. Given Lorenz-dominance, an ordinal ranking can be obtained from a broad set of inequality measures, so that $L_X > L_Y \Rightarrow I(X) > I(Y)$, where $I(\cdot)$ is some measure of inequality. Examples of these measures are the Gini index, the Theil indices, the Atkinson index and the coefficient of variation. Another group of measures, said to be weakly Lorenz-consistent, verify the condition $L_X > L_Y \Rightarrow I(X) \geq I(Y)$. The various quantile ratios are a case in point of the latter group of measures. The two other possible situations are: (i) Lorenz curves coincidence, meaning that both distributions are equally unequal; and (ii) Lorenz curves crossing, which allows no comparison based on the Lorenz criterion (Fields, 2001). Both inequality measures, strongly and weakly Lorenz-consistent, satisfy $L_X > L_Y \Rightarrow I(X) = I(Y)$. Our empirical analysis uses the statistical software R and Lorenz dominance analysis (and the corresponding inequality measures referred above)⁹.

3.2. Education inequality

In our analysis, educational attainment is given by the number of years of schooling corresponding to the highest completed level of education of workers from QP. (See Appendix B for definition and methodology.) The distribution of education and the corresponding inequality indices for 1986, 1996 and 2005 are presented in Tables 2-5. First we look at each cross-section separately and then overtime. In each case, we first analyse the chosen summary measure of education for the entire Portuguese economy and by region, and then present the inequality indices. Finally, we test the hypothesis of a Kuznets curve of education.

(i) The 1986 cross-section

According to Table 2, in 1986, average schooling in the Portuguese workforce was equal to 5.46 years¹⁰, ranging from 4.68 years in Braga to 6.37 years in Lisboa. Inter-regional dispersion is small, since the coefficient of variation for average years of schooling is only 4.2%. If one includes the region of Lisboa, the coefficient of variation increases to 7.3%, which means that dispersion remains small anyway.

Table 3 shows the estimated Gini coefficient. For the entire country, the coefficient is equal to 0.2844. The inter-regional dispersion is now larger at 10.9%, with Braga having the least unequal index (at 0.2129) and Portalegre the most unequal (at 0.3141). Considering the other inequality indices, in Table 4 and in Figure 2, the ranking of the regions according to the different measures does change. Nonetheless, Braga is the least unequal region irrespective of the selected index, which means that Braga Lorenz-dominates all the other regions.

e.g. Morrisson and Murtin (2007)). Our study is not influenced by this problem since we assume that the lowest level of education corresponds to one year of schooling, a common assumption when measuring human capital (also known as nonzero specification).

⁸ We consider ϵ equal to 0.5 in order to make the Atkinson index more sensitive to changes in the top end of the distribution as lower values of ϵ put greater weight in the highest incomes.

⁹ The software R is available from <http://www.r-project.org/>.

¹⁰ Despite some methodological differences, our values for average years of schooling are very close to those reported by Teixeira (2005).

TABLE 2 – Average years of schooling

Position	1986		1996		2005		Growth rate 1986-2005
	Portugal	5.46	Portugal	6.587	Portugal	7.805	42.9%
	Littoral	5.526	Littoral	6.655	Littoral	7.887	42.7%
	Interior	4.86	Interior	6.07	Interior	7.308	50.4%
1	Lisboa	6.373	Lisboa	7.602	Lisboa	8.706	36.6%
2	Setúbal	5.327	Setúbal	6.892	Setúbal	8.126	52.5%
3	Açores	5.325	Faro	6.665	Faro	7.877	49.0%
4	Faro	5.287	Coimbra	6.436	Coimbra	7.708	45.9%
5	Coimbra	5.284	Madeira	6.423	Santarém	7.58	54.0%
6	Porto	5.224	Bragança	6.412	Viana do Castelo	7.58	50.8%
7	Bragança	5.118	Porto	6.302	Madeira	7.567	49.9%
8	Leiria	5.059	Santarém	6.294	Leiria	7.555	49.3%
9	Madeira	5.049	Açores	6.248	Porto	7.541	44.4%
10	Viana do Castelo	5.026	Leiria	6.224	Bragança	7.523	47.0%
11	Aveiro	4.96	Viana do Castelo	6.206	Évora	7.388	56.6%
12	Viseu	4.943	Vila Real	6.196	Vila Real	7.349	49.3%
13	Vila Real	4.923	Viseu	6.128	Castelo Branco	7.33	50.6%
14	Santarém	4.921	Beja	6.073	Aveiro	7.329	47.8%
15	Castelo Branco	4.868	Guarda	6.016	Viseu	7.269	47.1%
16	Guarda	4.839	Castelo Branco	6.006	Portalegre	7.246	52.2%
17	Beja	4.792	Évora	5.975	Beja	7.235	51.0%
18	Portalegre	4.761	Aveiro	5.94	Guarda	7.219	49.2%
19	Évora	4.718	Portalegre	5.926	Açores	7.188	35.0%
20	Braga	4.677	Braga	5.738	Braga	7.037	50.5%
	<i>Coefficient of variation</i>	7.25%		6.52%		5.03%	
	<i>Unweighted mean</i>	5.0737		6.2851		7.51765	

Notes: Hereafter, Littoral regions are in light-grey and the Interior ones are in dark-grey. The insular regions are in italics.

