What are the causes of educational inequality and of its evolution over time in Europe? Evidence from PISA

DOI:10.1080/09645292.2012.736475

(Article begins on next page)

What are the Causes of Educational Inequality and of its Evolution over Time in Europe? Evidence from PISA

Abstract

This paper provides evidence on the sources of differences in inequality in educational scores and their evolution over time in four European countries. Using PISA data from the 2000 and the 2006 waves, the paper shows that inequality decreased in Germany and Spain (two "decentralised" schooling systems), whilst it increased in France and Italy (two "centralised" systems). The decomposition exercise shows that educational inequality does not only reflect background related inequality, but especially schools' characteristics. These characteristics are responsible for the observed evolution over time of inequality.

JEL Codes: I2, I38

Keywords: Educational inequalities, Oaxaca decomposition, Decentralisation of educational policies

1. Introduction

Income inequality is an important measure of differences characterising individuals in a society, but it is an incomplete one since it goes very often together with inequalities in other important aspects of human life, such as education, health, housing, and political participation. However, if research on income inequality has a long tradition (e.g., Lambert 1993), economic literature has only recently shown a growing interest in inequalities in other aspects of economic well-being. A burgeoning field is the measurement of health inequalities and the decomposition of their causes (e.g., Van Doorslaer and Koolman 2004; Van Doorslaer, Koolman, and Jones 2004; Lua et al., 2007; Chen, Eastwood, and Jen 2007). In particular, following the methodology proposed by Wagstaff, van Doorslaer, and Watanabe (2003), many studies focused on the decomposition of the causes of income related health inequalities within and across countries. On the contrary, economic research on educational inequality is quite modest, even though differences in educational attainment are considered an important determinant of aggregate wage inequality. An almost unique example is provided by Sahn and Younger (2007). They decompose within and between countries inequality using TIMSS data and find that, similarly to result on health inequality, within countries inequality is greater than the between component.

While economic research has produced little analysis on the causes of educational inequality, sociological research significantly studied to which extent parental education, occupational status, or class influence children’s educational achievements across countries and over time. Shavit and Blossfeld (1993) analyze the development of inequalities in educational attainment in the 20th century, concluding that the association between family background and educational attainment has remained stable over the 20th century for all countries they analyzed except for Sweden and the

Netherlands. Pfeffer (2008) measures inequality with mobility tables drawing on data from the 'International Adult Literacy Survey'. He finds a mostly stable strong association between parental education and children educational outcomes for all nations. The finding of persistent inequality has been contested by Breen et al. (2009), who use surveys from nine European countries, collected between 1970 and 2002 on men aged 30-69, to ensure a higher degree of across countries comparability. Their findings show a clear decline in educational inequality in several countries over the course of the 20th century.

This paper tries to fill the gap in the economic literature, by expanding the discussion on inequality beyond income and health to a different dimension: education. As for health, educational inequality can be assessed on different dimensions: access to education, performance at school, and wages. We concentrate here on inequalities in educational achievement, considering PISA test scores of 15 year-olds in four European countries, representative of different educational systems as for the role of regional governments in funding and managing schools. In particular, we first decompose observed inequality into its determinants, and then analyse its evolution over time. Following the same methodology proposed by Wagstaff, van Doorslaer, and Watanabe (2003), we decompose changes in inequalities of PISA test scores into changes in the means and inequalities of the scores determinants, and changes due to the rate of return of its determinants. Our results highlight that, besides parental background, also schools' characteristics are important determinants of inequality in achievements, with heterogeneous effects on the evolution of inequality over time apparently depending on the role of local governments in funding and managing schools. Since schools' characteristics are measured here with fixed effects, it is then crucial, in a policy perspective, to improve our understanding of the black box of schools, in order to identify which factors contribute more to influence students' outcomes.

The remainder of the paper is structured as follows: Section 2 describes the methodology for the measurement of inequality. Section 3 presents the data, while results are discussed in Section 4. A brief concluding section follows.

2. Methodology

2.1. Measuring Inequality

Consider a measure of educational performance or educational outcome to assess students' abilities. An educational system is unequal if students with different socioeconomic status (SES) are characterized by different outcomes (y). The measurement of inequality is generally based on concentration indices. Let us consider the distribution of the educational outcome measures by SES. The concentration curve, labeled L in Fig. 1, plots the cumulative proportion of the population, ranked by living standards, beginning with the most disadvantaged person and ending with the richest (x-axis), against the cumulative proportion of educational attainments (y-axis). If L overlaps the 45°-line, everyone enjoys the same educational performance irrespective of her living standards. Hence, the 45°-line can be labelled as the "line of equality". On the contrary, if L lies below the 45°-line, inequality in educational performance exists and favours the richer members of society. The further L lies from the diagonal, the greater the degree of inequality in educational performance across the income distribution.

[Figure 1 here]

The concentration index, denoted by C, is defined as the ratio between the area amid L and the diagonal, and the area between the 45°-line. Since we are considering the proportion of individuals, and the proportion of outcomes, it can be easily shown that C simplifies to twice the area between L and the diagonal. More formally, C can be expressed as follows:

$$C = \frac{2}{N\mu}\sum_{i=1}^{N} R_i - 1 \qquad (1)$$

where μ is the mean of educational performance, N is the number of individuals, and R_i is the fractional rank of the *i*th individual in the living standard distribution.

In the case where there is no income-related inequality, the concentration curve overlaps with the equality line, and the concentration index takes a value of zero. If the educational measure is a "good" – like school achievement –, inequality to the disadvantage of the poor pushes C above zero and the concentration curve below the equality line. More precisely, if there are inequalities in the distribution of educational attainment, the concentration curve lies below (above) the equality line, and the concentration index takes a positive (negative) value.

2.2. Decomposing Inequality and its Evolution Over Time

To decompose the degree of inequality into the contributions of different explanatory factors, we consider here the methodology proposed by Wagstaff, van Doorslaer, and

Watanabe (2003). One needs first to specify a linear additive regression model for educational performance on a set of k determinants (x_k):

$$y_i = \alpha + \sum_k \beta_k x_{ki} + \varepsilon_i \tag{2}$$

where α and β's are coefficients to be estimated, and ε_i is a standard random disturbance term. It is then possible to decompose the concentration index of *y* by using the means and the concentration indices of the explanatory variables. In particular, from the previous equation, the relationship between these indices can be written as:

$$C = \sum (\beta \bar{x}_k / \mu) C_k + G_\varepsilon / \mu \tag{3}$$

where μ is the mean of *y*, $\bar{x}_k$ is the sample mean of x_k, C_k is the concentration index for x_k, and G_ε is a generalized concentration index for ε_i defined as $G_\varepsilon = \frac{2}{n} \sum_{i=1}^{n} \varepsilon_i R_i$. In words, the concentration index C is equal to the sum of the concentration indices of the k regressors, weighted by the elasticity of y with respect to x_k (evaluated at the sample mean). The residual component reflects the inequality in educational outcomes that is not explained by systematic variations across income groups in the determinants of outcomes x_k.

