# Equality of Opportunity and Educational Achievement in Portugal

Pedro Carneiro*
University College London, Institute for Fiscal Studies
and Center for Microdata Methods and Practice

May 29, 2007

**Abstract**

Portugal has one of the highest levels of income inequality in Europe, and low wages and unemployment are concentrated among low skill individuals. Education is an important determinant of inequality. However, there are large differences in the educational attainment of different individuals in the population, and the sources of these differences emerge early in the life-cycle when families play a central role in individual development. We estimate that most of the variance of school achievement at age 15 is explained by family characteristics. Observed school inputs explain very little of adolescent performance. Children from highly educated parents benefit of rich cultural environments in the home and become highly educated adults. Education policy needs to be innovative: 1) it needs to explicitly recognize the fundamental long run role of families on child development; 2) it needs to acknowledge the failure of traditional input based policies.

## 1 Introduction

Portugal has one of the highest levels of income inequality in Europe. Using data from the OECD, we estimate that (in 1993) a worker at the 90th percentile of the earnings distribution earns 4.05 times more than a worker in the 10th percentile of the earnings distribution, and 2.47 times more than the worker in the 50th percentile. The ratio between earnings at the 50th and 10th percentile of the earnings distribution was 1.64 in that year. The level of inequality in Portugal is comparable to that observed in North America, which is thought to have the highest level of inequality in the developed world (see OECD, 2005). In Europe, other countries with a comparable (but smaller) level of inequality are Italy, Greece and the United Kingdom.

In this paper we examine the role of education as a source of inequality. Education directly affects an individual's employment and earnings and therefore it contributes to income inequality for a given cross section of individuals. Furthermore, children who are born from better educated parents enjoy a wider range of opportunities than those born from less educated parents. Parental education is not only associated with higher household income, but also with better school and home environments for all children in the household. Therefore, education contributes to intergenerational inequality by naturally creating inequality of opportunity for children born in different families.

We start by studying the relationship between education and wage inequality. We review the literature on the returns to schooling and inequality in Portugal and present some recent results

*This paper was prepared for the 2006 Conference on Economic Development in Portugal organized by the Bank of Portugal. I thank the Bank of Portugal for hospitality and financial and research support, and the Leverhulme Centre for Microdata Methods and Practice for research support. I also thank the hospitality of the Poverty Unit of the World Bank Research Group, and of Georgetown University. Finally I thank Ana Rute Cardoso, Colm Harmon, Pedro Martins, Pedro Telhado Pereira, Pedro Portugal, Vincent O'Sullivan, and two anonynimous referees for useful comments, information on relevant data, and bibliography for Europe and Portugal.

of our own. On average, one additional year of schooling leads to a roughly 1% increase in the employment probability and a 7% increase in net monthly wages. Age alone explains only 1-3% of the total variance of log wages. Education and age together explain 40 to 50% of the total variance of log wages in the portuguese economy in 2004. Loosely speaking, this implies that if everyone in the economy had the same level of education the variance of log wages would be halved. We can look at this result using two opposite but equally informative perspectives: 1) education is an important source of inequality and education policy can dramatically affect wage disparities in Portugal; 2) even if we could achieve equal education for all individuals in society there would still be a lot of inequality left.

Then we examine the sources of education inequality. We reproduce a version of the Coleman report (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld and York, 1966) for Portugal. The Coleman report was a study of U.S. schools which assessed what factors were behind student achievement. This study was extraordinarily influential in the policy and academic circles, and stimulated a substantial amount of research on this topic. Our conclusions are similar to those in the original Coleman report: 1) families play a crucial role in the educational achievement of adolescents (15 years of age) - student achievement is strongly affected by home and school environments, but the most important dimension of school environments is the family background of its students; 2) observed school resources have a weak effect on educational outcomes.

Our results raise important questions and implications. First, it is not desirable to think of education policy focusing only on schools or the classroom in a traditional way. Families have to be brought into the picture more explicitly. This principle has been applied in many developing countries, such as Brazil, Mexico or Colombia, where families are given incentives to keep their children in school, and for providing adequate nutrition and health care. Even in developed countries there are some interventions that consider both child and parent. The question of how to design family-education policy is both difficult and extremely important.