One interesting feature is that Braga is the least unequal region, whereas the region with the highest educational attainment, Lisboa, is the most unequal, except for the three districts, Portalegre, Beja and Évora, that altogether form one macro-region called *Alentejo*, one of seven NUTS2 regions¹¹. A look at the shape of both distributions, in Table 5, allows to derive some important conclusions: (i) Braga and Lisboa display a very similar bottom, with the first decile, the first quartile and the median presenting the same value (4 schooling years); (ii) Braga exhibits a very compressed top, the third quartile and the ninth decile present the same value (6 years of schooling), while Lisboa depicts a stretched and unequal top, corresponding to 9 years in the

11 Alentejo is one of the seven Portuguese NUTS2 regions; the others are North, Centre, Lisboa, Algarve, Madeira and Açores. As far as inequality is concerned, Alentejo displays a very particular behaviour compared to all the other regions. It started with one of the lowest education levels and the highest inequality value (very stretched top and bottom), instead of a compressed distribution around low values of years of schooling similar to the other education-poor districts. Furthermore, it exhibited a downward trend in inequality while increasing its education level (as the bottom became more compressed around medium schooling levels).

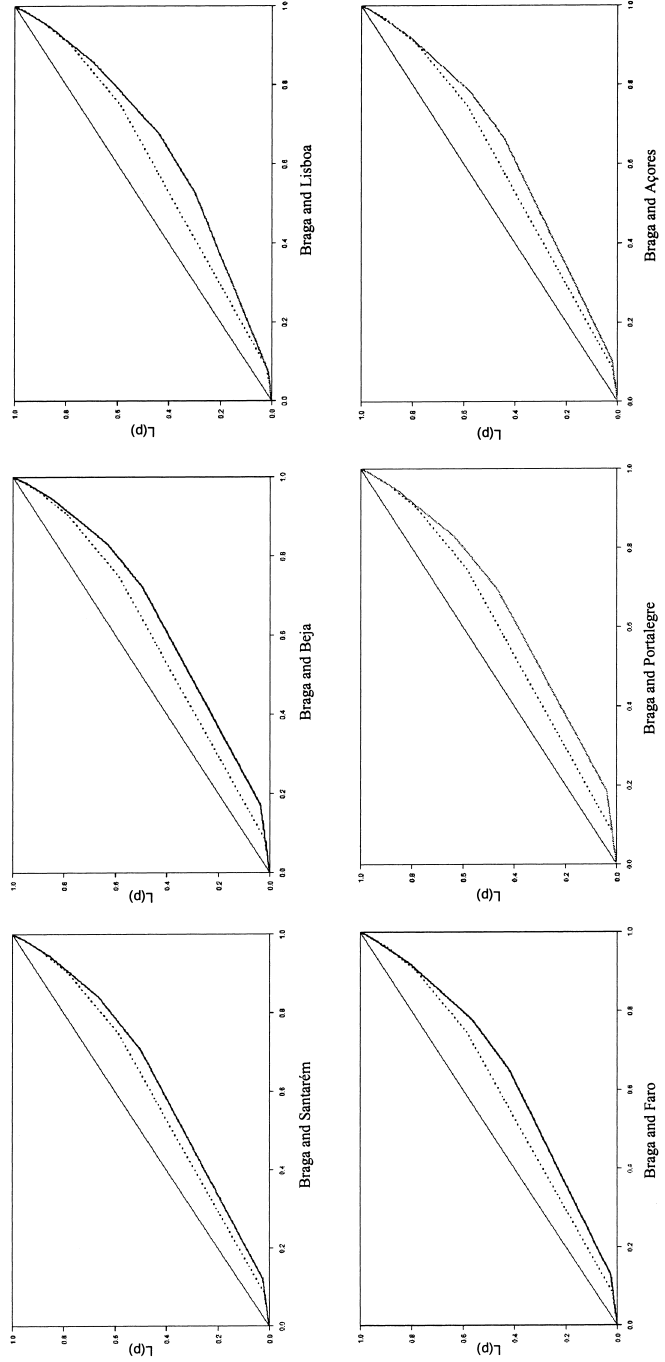


TABLE 3 – Gini coefficients

Position	1986		1996		2005		Growth rates		
	Portugal	0.2844	Portugal	0.2674	Portugal	0.2655	1986-2005	1986-1996	1996-2005
1	Portugal	0.2844	Portugal	0.2674	Portugal	0.2655	-10.3%	-6.0%	-4.6%
2	Litoral	0.285	Litoral	0.2678	Litoral	0.2534	-11.1%	-6.0%	-5.4%
3	Interior	0.2676	Interior	0.2591	Interior	0.2652	-0.9%	-3.2%	2.4%
4	Braga	0.2129	Braga	0.2263	Viana do Castelo	0.2348	-0.7%	-3.3%	2.6%
5	Viseu	0.2319	Viana do Castelo	0.2288	Braga	0.2358	10.8%	6.3%	4.2%
6	Viana do Castelo	0.2365	Viseu	0.2429	Setúbal	0.2428	-16.0%	-6.4%	-10.3%
7	Guarda	0.2404	Aveiro	0.2445	Lisboa	0.2431	-20.7%	-11.4%	-10.4%
8	Castelo Branco	0.2437	Castelo Branco	0.2508	Açores	0.2498	-12.1%	-9.4%	-3.0%
9	Aveiro	0.248	Guarda	0.2519	Coimbra	0.2504	-6.7%	-4.2%	-2.5%
10	Vila Real	0.2602	Coimbra	0.2569	Faro	0.2511	-16.4%	-11.4%	-5.7%
11	Leiria	0.2615	Açores	0.2576	Madeira	0.2517	-13.6%	-7.6%	-6.5%
12	Porto	0.2641	Vila Real	0.2588	Leiria	0.2527	-3.4%	-0.5%	-2.9%
13	Coimbra	0.2683	Porto	0.2598	Viseu	0.254	9.5%	4.7%	4.6%
14	Santarém	0.2737	Leiria	0.2602	Santarém	0.2552	-6.8%	-3.6%	-3.3%
15	Bragança	0.2816	Bragança	0.2626	Aveiro	0.2556	3.1%	-1.4%	4.5%
16	Açores	0.2843	Santarém	0.2638	Porto	0.2577	-2.4%	-1.6%	-0.8%
17	Setúbal	0.2892	Faro	0.2662	Castelo Branco	0.2643	8.5%	2.9%	5.4%
18	Madeira	0.2912	Madeira	0.2691	Évora	0.2657	-15.0%	-10.2%	-5.4%
19	Faro	0.3005	Setúbal	0.2708	Guarda	0.2687	11.8%	4.8%	6.7%
20	Lisboa	0.3064	Lisboa	0.2714	Portalegre	0.2688	-14.4%	-13.2%	-1.4%
21	Beja	0.3091	Portalegre	0.2726	Bragança	0.2697	-4.2%	-6.7%	2.7%
22	Évora	0.3127	Beja	0.2746	Beja	0.2708	-12.4%	-11.2%	-1.4%
23	Portalegre	0.3141	Évora	0.2809	Vila Real	0.2756	5.9%	-0.5%	6.5%
Coefficient of variation		10.9%	5.6%		4.6%				