Inequality can be also considered in its evolution over time. In this case, once decomposition of inequality into its observable components has been carried out in two different time periods (*t-1* and *t*), it is interesting to decompose the observed differences over time in inequality in variations due to (i) the determinants of educational performance, and (ii) the impact of these determinants on performance. The approach proposed by Wagstaff van Doorslaer, and Watanabe (2003) consists in applying an Oaxaca-type decomposition to the expression of the concentration index C (Eq. 3). If we denote by $\eta_k = \frac{\beta_k \bar{x}_k}{\mu}$ the elasticity of y with respect to x_k at time t, and apply the Oaxaca's method, we obtain:

$$\Delta C = C_t - C_{t-1} = \sum_k \eta_{kt}(C_{kt} - C_{kt-1}) + \sum_k C_{kt-1}(\eta_{kt} - \eta_{kt-1}) + \Delta(GC_{\varepsilon t} / \mu) \qquad (4)$$

which can be alternatively rewritten as:

$$\Delta C = \sum_k \eta_{kt-1}(C_{kt} - C_{kt-1}) + \sum_k C_{kt}(\eta_{kt} - \eta_{kt-1}) + \Delta(GC_{\varepsilon t} / \mu) \qquad (4\text{ bis})$$

The approach allows identifying for each x_k the extent to which changes in educational inequality are due to changes in inequality in the determinants of educational performance (the first term on the right-hand side of each equation), rather than in their elasticities (the second term on the right-hand side).

3. Data

To assess educational inequality and its causes, we need a measure of educational outcome and a set of its determinants. We consider here information provided by the Program for International Student Assessment (PISA), a study conducted by the

OECD every three years since 2000 in order to obtain an internationally comparable database on the competencies of 15 year-old students in reading, math, science, and problem solving across countries. Data used in this paper originate from two waves, 2000 and 2006. 15-year-old students in each country had to be enrolled in an educational institution, regardless of the grade level or type of institution. The survey is not structured as a panel; rather schools are randomly selected at every round. Sampling procedures are, however, designed to allow comparability of test scores both over time and across countries.[1]

In addition to the performance tests, students as well as schools' teacher heads answered respective questionnaires, yielding rich background information on students' individual characteristics and family origins, as well as on schools' resources endowment and educational practices.

While the international dimension of the PISA data set can help assess the empirical contribution to inequality of test score determinants, the data as usual leaves much to be desired in other respects. Some survey questions and the unit scale used to measure the same variables, have changed from the 2000 to the 2006 survey, reducing the number of explanatory variables that we can use when pooling the two surveys. This is especially true at the school level: for this reason, we account for school level variables with school fixed effects.[2]

---

[1] In particular, the sample design is generally referred to a two-stage stratified sample: the first-stage sampling units consist of individual schools with eligible students, allowing for a minimum of 150 schools in each country; the second-stage sampling units are eligible students within sampled schools, selected to achieve a minimum of 4,500 students. For additional details, see OECD (2005). On the robustness of findings in economic research based on students achievement tests, see Hanushek and Woessman (2011).

[2] We had hoped to isolate the contribution of inequalities in different regions within the same country but this proved not possible: data on individuals' region of residence were collected only for 2006, and even then were collected only for some Italian regions.

The relevant notion of competencies assessed in PISA concerns knowledge and skills that can be applied in real world settings. In particular, our measure for educational performance is the score obtained in the reading test. Reading literacy is defined as "understanding, using and reflecting on written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society" (OECD 2003). The PISA definition clearly goes beyond literal comprehension, and takes into account the role of reader in gaining meaning from written texts. There are two main reasons for considering reading literacy. First, poor reading abilities are consistently associated with poor performances in the labour market, but also in other domains of well-being, such as family life, health, and civic engagement (e.g., Dugdale and Clark 2008). Second, among the four domains tested in PISA, only reading literacy is fully comparable between 2000 and 2006 (OECD, 2010; Jerrim, 2011).[3]

The measure for the socioeconomic status (the ranking variable) is the sum of the father and mother socioeconomic index of occupational status. Absent information on income at the family level, not collected by PISA, the index captures the attributes of occupations that convert parents' education into income.[4] Taking the sum rather than one parental indicator allows us to reduce problems involved with repetitive values of the indicator of economic welfare (Chen and Roy 2008).

---

[3] To be sure, however, that our conclusions are not driven by the choice of the peculiar domain used to measure students' achievements, we re-run our exercise using mathematical literacy, despite it is comparable only between 2003 and 2006. Results are fully consistent with those presented below. They are not reported here for brevity, but are available upon request from the authors.

[4] The index was derived using an optimal scaling procedure that assigns scores to each of 271 distinct occupation categories in such a way as to maximize the indirect effect of education on income through occupation and to minimize the direct effect of education on income, net of occupation (both effects being net of age). See Ganzeboom, De Graaf, and Treiman (1992) for further details on this methodology.

As already discussed, the explanatory variables included in the analysis are limited by their availability in the two surveys. In particular, we consider the following: students' gender, defining a dummy variable equal to one for females; students' background, defining a set of dummies for the highest level of education completed by both parents; a dummy for being born in another country; dummies for the grade at which the student is enrolled; and, finally, school fixed effects. Whilst the first set of variables captures characteristics at the individual and family level, school fixed effects clearly capture characteristics at the school level.

4. Results

4.1. The Evolution of Inequality

Table 1 presents the concentration indices of the PISA reading test score by years and the difference over time for a number of Western European countries included in both the 2000 and the 2006 waves. We consider, in particular, the following member states of the European Union: France, Germany, Greece, Italy, Portugal, Spain, Sweden and the United Kingdom. Students have been ranked by the sum of paternal and maternal socioeconomic index of occupational status, our measure of SES. The value of the concentration index is always positive, meaning that - as the measure is attainment - high scores are concentrated amongst the better off. The closest the values are to zero the less concentration in the distribution of the reading score is observed. From the point of view of the level of inequality, there are no clearly identifiable patterns emerging from the data. Italy and Spain, together with Sweden, present the lowest concentration indices; Portugal and Greece, two other Mediterranean countries, the highest. A positive value of the difference over time means that concentration has increased and inequality rose, whilst the contrary holds when the value of the difference is negative. Germany displays the highest reduction in inequality, followed by Spain. In France, Italy, and Greece inequality has increased, as in Sweden but to a lesser extent.[5]

An interesting insight to interpret this evolutionary pattern is that Germany and Spain are two countries where regional governments play a substantial role in the provision of education, not only in terms of managing and running schools, but also - and more importantly - in terms of education spending and funding. Also in Sweden, provision of education is decentralised, but differently from Germany and Spain, municipalities are in charge of funding schools, both public and private ones (e.g., Böhlmark and Lindahl, 2008). On the contrary, in France, Greece, and Italy education policies are (almost) completely centralised, especially in terms of spending and funding, with very little autonomy for each school.

The link between decentralisation and inequality has recently received attention in the literature, besides the traditional connection between decentralisation and efficiency (e.g., Rodrìguez-Pose and Ezcurra 2010). At the theoretical level, there are arguments supporting the view that decentralisation may harm equality, and the opposite view, that decentralisation may favour equality. As for the former, poorer regions may not have resources to exploit the potential benefits associated to the better matching between policies and local needs, and may be forced to deliver basic goods and services in a less efficient way, because of the loss of economies of scale. Further,

[5] To check the robustness of our results, we compute the concentration index ranking individuals by the sum of maternal and paternal educational level. Differences over time display the same variation in France, Germany, Italy, Spain and Sweden, while results slightly change for the other countries. The socio-economic index of occupational status is a preferable and more reliable indicator of economic welfare because it has a higher number of categories (271) than the educational level (6), which reduces problems related to its repetitive values.

decentralisation may imply a loss of political influence for poorer regions in the allocation of federal funds, resulting in fewer resources available at the local level. As for the opposite view, decentralisation may imply greater transparency for the differences in the provision of services across jurisdictions. This should put the central government and the sub-national governments under pressure to improve both efficiency and equality in the delivery of services. In addition, the expected competition across jurisdictions should promote improvements in efficiency, to reach a common standard. The arguments linking decentralisation and equality are likely, however, to operate differently in different institutional environments. In particular, the argument for increasing inequality seems to be more justified in poorer and less developed countries, where local governments may be called to provide basic services to citizens. On the contrary, the argument for decreasing inequality is more likely to be verified in richer and well-developed countries, characterized by lower territorial inequality. Indeed, by analyzing regional per capita GDP as a measure of inequality, Rodrìguez-Pose and Ezcurra (2010) empirically show these differences between the two groups of developed and developing countries. The evidence presented here is in accordance with the view that decentralised systems of education perform better than centralised systems in terms of reducing students' achievements inequality in developed countries; and this may contribute to explain also their performance in terms of lower inequality in the income per capita. This story is confirmed also by Causa and Chapuis (2009), using PISA data on OECD countries: their regression analysis confirms that decentralised schooling systems are positively associated with equity in educational achievement.

In what follows, we concentrate on the cases of France, Germany, Italy, and Spain, since they display some of the greatest changes in inequality over the period under consideration, and – more importantly - they are representative of different school systems. France and Italy feature "centralised" schooling systems as opposed to the German and the Spanish "decentralised" systems.

Figure 2 presents the difference between the concentration curve and the equality line plotted against the cumulative percentage of the sample, ranked by the sum of paternal and maternal socio-economic index of occupational status for Germany, Spain, France and Italy in 2000 and 2006. If the concentration curve for one period lies everywhere closer to the x-axis than the other, the first curve dominates the second and the ranking by degree of inequality is unambiguous. Non-dominance emerges when the concentration curves cross. Fig. 2 confirms the descriptive evidence presented above on the evolution of inequality: the 2006 concentration curve for Germany and Spain dominates the 2000 concentration curve, suggesting that inequality reduced over time for each individual in the i*th* ranking from 2000 to 2006. The opposite occurred in France and Italy, where the 2000 concentration curves dominates the 2006 ones, indicating an increase of inequality over time.

[Table 1 and Figure 2 here]

4.2. Regression Results

To explain variations in the PISA test score, we adopt a standard educational production-type regression framework, in which the student's test score is specified to be a linear function of a vector of student-level variables, X_1, a vector of household-level variables, X_2, and a commune fixed effect at the level of the student's school. Specifically, the model we estimate is:

$$y_i = \alpha + X_1\beta_1 + X_2\beta_2 + \varepsilon_i \tag{2}$$

The X_1 vector includes the students' gender, a dummy for being born in another country, and a set of dummies indicating the grade at which the student is enrolled. The vector X_2 includes dummies for the highest level of education completed by both parents at three levels: recognised third level education (ISCED 5–7), second stage of secondary level of education (ISCED 3), and less than second stage of secondary education (ISCED 0–2), treated as the omitted category; a dummy variable for not speaking the national language at home. Finally, α denotes the average school fixed effect, conditional on the two sets of observables X_1 and X_2.

The production function regression results are presented in Table 2. The constant term captures the school fixed effects. In particular, according to our specification, it represents the across school average PISA test score of male students enrolled in the lowest grade whose parents have less than second stage of secondary education. Its value is slightly lower than the average country reading score, which represents the unconditional mean (see the last row in Table 2-6). An increase of the constant term indicates that the average test score across schools rose, while the opposite holds in case of a decline. The mean of the school fixed effects increases between 2000 and 2006 in Germany and Spain, the two countries that display over time a reduction in concentration indexes, whilst it decreases in France and Italy, which display increases in concentration index between the two years. The increase registered in Germany and Spain is also statistically significant: the hypothesis of constant school fixed effect is rejected at 5% level in Germany and Spain, but it is not in France and Italy. Apparently, an increase in the average test score goes together with a reduction in the observed inequality.