Unweighted mean		0.2715	0.2585		0.2559				



Figure 2 – Comparison between Braga and some selected regions (Lorenz curves)



Notes: The dotted Lorenz curve in each graphic corresponds to Braga. The other regions (solid lines) are, by order of appearance in the graphics from left to right, Santarém, Beja, Lisboa, Faro, Portalegre and Açores.



(ii) The 1996 cross-section

Average years of schooling of the Portuguese workforce in 1996 was 6.59, which represents an important increase in the level of education from 1986. Braga and Lisboa maintained their positions as the regions with, respectively, the lowest and the highest level of schooling at 5.74 and 7.60 years. The other regions exhibited a very similar upward trend in schooling.

As shown in Tables 3 and 4, all selected inequality measures are lower in 1996 than in 1986 for most regions. Braga and Lisboa are at the top and the bottom positions, respectively, with a Gini coefficient of 0.2263 in the former region, and 0.2714 in the latter. Again, considering the three inequality indices, the ranking of the regions is not exactly the same, not even at the top positions, with Braga and Viana do Castelo exchanging the first position depending on the selected inequality index. One important feature is that, irrespective of the selected index, inter-regional dispersion in 1996 is smaller than in 1986. In particular, the inter-regional dispersion for the Gini coefficient reduced to half, from 10.9% to 5.6%.

Looking at the shape of the distributions (see Table 5), we conclude that Braga and Lisboa still present a very similar distribution up to the median. Differences are more visible in the second half of the distribution: Braga displays a compressed top with schooling levels close to those of the first half of the distribution, while the distribution of Lisboa has a clear stretched top. This conclusion is also supported by the fact that the ratio Q50/Q10 is the same for Lisboa and Braga while the ratio Q90/Q50 is higher in Lisboa.

(iii) The 2005 cross-section

The level of education of the Portuguese workforce kept its upward trend between 1996 and 2005. For the entire country, average years of schooling is 7.8, while in Lisboa is 8.71 years, and in Braga 7.04 years. Meanwhile, inter-regional dispersion continues to narrow, since the coefficient of variation was 5.03%, considering all regions (3.51% ignoring Lisboa).

Most regions also show a reduction in inequality (see Tables 3 and 4). Viana do Castelo was, in all cases, the least unequal region, whereas Vila Real is in the last position, showing a considerable increase in inequality over the 1986-2005 period. Despite different trends among regions, in 2005 inter-regional dispersion with respect to inequality is visibly smaller than in 1986 – the coefficient of variation for all the three inequality measures is smaller in 2005 than in 1986.

(iv) Education inequality over the period 1986-2005

The main conclusion from the previous analysis is that the level of education is steadily increasing while, at the same time, for the majority of the Portuguese regions, there is an unambiguous decrease in within-region education inequality.