The fact that the pattern of the average fixed effect mirrors the evolution of the concentration index suggests that school characteristics are important factors in the evolution of inequality. In the attempt to better understand their role, we also report in Table 2 (last rows) the two variance components, together with the share of variance in reading scores due to between school variation (rho), which gives us additional information on the role of school fixed effects. While Germany displays the highest level of between school variation, Spain features the lowest, suggesting that institutional heterogeneity across schools is larger in the former than in the latter country. Correspondingly, in Germany, within school variation is a small component of the total variance in test scores, while in the other countries it accounts for a larger share of total variance. As for the evolution over time, the between school component of variance increased in all countries, especially in Germany and France. On the contrary, changes in within school variance are not sizeable in all countries, but Italy. These differences show that characteristics at the school level may play an important role in affecting educational inequality, despite the heterogeneity of the institutional context schools operate in.

A potentially relevant shortcoming of our strategy is the fact that it cannot account for the role of family unobservable characteristics in school selection. If parents who care about school quality sort their children in better schools, the school level fixed effect would capture both unobservable parental and school characteristics. The extent to which this is true depends on the ability of observable variables to assess school selection. As Altonji, Elder, and Taber (2005) point out, if the observable variables have enough explanatory power on the outcome variable, then the effect of unobservable variables cannot be too large. As Table 2 shows, this is our case: the variables in the X_2 vector have sizeable coefficients and are almost all significant, except in the France 2006 regression.

Girls get lower reading scores than boys do, and the gender gap widened over the period 2000-2006 in Germany, Spain, and Italy, whilst it slightly narrowed in France. The estimated coefficients of maternal secondary and tertiary educational level are positive across all countries in 2000 (though not significant for Italy). However, the absolute values of the estimated coefficients decline over time, and become less significant in Spain and not significant in France in 2006. The coefficients of paternal educational level are not highly significant in Germany, France and Italy in 2000, while they are in Spain. Maternal education has a sizeable and significant effect in all countries except for Italy. Paternal education becomes positive and significant in Germany over time, whilst coefficients decline in France, Italy, and Spain.

Speaking a non-national language at home always has a negative and significant effect on reading score, but whilst the gap falls in Germany and France, becoming not significant in France, it slightly rises in Italy. The coefficient estimates for the immigrant dummy variable are all significant and negative for all countries but Spain, and for all years (although almost zero and not significant in Italy in 2000), indicating that immigrants children achieve lower educational scores status. The declining over time pattern of the coefficient observed in Germany and France may reflect the increasing integration of immigrants in the school system, whilst the opposite is observed in Italy and Spain, where the gap either widened or became significant over the period under consideration. Being enrolled in higher grades is always associated with higher reading scores, with the effect increasing over time in France, and slightly decreasing in the other countries.

We test time invariant slope coefficients based on zero slope effects from regression on pooled sample, which includes observation from the 2000 and the 2006 surveys. The joint hypothesis of time-invariant slope coefficient is rejected at the 1% level for all countries.

The chosen specification assumes that all individuals within the same country face the same coefficient vectors irrespective of their socio-economic status. The reason is that the regressions already control for mother and father's educational level, which capture the parental socio-economic status. Thereby, the vector b_k includes the effect of each variable once we have netted out the effect of parental education. However, to check the importance of interactions between our control variables and parental educational level dummies we re-estimate a specification including these additional covariates. An F-test does not reject the null hypothesis of no joint significance of the interaction terms, suggesting that we can exclude these variables from the model.

The results of the production function are qualitatively coherent with the existing burgeoning literature on PISA scores, although specifications usually differ from ours in that explanatory variables are included in place of school fixed effects. The impact of parental education on school performance has a steeper coefficient in Germany, whilst a flatter one in France, Spain, and Italy, coherently with the findings in Baumert, et al. (2003), Stanat (2003), and Entorf and Minoiu (2004). Similarly, the educational gap between natives and students born in other countries, and the gap between students with the national language as their major language spoken at home and students from foreign language speaking backgrounds are higher in Germany than in France, Spain, and Italy (Entorf and Minoiu 2004 and Ammermueller 2007).

[Table 2 here]

4.3. Decomposition Results

Table 3, 4, 5, and 6 present the decomposition of our measures of inequality for the two years and each of the four countries. The first two columns of the tables show the coefficients of the production function estimates as presented in Table 2. The third and fourth columns show the mean of the dependent variables for 2000 and 2006; the fifth and sixth columns present the elasticties $\eta_{ki} = \frac{\beta_k \bar{x}_k}{\mu}$, whilst the last two columns indicate the values of the concentration indices of the explanatory variables. A positive (negative) value of the concentration index suggests a pro-rich (pro-poor) distribution of the x_k determinants of inequality. In almost all countries, parental secondary education shows a pro-poor distribution, while tertiary education highlights exactly the opposite. This suggests that parental background is better amongst the better-off. Notice that Italy is somewhat an exception, since in 2006 also parental secondary education shows a pro-rich distribution. Moreover, as for parental background, the level of pro-rich inequality is higher in Italy and Spain than in the other two countries.

Table 7 gives the decomposition results based on the Oaxaca decomposition, which are estimates of the contributions of the explanatory variables to the concentration indices as well as the change between 2000 and 2006. Looking first at the contribution of observable factors to the concentration index, being a girl has no contribution to the inequality in all the four countries. Parental tertiary education disfavours the poor in the 2000 surveys, but the effect almost vanishes in the 2006 survey. The effect of being born in another country or speaking a different language is null in all countries and years, although speaking a different language slightly disfavours the poor in Italy.

Being enrolled in higher grade increases inequality towards the poor. The effect is stronger in Germany, Italy and Spain, lower in France.

However, the bulk of inequality in education in both 2000 and 2006 was caused by inequality in school fixed effects, disfavouring the poor. The inference is that in both years, poor pupils went to schools that were likely to have worst characteristics with respect to those attended by rich pupils.