One interesting aspect can also be identified by looking at the changes in the inequality ranking over the period 1986-2005. In the first year under analysis, 1986, the top positions of the inequality ranking (more equal regions) were occupied by both *litoral*¹² and *interior*¹³ regions indistinctly, and the same applies to the bottom positions. In 2005, on the contrary, only one *interior* region, Viseu, appears in the top positions, which implies that the *litoral* regions became the least unequal and also exhibit higher levels of education. The *interior* regions became the most unequal regions, with the lowest educational attainment¹⁴.

A second feature observed is that the regions that started with higher initial levels of education are the ones that exhibit a decrease in education inequality, whereas the regions that started with

12 The following ten *Distritos* located near the coastline are classified as *litoral*: Aveiro, Braga, Coimbra, Faro, Leiria, Lisboa, Porto, Santarém, Setúbal, Viana do Castelo.

13 The following eight, landlocked *Distritos* are classified as *interior*: Beja, Bragança, Castelo Branco, Évora, Guarda, Portalegre, Vila Real, Viseu.

14 Only Viseu is in the top-ten positions in the inequality rankings for the various years. Bragança is the tenth education-richest region in 2005.

a lower schooling level recorded increases in inequality. This interesting empirical result suggests that the relationship between the education level and education inequality can indeed be close to an inverted-U shape. We investigate this issue in the next section.



(v) The Kuznets curve of education

The process of increasing the level of education will tend to produce an inverted-U-shaped evolution of education inequality. Each region/country starts from a situation characterised by a low level of education, in which the large majority of the population has a very low level of education. Therefore, the distribution will be considerably compressed around the lowest levels of education¹⁵. As the level of education rises, particularly through the access of young cohorts to higher levels of education, the result is a stretching of the top end of the distribution, which will lead to an increase in education inequality. This intermediate situation reflects a dualistic structure, with two main groups: (i) a large group with low levels of education; and (ii) a small, but growing group with medium to high levels of education. In a third stage, as an increasing share of the population obtains higher levels of education, there will be a shifting of the distribution to the right. Education distribution is thus expected to return to its initial compressed shape, but now around high levels of education. In this final stage there will be a homogeneous and highly educated population.

In contrast to the original Kuznets curve, which is based on the idea of inter-sectoral shifts, our framework reflects mainly inter-generational changes. Since we are using years of schooling as our proxy for education, the individual level of education can be considered rather fixed. Therefore, changes in the level of education of the workforce in a certain economy emerge mainly via inter-generational forces.

(v.1) The empirical model

We will test two main hypothesis:

- **hypothesis 1:** the regions with a higher initial level of education (i.e. above a certain threshold) tend to have a decrease in education inequality, whereas the poor regions tend to have an increase in inequality;
- **hypothesis 2:** the relationship between the level of education and education inequality has an inverted-U shape.

Accordingly, the regressions to be estimated are the following:

$$\Delta X_i = \beta_0 + \beta_1 H_{i,1986} + u_i \quad (1)$$

$$X_{it} = \beta_0 + \beta_1 H_{it} + u_{it} \quad (2)$$

$$X_{it} = \beta_0 + \beta_1 H_{it} + \beta_2 + H_{it}^2 + u_{it} \quad (3)$$

where X_{it} and ΔX_{it} are, respectively, the level and the variation in education inequality between 1986 and 2005, H_{it} is the level of education given by the average years of schooling, α_i is the

¹⁵ The usual assumption, when measuring the stock of human capital, concerning the most basic level of education is that the individuals that have not completed primary education (a 4-year cycle) have one year of schooling, instead of zero years, since even when an individual has no formal education there is a minimum level of human capital, which is acquired through informal education and on-the-job training, for instance. We follow this specification, also known as the nonzero specification/hypothesis, which may influence our conclusions, since when the average education level is very low, education inequality will also be very low, as all the population will have a very similar education level given our lower bound (one year of schooling).



fixed effects component, i is the geographic unit (the eighteen *Distritos* plus the two insular regions in the complete sample¹⁶), and t is a time index, $t = 1986, 1996, 2005$.

Equation (1) gives the relationship between the initial level of education and the change in education inequality. In particular, it explains the change in inequality as a function of the beginning period level of education. Equation (2) tests a linear specification for the contemporaneous relationship between the level of education and inequality, while equation (3) tests the quadratic specification. The first model uses a single data point; models (2) and (3) use panel data for the 20 regions and three data points.

The results are presented in Tables 6 and 7. As we can see, there is a statistically significant negative relationship between the initial level of education and the change in inequality (column (a)). Each additional year of schooling in 1986 reduced the Gini coefficient by 14 percentage points for the period 1986-2005. Excluding Alentejo and Lisboa, the magnitude of the effect for this restricted sample triplicates to a value of 41 percent (column (c))¹⁷.

TABLE 6 – Regression of the evolution of education inequality on the initial level of education

Dependent variable Sample	$\Delta X_{i, 1986-2005}$	$\Delta X_{i, 1986-2005}$	$\Delta X_{i, 1986-2005}$	$\Delta X_{i, 1986-1996}$	$\Delta X_{i, 1986-1996}$	$\Delta X_{i, 1986-1996}$
	Full	Without Alentejo	Without Alen- tejo and Lisboa	Full	Without Alentejo	Without Alen- tejo and Lisboa
	(a)	(b)	(c)	(d)	(e)	(f)
Constant	0.661316** (2.2793)	1.04128*** (4.2949)	2.03402*** (5.4156)	0.263922 (1.4528)	0.518782*** (3.7739)	1.07598*** (5.0183)
$H_{i,86}$	-0.139739** (-2.4497)	-0.209132*** (-4.4357)	-0.406692*** (-5.4741)	-0.0602908 (-1.6881)	-0.106801*** (-3.9951)	-0.217686*** (-5.1326)
Number of observations	20	17	16	20	17	16
R^2	0.250029	0.567415	0.681570	0.136671	0.515519	0.652981

Notes: OLS estimations of model (1). t-statistics are in parentheses. *, ** and *** mean significance at the level of 10%, 5% and 1%, respectively.

As far as the shape of the evolution of education inequality is concerned (see table 7 and Figure 3), our results suggest an inverted-U curve. In model (2), columns (g)-(h), the coefficient is negative and statistically significant, revealing a negative relationship between the level of education and education inequality (using the Gini index). Despite the statistical significance and the large value of R^2 , the coefficient of schooling is quite small (the value for the full sample is -0.0065, which means that an increase in average schooling by one year leads to a decrease in the Gini coefficient by only 0.0065 – column (g)). In model (3), we obtain the expected signs for the coefficients on the level of education and its square, but for the full sample the coefficients are not statistically significant (column (i)). However, leaving Alentejo out of the sample, the coefficients become statistically significant (column (j)). Moreover, the results are sensitive to the initial year of the sample, since after eliminating the observations for 1986 the absolute values of the coefficients more than triplicate and the R^2 reaches almost 90% (column (k)).

One important implication of these results is that we can establish a threshold which separates two main groups of regions: those in which we observe an increase and those in which there is a

¹⁶ Lisboa is manifestly an outlier, as it presents very high levels of average years of schooling relative to all the other regions, although the differences have narrowed over the sample years. We thus exclude Lisboa from the sample in some of the estimations.

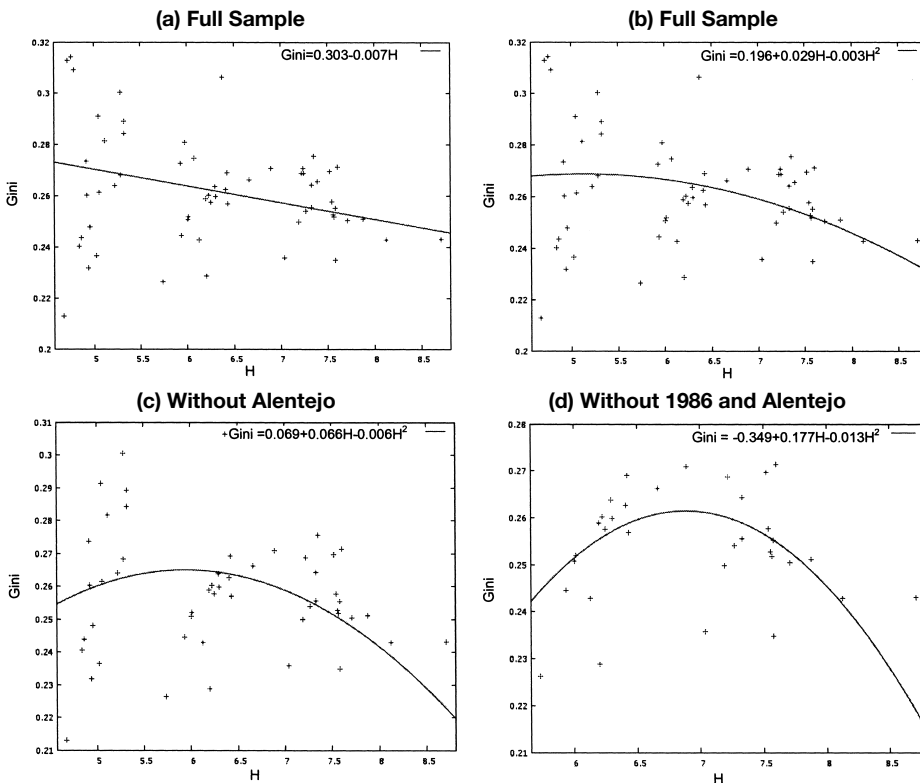
¹⁷ The results are robust to changes in the inequality measure used. For instance, the estimated coefficient for the full sample when using the Atkinson index is -0.105, but not statistically significant. The use of the Theil measure leads to a statistically significant coefficient of -0.121.