The contribution of school fixed effects is especially sizeable in Germany, the country characterized by the highest between school variance. The effect rises over time in France and Italy, the two "centralised" systems, and it decreases over time in Spain and Germany. The column headed "Change 2000-2006" shows that the reduction of inequality in Germany and Spain between 2000 and 2006 was mainly due to changes at the school level, whilst the same variable was responsible for the increase in inequality in Italy and France. This result suggests that centralised schooling systems may not be a guarantee for equality in the provision of educational services, even in countries where the role of public schools is overwhelming, as in Italy and France.[6]

However, the decomposition in Table 7 does not enable us to see whether these changes are due to changes in the elasticities rather than to changes in inequality. The Oaxaca decomposition in Table 8 addresses this issue. For a given observable variable, the columns with header "$\Delta C\times\eta$" are the contributions of the respective explanatory variable to the change in inequality in the total concentration index due to the change of the concentration index of the explanatory variable itself. The columns with header "$\Delta\eta\times C$" indicate the contribution due to the change in the elasticity of the explanatory variable. The column "Total" corresponds to the sum of the two

[6] Evidence on the inequality of opportunity in the Italian schooling system is provided, e.g., by Checchi and Peragine (2010). The authors highlight the importance of regional differences as opposed to cross-school differences.

components, and coincides with the column headed “Change” in Table 7, but for the residual (not included here). Finally, the “%” column computes the relative contribution of each observable variable to the total concentration. Interesting insights emerge from the decomposition exercise.

First, looking at school fixed effects, in all countries changes in inequality and changes in elasticities go in opposite directions, but the former offset the latter. This suggests that it is the change in inequality – rather than elasticity – that accounts for the bulk of the variation in inequality associated with school characteristics. In particular, while Germany and Spain show consistent reductions in inequality at school level, Italy and France show an increase in between school inequality. This may appear to be in contrast with the evolution of between school variance in Germany, which rose consistently over the period 2000 2006 as reported at the bottom of Table 2. However, variance is not scale invariant, meaning that it increases if test scores increase proportionately, as it is the case for Germany over the period under consideration.

These results further reinforce findings by Rodrìguez-Pose and Ezcurra (2010) and Causa and Chapuis (2009) on the relationship between inequality and decentralisation. Of course, cross-school variations may reflect cross-regional variations (e.g., Checchi and Peragine, 2010, for Italy), but the general conclusion that centralisation appears to hamper the reduction of territorial differences remains, and opens interesting issues on the impact of decentralisation on school sorting and stratification by students ability, which has been shown to persist in the mostly public UK school system (e.g., Gibbons and Telhaj, 2007).

Second, changes in parental background inequality appear to be unimportant in all countries but Germany. However, in Germany, the reduction of inequality in mother tertiary education and in father secondary education is approximately compensated by the increase in inequality in mother secondary education and in father tertiary education, so that the net effect of inequality in parental background is negligible. Looking at changes in the elasticities for parental background, we record in all countries a reduction in the elasticity of mother and father tertiary education (with the exception of Germany for fathers). Hence, in all countries but Germany, parental background is contributing to reduce inequality more via its reduced impact on attainments than via changes in inequality in parental background.

5. Conclusions

This paper analyses the determinants and the evolution of educational inequality by comparing PISA test scores in reading literacy in the 2000 and the 2006 waves. We first measure concentration indices in a number of different European countries, showing that inequality has been mostly reducing in “decentralised” schooling systems, while it is increasing in “centralised” ones. Concentrating on France, Italy, Germany and Spain as representatives of typical school systems for the role played by regional governments, we then decompose observed inequality into its causes, and analyse its evolution over time. In particular, following Wagstaff et al. (2003), we decompose change in educational inequality on standardised tests into shares due to changes in the mean and inequality of the determinants of educational outcomes, and changes due to the rate of return of its determinants. Our results highlight that, besides parental background, also schools’ characteristics are important determinants of inequality in achievements among students. We observe a positive contribution toward reducing inequality coming from the evolution of inequality across schools in Germany and Spain, while the opposite occurs in France and Italy.

This decomposition exercise allows policy makers to target areas that may make the largest contribution to reducing educational inequality. As schools are treated in this exercise as fixed effects, it is crucial in a policy perspective to open the black box of schools, in order to understand what the most important determinants of inequality between schools are. This is left for future research.

References

Altonji, J., T. Elder, C. Taber 2005. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *Journal of Political Economy* 113: 151-184.

Ammermueller A. 2007. Poor background or low returns? Why immigrant students in Germany perform so poorly in the Programme for International Student Assessment. *Education Economics* 15: 215-230.

Baumert, J., C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, G. Schümer, P. Stanat, K.-J. Tillmann and M. Weiß 2003. PISA 2000. *Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Zusammenfassung zentraler Befunde*. Berlin: Max-Planck-Institut für Bildungsforschung.

Blossfeld H., Y. Shavit 1993. Persistent inequality: changing educational attainment in thirteen countries, Boulder, Colo.: Westview.

Böhlmark A., M. Lindahl 2008. Does School Privatization Improve Educational Achievement? Evidence from Sweden's Voucher Reform. IZA Discussion Paper No. 3691.

Breen R., R. Luijkx, W. Muller, R. Polla 2009. Nonpersistent inequality in educational attainment: evidence from eight European countries. *The American Journal of Sociology* 114: 1475-1521.

Causa O., C. Chapuis 2009. Equity in student achievement across OECD countries: An investigation of the role of policies. OECD Economics Department Working Papers No. 708.

Checchi D., V. Peragine 2010. Regional disparities and inequality of opportunity: the case of Italy. *Journal of Economic Inequality* 8: 429-450.

Chen, Z. and K. Roy 2009. Calculating concentration index with repetitive values of indicators of economic welfare. *Journal of Health Economics* 28: 169-175.

Chen Z., D. Eastwood, S. Yen 2007. A decade's story of childhood malnutrition inequality in China: where you live does matter. *China Economic Review* 18: 139–154.

Dugdale, G. and Clark C. (2008). *Literacy changes lives: An advocacy resource*. London: National Literacy Trust.

Entorf H. and N. Minoiu 2004. What a difference immigration law makes: PISA results, migration background, socioeconomic status and social mobility in Europe and traditional countries of immigration. ZEW Discussion Paper No. 04-17.

Ganzeboom H., P. De Graaf, D. Treiman 1992. A standard international socio-economic index of occupational status. *Social Science Research* 21: 1-56.

Gibbons S., S. Telhaj 2007. Are schools drifting apart? Intake stratification in English Secondary schools. *Urban Studies* 44 (7): 1281-1305.

Hanushek E. and Woessman L. 2011. Sample Selectivity and the Validity of International Student Achievement Tests in Economic Research. *Economics Letters* 110 (2): 79-82.

Jerrim J. 2011. England's “plummeting” PISA test scores between 2000 and 2009: Is the performance of our secondary school pupils really in relative decline? DoQSS Working Paper No. 11-09. London: Institute of Education, University of London.

Lambert P. J. 1993. The distribution and redistribution of income. Manchester: Manchester University Press.