TABLE 7 – The education inequality curve

Dependent variable Sample	$X_{i,t}$	$X_{i,t}$	$X_{i,t}$	$X_{i,t}$	$X_{i,t}$
	Full	Without Alentejo	Full	Without Alentejo	Without 1986 and Alentejo
	(g)	(h)	(i)	(j)	(k)
Constant	0.302901*** (24.8751)	0.286239*** (22.0821)	0.195648** (2.6063)	0.0694279 (0.9551)	-0.349125*** (-3.2677)
$H_{i,t}$	-0.0065028*** (-3.4025)	-0.00447848** (-2.2178)	0.0285482 (1.1753)	0.0657605*** (2.8201)	0.177355*** (5.7144)
$H_{i,t}^2$			-0.00278169 (-1.4474)	-0.00552911*** (-3.0213)	-0.012879*** (-5.7597)
Number of observations	60	51	60	51	34
R^2	0.671967	0.599569	0.689107	0.688442	0.889366

Notes: Fixed effects estimations of models (2) and (3). t-statistics are in parentheses. *, ** and *** mean significance at the level of 10%, 5% and 1%, respectively.

Figure 3 – Returns to schooling in selected developed countries





decrease in inequality. Taking into account the full sample (60 observations) the threshold is estimated in 5.13 schooling years and when ignoring Alentejo (51 observations) the threshold becomes equal to 5.95 years of schooling.

4. Concluding remarks

The capacity of a country to create *new things and ideas* (innovation) and/or its ability to absorb and adapt to *new knowledge and technologies* (imitation) is a major determinant of its economic performance. Both innovation and imitation activities require a workforce endowed with a high stock of human capital. The importance of human capital thus calls for a more thorough analysis of its level and distribution.

In this paper, we investigated the distribution of human capital acquired through formal education in Portugal in great detail, filling a gap in the literature. Although Portugal still presents a low average level of education relative to other developed countries, our paper confirms the increase in average years of schooling of the workforce reported in previous studies, from 5.46 years in 1986 to 7.81 years in 2005. We also found a downward trend in education inequality for Portugal, with the Gini coefficient decreasing from 0.284 in 1986 to 0.255 in 2005, and the Atkinson coefficient and the Theil measure decreasing by 29% and 31%, respectively. Generally speaking, the decrease in the values of the three indices suggests: (i) a more equal distribution of education; (ii) lower inequality at the top end of the distribution; and (iii) lower inequality at the bottom end of the distribution, the latter indicating that a higher share of population is gaining access to intermediate and high educational levels.

We also tested whether the relationship between the level of education and education inequality during the process of increasing average education would be close to an inverted-U. The process of human capital accumulation (through formal education) can lead to an increase in education inequality, at initial and intermediate stages, and only after a certain threshold will education inequality start to decrease. The stretching of the top end of the distribution at initial stages is responsible for the increase in inequality, whilst the subsequent compression of the distribution around higher levels of education will force inequality to decrease. We found that the initially relatively high educated regions were those that experienced decreases in inequality, whereas the initially relatively poor educated regions show an opposite shift, with the threshold that separates the two groups of regions occurring at 5.13 years of schooling in the full sample. This result highlights the importance of public policies addressing education, since investments in education, besides all other positive impacts, will lead to a reduction in education inequality, after a certain threshold. This can transmit to a reduction in earnings inequality and, later, in income inequality, demanding less redistribution from the state.

Our results can also help to explain the evolution of earnings inequality in the Portuguese economy reported in previous studies. The majority of the studies have associated the increase in earnings inequality with the increase in schooling levels and have observed that this increasing inequality is related to a stretching of the top end of the distribution. Since we have shown that for some regions there was indeed an increase in education inequality this could have fostered between-group earnings inequality. However, the main impact resulted from the stretching of the top of education distribution that most regions have experienced, which might have spurred the referred stretching of the top end of the earnings distribution. Nevertheless, one should expect a future decrease in between-group earnings inequality as the continuing increase in the level of education will supposedly lead to a compression of its distribution and a reduction of its inequality (a process still occurring). Lisboa, Setúbal and Faro are the most advanced districts in this process since the top end of the respective distributions have become increasingly compressed (for instance, in the last year of the analysis, Q75 and Q90 display the exact same value). Consistent with our claims, these three districts were the ones that presented the highest decreases in education inequality, measured by the Gini coefficient.

The evidence presented in this paper is of special interest to future studies investigating the impact of inequality on growth in Portugal: assessing the direct impact of education (or human capital) inequality on growth or using it as a control variable when analysing the influence of income (earnings) inequality on growth. Nevertheless, some drawbacks of the analysis must be taken into consideration and addressed in future research: the high level of disaggregation leads us to capture very specific regional characteristics that enrich our analysis but may bias our results; and our data excludes a large group of the Portuguese workforce, public administration workers, which can potentially bias our estimations.

Further research on the subject should include the measurement of human capital inequality considering a Mincerian specification, as previous studies have found evidence that even if educational attainment is an accurate proxy of the human capital stock, using education inequality as a proxy for human capital inequality can be misleading (see e.g. Lim and Tang, 2008). Furthermore, empirical studies regarding the relationship between education or human capital inequality and earnings inequality for Portugal are needed to test whether or not the impact is positive, as expected.