Lua J., G. Leung, S. Kwon, K. Tin, E. Van Doorslaer, O. O’Donnell 2007. Horizontal equity in health care utilization evidence from three high-income Asian economies. *Social Science and Medicine*, 64: 199–212.

OECD 2003. The PISA 2003 Assessment Framework. Mathematics, Reading, Science and Problem Solving Knowledge and Skills. OECD: Paris.

OECD 2005. School sampling preparation manual. PISA 2006 main study, OECD: Paris.

OECD 2010. PISA 2009 Results: Learning Trends: Changes in Students Performance Since 2000. OECD: Paris.

Pfeffer F. 2008. Persistent inequality in educational attainment and its institutional context. *European Sociological Review* 24: 543–565.

Rodrìguez-Pose A., R. Ezcurra 2010. Does decentralization matter for regional disparities? A cross-country analysis. *Journal of Economic Geography* 10: 619-644.

Sahn D., S. Younger 2007. Decomposing World Education Inequality. Cornell Food and Nutrition Policy Program Working Paper No. 187.

Stanat, P. 2003. Schulleistungen von jugendlichen mit migrationshintergrund: differenzierung deskriptiver befunde aus PISA und PISA-E. In: Ein differenzierter blick auf die länder der bundesrepublik Deutschland. 243-260. Opladen: Deutsches PISA Konsortium.

Van Doorslaer E., X. Koolman 2004. Explaining the differences in income-related health inequalities across European countries. *Health Economics* 13: 609–628.

Van Doorslaer E., X. Koolman, A. Jones 2004. Explaining income-related inequalities in doctor utilisation in Europe: a decomposition approach. *Health Economics* 13: 629-647.

Wagstaff A., E. van Doorslaer, N. Watanabe 2003. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics*, 112: 207 – 223.

Table 1. Concentration index, 2000 and 2006 (std. dev. in brackets)

	2006	2000	Difference
France	0.0380	0.0327	**0.0053**
	[0.0018]	[0.0016]	
Germany	0.0348	0.0472	**-0.0124**
	[0.0020]	[0.0023]	
Italy	0.0288	0.0233	**0.0055**
	[0.0016]	[0.0019]	
Greece	0.0444	0.0390	**0.0054**
	[0.0022]	[0.0024]	
Portugal	0.0442	0.0438	**0.0005**
	[0.0023]	[0.0020]	
Spain	0.0289	0.0332	**-0.0043**
	[0.0017]	[0.0016]	
Sweden	0.0330	0.0303	**0.0026**
	[0.0018]	[0.0016]	
Great Britain	0.0376	0.0374	**0.0002**
	[0.0015]	[0.0016]	

Table 2. Production function (dependent variable: read score)

	Germany		France		Italy		Spain	
	2000	*2006*	*2000*	*2006*	*2000*	*2006*	2000	2006
Female	-15.92	-27.968	-16.012	-15.156	-14.159	-18.772	-14.158	-21.985
	[2.159]***	[2.103]***	[2.039]***	[2.187]***	[2.205]***	[1.901]***	[1.970]***	[1.750]***
Mother second educ	35.626	17.362	10.156	3.89	5.144	4.297	7.018	3.575
	[10.941]***	[3.433]***	[4.445]**	[3.062]	[3.23]	[1.957]**	[2.959]**	[2.176]
Mother tertiary educ	47.571	21.18	14.502	-1.437	5.643	1.894	12.614	3.595
	[11.237]***	[3.762]***	[4.666]***	[3.605]	[3.532]	[2.919]	[2.926]***	[2.568]
Father second educ	-18.2	10.415	6.072	-0.207	7.649	0.632	9.785	4.758
	[9.469]*	[3.462]***	[4.038]	[2.91]	[3.382]**	[1.926]	[2.994]***	[2.272]**
Father tertiary educ	-18.452	5.359	6.967	2.262	7.947	-4.228	16.904	4.630
	[10.056]*	[3.551]	[4.224]*	[3.442]	[3.738]	[2.926]	[2.761]***	[2.500]**
Not national language spoken at home	-27.012	-12.312	-13.658	-2.098	-7.829	-9.566	-3.748	-5.097
	[7.693]***	[4.826]**	[-5.029]***	[6.012]	[3.072]**	[2.58]***	[3.526]	[3.457]
Born in another country	-24.334	-14.584	-8.789	-5.509	0.765	-5.572	-5.922	-5.750
	[3.575]***	[3.177]***	[2.555]***	[2.818]*	[4.109]	[2.943]*	[4.683]	[3.379]*
Grade 8	26.907	20.701						
	[11.545]**	[12.258]*						
Grade 9	69.979	48.389	45.74	38.533			60.748	51.061
	[11.545]***	[12.258]***	[4.413]***	[6.178]***			[6.517]***	[4.440]***
Grade 10	105.151	82.548	106.332	123.521	44.669	41.399	145.614	122.752
	[11.471]***	[12.126]***	[7.508]***	[9.548]***	[3.03]***	[2.707]***	[6.389]***	[4.219]***
Grade 11			149.983	180.944	71.229	68.616		
			[9.43]***	[10.768]***	[5.037]***	[5.709]***		
School fixed effects	421.528	451.827	424.522	413.765	449.178	447.675	367.699	383.236
	[13.203]***	[12.069]***	[7.059]***	[8.069]***	[4.558]***	[2.955]***	[6.380]***	[4.378]***
Observations	4,158	3,893	4,158	3,893	4,705	20,433	5700	16921
R-squared	0.649	0.688	0.649	0.688	0.55	0.545	0.480	0.460
F test: No change in intercept	6.300		0.930		0.260		12.590	
Prob > F	0.012		0.336		0.611		0.000	
F test: No change in slope coefficients	5.260		2.990		5.290		12.110	
Prob > F	0.00		0.00		0.00		0.000	
sigma_u (between var)	69.81	88.83	39.35	53.81	56.70	72.67	30.447	33.756
sigma_e (within var)	60.87	60.17	60.05	62.68	60.44	70.71	61.140	63.052
rho (share between school var.)	0.57	0.69	0.3	0.424	0.468	0.514	0.199	0.223

Robust standard errors in parentheses. * significant at 5%; ** significant at 1%. F-test to test hypothesis of no change in mean school fixed effect between 2000 and 2006 is based on school fixed effects obtained from regression on 2000 sample. F-test for the joint hypothesis of no change in slope coefficient is based on interactions between X-variables and time dummy in pooled sample.

Table 3. Germany - Inequality Decomposition for 2000 and 2006

	Coefficients		*Mean*		*Elasticities*		*Concentration indices*	
	2000	2006	2000	2006	2000	2006	2000	2006
Mean read score			483.969	499.662				
Female	-15.920	-27.968	0.498	0.517	-0.016	-0.029	0.010	0.021
Mother second educ	35.626	17.362	0.650	0.538	0.048	0.019	-0.137	-0.067
Mother tertiary educ	47.571	21.180	0.302	0.286	0.030	0.012	0.348	0.255
Father second educ	-18.200	10.415	0.524	0.422	-0.020	0.009	-0.227	-0.129
Father tertiary educ	-18.452	5.359	0.445	0.406	-0.017	0.004	0.290	0.236
Not national language spoken at home	-27.012	-12.312	0.072	0.064	-0.004	-0.002	-0.272	-0.234
Born in another country	-24.334	-14.584	0.210	0.165	-0.011	-0.005	-0.168	-0.214
Grade 8	26.907	20.701	0.147	0.116	0.008	0.005	-0.171	-0.231
Grade 9	69.979	48.389	0.606	0.568	0.088	0.055	0.005	-0.018
Grade 10	105.151	82.548	0.235	0.302	0.051	0.050	0.108	0.132
School fixed effects	421.528	451.827	483.969	499.662	4.215	4.518	0.008	0.006

Table 4. France - Inequality Decomposition for 2000 and 2006

	Coefficients		*Mean*		*Elasticities*		*Concentration indices*	
	2000	2006	2000	2006	2000	2006	2000	2006
Mean read score			504.899	491.203				
Female	-16.012	-15.156	0.484	0.486	-0.015	-0.015	0.012	0.017
Mother second educ	10.156	3.890	0.445	0.482	0.009	0.004	-0.222	-0.086
Mother tertiary educ	14.502	-1.437	0.475	0.308	0.014	-0.001	0.258	0.340
Father second educ	6.072	-0.207	0.457	0.450	0.005	0.000	-0.219	-0.123
Father tertiary educ	6.967	2.262	0.461	0.328	0.006	0.001	0.271	0.383
Not national language spoken at home	-13.658	-2.098	0.048	0.055	-0.001	0.000	-0.096	-0.078
Born in another country	-8.789	-5.509	0.250	0.232	-0.004	-0.003	-0.020	0.053
Grade 9	45.740	38.533	0.367	0.337	-0.021	-0.019	-0.317	-0.264
Grade 10	106.332	123.521	0.535	0.587	-0.076	-0.097	-0.173	-0.195
Grade 11	149.983	180.944	0.027	0.025	-0.046	-0.069	0.140	0.112
School fixed effects	424.522	413.765	504.899	491.203	4.245	4.138	0.003	0.004

Table 5. Italy - Inequality Decomposition for 2000 and 2006

	Coefficients		*Mean*		*Elasticities*		*Concentration indices*	
	2000	2006	2000	2006	2000	2006	2000	2006
Mean read score			489.242	472.920				
Female	-14.159	-18.772	0.501	0.492	-0.014	-0.020	0.006	0.015
Mother second educ	5.144	4.297	0.429	0.424	0.005	0.004	-0.272	0.009
Mother tertiary educ	5.643	1.894	0.444	0.192	0.005	0.001	0.348	0.562
Father second educ	7.649	0.632	0.444	0.408	0.007	0.001	-0.261	0.007
Father tertiary educ	7.947	-4.228	0.444	0.194	0.007	-0.002	0.362	0.589
Not national language spoken at home	-7.829	-9.566	0.169	0.122	-0.003	-0.002	-0.263	-0.248
Born in another country	0.765	-5.572	0.057	0.097	0.000	-0.001	0.076	-0.106
Grade 10	44.669	41.399	0.776	0.151	0.071	0.072	-0.007	0.020
Grade 11	71.229	68.616	0.059	0.820	0.009	0.004	0.438	0.486
School fixed effects	449.178	447.675	489.242	472.920	4.492	4.477	0.004	0.005

Table 6. Spain - Inequality Decomposition for 2000 and 2006

	Coefficients		*Mean*		*Elasticities*		*Concentration indices*	
	2000	2006	2000	2006	2000	2006	2000	2006
Mean read score			492.840	466.594				
Female	-14.158	-21.985	0.482	0.508	-0.014	-0.024	0.043	0.011
Mother second educ	7.018	3.575	0.209	0.292	0.003	0.002	-0.112	-0.010
Mother tertiary educ	12.614	3.595	0.336	0.270	0.009	0.002	0.444	0.482
Father second educ	9.785	4.758	0.194	0.290	0.004	0.003	-0.133	-0.047
Father tertiary educ	16.904	4.630	0.399	0.305	0.014	0.003	0.398	0.455
Not national language spoken at home	-3.748	-5.097	0.139	0.159	-0.001	-0.002	0.045	-0.092
Born in another country	-5.922	-5.750	0.065	0.120	-0.001	-0.002	0.053	-0.059
Grade 9	60.748	51.061	0.253	0.310	0.031	0.034	-0.214	-0.224
Grade 10	145.614	122.752	0.725	0.631	0.214	0.166	0.083	0.139
School fixed effects	367.699	383.236	492.840	466.594	3.677	3.832	0.003	0.002