References

- Aghion, P. Caroli, E. and García-Penalosa, C. (1999) Inequality and economic growth: the perspective of the new growth theories, *Journal of Economic Literature*, 37-4, 1615-1660.
- Alesina, A.; Perotti, R. (1996) Income distribution, political instability, and investment, *European Economic Review*, 40-6, 1203-1228.
- Alesina, A.; Rodrik, D. (1994) Distributive Politics and Economic Growth, *Quarterly Journal of Economics*, 109-2, 465-490.
- Atkinson, A.B. (2003) Income Inequality in OECD Countries: Data and Explanations, CESifo Group Munich, CESifo Working Paper No. 881.
- Alves, N. (2009) Novos factos sobre a pobreza em Portugal, *Boletim Económico (Primavera)*, Banco de Portugal, 125-154.
- Barro, R. J. (2000) Inequality and growth in a panel of countries, *Journal of Economic Growth*, 5-1, 5-32.
- Barro, R. J.; Lee, J.-W. (1993) International comparisons of educational attainment, *Journal of Monetary Economics*, 32-3, 363-394.
- Barro, R. J.; Lee, J.-W. (1996) International measures of schooling years and schooling quality, *American Economic Review*, 86-2, 218-223.
- Barro, R. J.; Lee, J.-W. (2001) International data on educational attainment: Updates and implications, *Oxford Economic Papers*, 53, 541-563.
- Bertola, G. (1993) Factor Shares and Savings in Endogenous Growth, *American Economic Review*, 83-5, 1184-1198.
- Birdsall, N.; Londoño, J.L. (1997) Asset inequality matters: An assessment of the World Bank's approach to poverty reduction, *American Economic Review*, 87-2, 32-37.
- Bourgignon, F.; Morrison, C. (2002) Inequality among world citizens: 1820-1992, *American Economic Review*, 92-4, 727-743.
- Budría, S. (2007) Economic inequality in Portugal: A picture in the beginnings of the 21st century, *Centro de Estudos de Economia Aplicada do Atlântico*, MPRA paper No. 1784.
- Cardoso, A.R. (1998) Earnings inequality in Portugal: High and rising?, *Review of Income and Wealth*, 44-3, 325-341.
- Carneiro, P. (2008) Equality of opportunity and educational achievement in Portugal, *Portuguese Economic Journal*, 7-1, 17-41.
- Castelló, A.; Doménech, R. (2002) Human capital inequality and economic growth: some new evidence, *The Economic Journal*, 112, 187-200.
- Checchi, D. (2001) Education inequality and income inequality, LSE STICERD Research Paper No. 52.
- De Gregorio, J.; Lee, J. (2002) Education and income inequality: new evidence from cross-country data, *Review of Income and Wealth*, 48-3, 395-416.
- Dominicis, L.; Groot, H.; Florax, R. (2006) Growth and Inequality: A Meta-Analysis, Tinbergen Institute Discussion Paper No. TI 2006-064/3.
- Duarte, A.; Simões, M. (2009) Channels of transmission of inequality to growth: A survey of the theory and evidence from a Portuguese perspective, *Estudos GEMF* No. 7, pp. 43.
- Ehrhart, C. (2009) The effects of inequality on growth: a survey of the theoretical and empirical literature, *ECINEQ Working Paper Series* No. 2009-107.

- Fields, G.S. (2001) *Distribution and Development*, Cambridge, MIT Press.
- Galor, O.; Zeira, J. (1993) Income distribution and macroeconomics, *Review of Economic Studies*, 60, 35-52.
- Gouveia, M.; Tavares, J. (1995) The distribution of household income and expenditure in Portugal: 1980 and 1990, *Review of Income and Wealth*, 41-1, 1-17.
- Kuznets, S. (1955) Economic growth and income inequality, *American Economic Review*, 45, 1-28.
- Lim, A. S. K.; Tang, K. K. (2008) Human capital inequality and the Kuznets curve, *The Developing Economies*, 46-1, 26-51.
- Lin, C. A. (2007) Education expansion, educational inequality, and income inequality: evidence from Taiwan, 1976-2003, *Social Indicators Research*, 80, 601-615.
- Lucas, R. (1988) On the mechanics of economic development, *Journal of Monetary Economics*, 22, 3-42.
- Machado, J.; Mata, J. (1998) A distribuição dos salários em Portugal: 1982-1994, *Boletim Económico(Março)*, Banco de Portugal, 31-37.
- Mankiw N.; Romer D.; Weil D. (1992) A contribution to the empirics of economic growth, *The Quarterly Journal of Economics*, 106-2, 407-437.
- Martins, P.; Pereira, P. (2004) Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11-3, 355-371.
- Morrison, C.; Murtin, F. (2007) Education inequalities and the Kuznets curves: a global perspective since 1870, *Laboratoire D'Économie Appliquée – INRA*, Working Paper No. 2007 – 12.
- Nehru, V.; Swanson, E.; Dubey, A. (1995) A new database on human capital stock in developing and industrial countries: sources, methodology, and results, *Journal of Development Economics*, 46-2, 379-401.
- OECD (2008) *Growing Unequal? Income Distribution and Poverty in OECD Countries*, Paris: OECD.
- Perotti, R. (1996) Growth, income distribution and democracy: What the data say, *Journal of Economic Growth*, 1, 149-187.
- Persson, T.; Tabellini, G. (1994) Is Inequality Harmful to Growth?, *American Economic Review*, 84-3, 600-621.
- Ram, R. (1990) Educational expansion and schooling inequality: international evidence and some implications, *Review of Economics and Statistics*, 72-2, 266-274.
- Rodrigues, C.F. (1994) Repartição do rendimento e desigualdade: Portugal nos anos 80, *Estudos de Economia*, 14-4, 399-427.
- Rodrigues, C.F. (2007) *Distribuição do rendimento, desigualdade e pobreza: Portugal nos anos 90*, Coimbra, Almedina.
- Romer, P. (1986) Increasing returns and long-run growth, *Journal of Political Economy*, 94-5, 1002-1037.
- Romer, P. (1990) Endogenous technological change, *Journal of Political Economy*, 98-5, S71-S101.
- Sianesi, B.; Van Reenen, J. (2003) The returns to education: Macroeconomics, *Journal of Economic Surveys*, 17-2, 157-200.
- Sala-i-Martin, X. (2006) The world distribution of income: Falling poverty and convergence, Period, *The Quarterly Journal of Economics*, 121-2, 351-396.