Table 7. Contribution of observables and change between 2000 and 2006

	Contribution to C								Change 2006 - 2000			
	Germany		France		Italy		Spain		Germany	France	Italy	Spain
	2000	2006	2000	2006	2000	2006	2000	2006				
Female	-0.0002	-0.0006	-0.0002	-0.0003	-0.0001	-0.0003	-0.0006	-0.0003	-0.0005	-0.0001	-0.0002	0.0003
Mother second educ	-0.0065	-0.0013	-0.0020	-0.0003	-0.0012	0.0000	-0.0003	0.0000	0.0053	0.0017	0.0013	0.0003
Mother tertiary educ	0.0103	0.0031	0.0035	-0.0003	0.0018	0.0004	0.0038	0.0011	-0.0072	-0.0038	-0.0013	-0.0028
Father second educ	0.0045	-0.0011	-0.0012	0.0000	-0.0018	0.0000	-0.0005	-0.0001	-0.0056	0.0012	0.0018	0.0004
Father tertiary educ	-0.0049	0.0010	0.0017	0.0005	0.0026	-0.0010	0.0054	0.0013	0.0060	-0.0012	-0.0036	-0.0041
Not national language spoken at home	0.0011	0.0004	0.0001	0.0000	0.0007	0.0006	0.0000	0.0002	-0.0007	-0.0001	-0.0001	0.0002
Born in another country	0.0018	0.0010	0.0001	-0.0001	0.0000	0.0001	0.0000	0.0001	-0.0007	-0.0002	0.0001	0.0001
Grade 8	-0.0014	-0.0011							0.0003			
Grade 9	0.0004	-0.0010	0.0067	0.0049			-0.0067	-0.0076	-0.0015	-0.0018		-0.0009
Grade 10	0.0055	0.0066	0.0131	0.0189	-0.0005	0.0014	0.0178	0.0230	0.0011	0.0058	0.0019	0.0053
Grade 11			-0.0065	-0.0078	0.0038	0.0021				-0.0013	-0.0017	
School fixed effects	0.0334	0.0294	0.0131	0.0158	0.0169	0.0222	0.0113	0.0086	-0.0041	0.0028	0.0053	-0.0027
"Residual"	0.0045	-0.0005	0.0057	0.0079	0.0023	0.0029	0.0044	0.0038	-0.0049	0.0022	0.0006	-0.0007
Total	0.0485	0.0359	0.0341	0.0393	0.0244	0.0285	0.0345	0.0300	-0.0126	0.0052	0.0041	-0.0045

Table 8. Oaxaca-type decompositions for change in inequality 2000 and 2006

	Germany				France				Italy				Spain			
	ΔC*η	Δη*C	Total	%	ΔC*η	Δη*C	Total	%	ΔC*η	Δη*C	Total	%	ΔC*η	Δη*C	Total	%
Female	-0.0002	-0.0003	-0.0005	5.85	-0.0001	0.0000	-0.0001	-2.37	-0.0001	-0.0001	-0.0002	-6.21	0.0004	-0.0001	0.0003	4.24
Mother second educ	0.0033	0.0020	0.0053	-68.63	0.0012	0.0005	0.0017	56.93	0.0013	0.0000	0.0013	36.50	0.0003	0.0000	0.0003	4.05
Mother tertiary educ	-0.0028	-0.0045	-0.0072	93.92	0.0011	-0.0049	-0.0038	-128.06	0.0011	-0.0024	-0.0013	-38.97	0.0003	-0.0031	-0.0028	-35.88
Father second educ	-0.0019	-0.0037	-0.0056	72.99	0.0005	0.0007	0.0012	41.02	0.0019	0.0000	0.0018	52.38	0.0003	0.0000	0.0004	4.84
Father tertiary educ	0.0009	0.0050	0.0060	-77.44	0.0007	-0.0020	-0.0012	-41.91	0.0016	-0.0053	-0.0036	-104.95	0.0008	-0.0049	-0.0041	-53.92
Not national language spoken at home	-0.0002	-0.0006	-0.0007	9.40	0.0000	-0.0001	-0.0001	-3.59	0.0000	-0.0001	-0.0001	-2.88	0.0001	0.0001	0.0002	2.75
Born in another country	0.0005	-0.0012	-0.0007	9.73	-0.0003	0.0001	-0.0002	-7.70	0.0000	0.0001	0.0001	3.32	0.0001	0.0000	0.0001	1.70
Grade 8	-0.0005	0.0008	0.0003	-3.78												
Grade 9	-0.0020	0.0006	-0.0015	18.86	-0.0011	-0.0007	-0.0018	-59.64					-0.0003	-0.0006	-0.0009	-11.95
Grade 10	0.0012	-0.0001	0.0011	-13.76	0.0016	0.0041	0.0058	194.36	0.0019	0.0000	0.0019	55.64	0.0119	-0.0066	0.0053	68.46
Grade 11					0.0013	-0.0026	-0.0013	-42.70	0.0004	-0.0021	-0.0017	-48.96				
School fixed effects	-0.0060	0.0020	-0.0041	52.85	0.0032	-0.0004	0.0028	93.67	0.0054	-0.0001	0.0053	154.15	-0.0030	0.0003	-0.0027	-34.75
"Residual"			-0.0049	61.92			0.0022	75.83			0.0006	17.70	0.0000	0.0000	-0.0007	-8.80
Total	-0.0077	0.0000	-0.0077		0.0082	-0.0052	0.0030		0.0134	-0.0099	0.0035		0.0000	0.0000	0.0077	

Figure 1. Concentration Curve

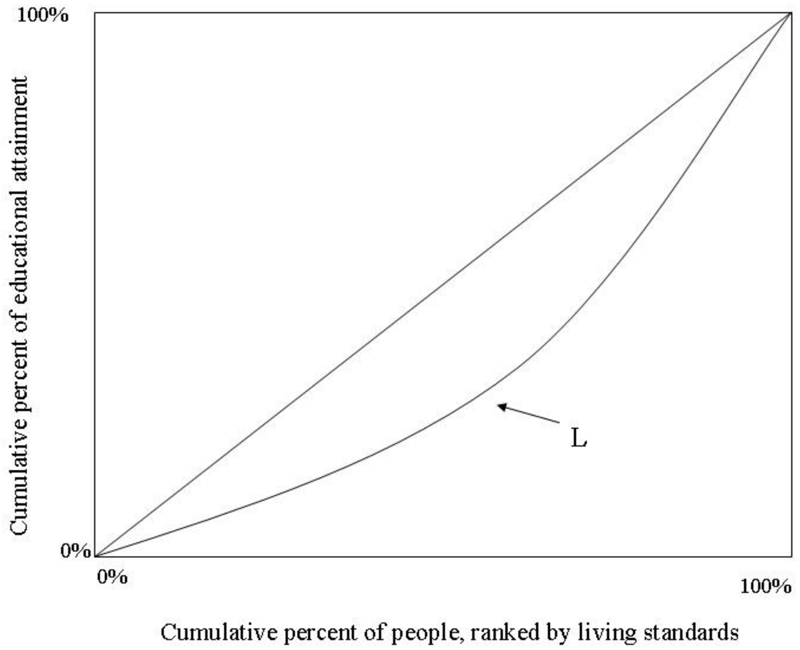

Figure 2. Difference between the concentration curve and the equality line for the reading score, 2000 and 2006