Sologon, D.; O'Donoghue, C. (2009) Earnings dynamics and inequality among men across 14 EU Countries, 1994-2001: Evidence from ECHP, IZA DP No. 4012.

Teixeira, A. (2005) Measuring aggregate human capital in Portugal: 1960-2001, *Portuguese Journal of Social Science*, 4-2, 101-120

Teixeira, A.; Fortuna, N. (2004) Human capital, innovation capability and economic growth in Portugal, 1960-2001, *Portuguese Economic Journal*, 3, 205-225.

Thomas, V.; Wang, Y.; Fan, X. (2001) A new dataset on inequality in education: gini and theil indices of schooling for 140 countries, 1960-200, World Bank research paper No. 2525.

Vieira, J., Couto, J.P. and Tiago, M.T. (2006) Inter-regional wage dispersion in Portugal, *Regional and Sectoral Economic Studies*, 6-1, 85-106.

Voitchovsky, S. (2005) Does the profile of income inequality matter for economic growth? Distinguishing between the effects of inequality in different parts of the income distribution, *Journal of Economic Growth*, 10, 273-296.

World Bank (2006) *World Development Report 2006: Equity and Development*, The World Bank, Washington D.C.

Appendix A



TABLE A.1 – Inequality indices and definitions

Inequality index	Definition
Gini index	$G = 1 + \left(\frac{1}{n}\right) - \left(\frac{2}{n2\mu}\right) \sum_{i=1}^n (n-i+1)h_i$
Atkinson index (with $\varepsilon=0.5$)	$A = 1 - \left(\frac{1}{n}\right) \left[\left(\frac{1}{n}\right) \sum_{i=1}^n (h_i^\varepsilon) \right]^{\frac{1}{1-\varepsilon}}, \varepsilon = 0.5$
Theil's first measure	$T = \left(\frac{1}{n}\right) \sum_{i=1}^n \left(\frac{h_i}{\mu}\right) \ln\left(\frac{h_i}{\mu}\right)$

Notes: h_i represents years of schooling of individual i , n is the population, μ is the average education level, ε is the Atkinson parameter for inequality aversion, and $0 < h_1 < h_2 < \dots < h_n$.

Appendix B

QP provides information on the highest schooling level completed by each worker, according to the six levels presented in Table B.1 below. The level of education of each worker corresponds to the assumed cumulated duration in years of the highest schooling level she/he has completed.

TABLE B.1 – Schooling levels classification

Level	Definition	Assumed cumulated duration in years
H0	Less than Basic 1 st cycle ¹⁸	1
H1	Basic 1 st cycle	4
H2	Basic 2 nd cycle	6
H3	Basic 3 rd cycle	9
H4	Secondary schooling	12
H5	Higher education, 1 st cycle (<i>Bacharelato</i>)	15
H6	Higher education, 2 nd cycle (<i>Licenciatura</i>)	17

Following the assumed correspondence between schooling years and each schooling level, we can compute the average years of schooling of the workforce employed in geographic unit i at time t , as

$$H_{it} = \frac{1x \sum_{j=1}^{17} L_{ijtH0} + 4x \sum_{j=1}^{17} L_{ijtH1} + 6x \sum_{j=1}^{17} L_{ijtH2} + 9x \sum_{j=1}^{17} L_{ijtH3} + 12 \sum_{j=1}^{17} L_{ijtH4} + 15 \sum_{j=1}^{17} L_{ijtH5} + 17 \sum_{j=1}^{17} L_{ijtH6}}{\sum_{j=1}^{17} L_{ijt}}$$

where H0-H6 are the schooling levels, i (geographic unit)=1,...,20; j (CAE)=1,...,17 (in 1986 there are only 9 sectors); t (year)=1986, 1996 and 2005; L_{ijts} = number of workers in geographic unit i and sector j at time t for which s (H0-H6) is the highest schooling level completed. For instance, L_{ijtH4} is the number of workers in geographic unit i and sector j at time t for which H4 is the highest schooling level completed.

18 In 2005, the description for H0 is 'Cannot read or write'.