Several recent studies in the US such as Cameron and Heckman (2001) and Carneiro and Heckman (2002, 2003) also emphasize the role of long run family environments in the schooling decisions of adolescents. Furthermore, they document that the differences of achievement in children from different family backgrounds emerge very early and if anything they explain. Skill formation is a cumulative process where "skill begets skill" (Heckman, 2001), and therefore education policy needs to be considered in a life-cycle perspective, recognizing that an individual's achievement and ability to learn at a point in time are a result of the history of past investments (and other events). Once this is recognized, the fundamental importance of early education interventions for children of disadvantaged backgrounds becomes very clear.

Second, our work and that of a whole literature suggests that increases in school resources, at least in the ones we generally observe, will not generate large achievement gains. In the large empirical literature on school quality we can rarely find cost effective interventions. Even when school resources have some impact on student outcomes the interventions are too costly to be justified (see, e.g., Hanushek, 2002, or Carneiro and Heckman, 2003). Of course, there are successful exceptions, such as "the literacy hour", a program evaluated by Machin and McNally (2005), although this is quite an unusual program.[1] As a result, education policy needs to be innovative (such as the one studied by Machin and McNally, 2005): the case for traditional policies of increasing school resources is not strong. Some proposed alternatives are, for example, improvement of incentives in schools through high stakes testing or school choice, or even improvements in incentives to students, although much more research is needed on this topic. This is not to say that schools are ineffective, far from it. In fact, teachers seem to play a major role in student achievement, but it is usually very difficult to identify the observable attributes of a good teacher (for recent evidence, see Hanushek, Kain and Rivkin, 2005). Our point, and that of a whole literature on this topic is that, after a long history of experimenting, traditional input based policies did not prove to be effective.

Third, our regressions show that there is a strong association between an individual's achievement

[1] Krueger and Whitmore (2001) also argue that class size reductions have strong effects on student achievement in elementary school.

in school and the family background of the other students in the same school. Unfortunately, we cannot argue that this is evidence of peer effects. Peer effects are difficult to identify in any dataset (e.g., Manski, 2000), and only recently have economists used research designs suitable for studying this topic (e.g., Sacerdote, 2000), even though this is a topic with substantial policy relevance. In any case, whether peer effects are important for student achievement or not, our work shows that there is a tendency for individuals with similar family backgrounds to attend the same schools, generating a somewhat segregated system. In the dataset we analyze, school effects explain slightly more than 30% of the variance of test scores of 15 year old adolescents in portuguese schools. In comparison with other countries, this is a relatively high number, indicating an above average degree of inequality between schools (OECD, 2002a). Since we cannot accurately describe the empirical importance of peer effects, whether such segregation is harmful or not for the achievement of disadvantaged students is a question we cannot answer. However, segregated educational systems are often cause for concern since they may have adverse effects on social cohesion.

Fourth, even though in our data we have available an unusually rich set of individual, family and school characteristics, a large fraction of achievement inequality remains unexplained. Our individual test score regressions have at most an R-squared of 40% and the school level regressions have at most an R-squared of 53%. Even when we allow the effect of home variables to vary across schools (interacting the school dummies with the home environment variables in the regressions) we get at most an R-squared of 58% in the individual test score regression. This raises an important question that needs to be addressed in future research: why is the explanatory power of these variables so low and what other variables should we be looking at?[2] Furthermore, this reiterates the point that schools can only explain a limited portion of the variation in student performance.

Once we establish the crucial role of family background for student achievement, we study how unequal are the home and school environments of individuals from different family backgrounds. We show that better educated parents provide home environments more conducive to learning than less educated parents. However, there are not many significant differences in the resources of the schools attended by children from low and high education parents. The largest (observable) difference in school environments for these two types of children is in the composition of their peer group. Children of highly educated parents attend school with children who also have highly educated parents, while children whose parents have low levels of education attend school with similar children. In summary, the most significant differences between educational opportunities of children of different family backgrounds are in home environments and peers, not in school resources.

Finally, we document how inequality of opportunity translates into inequality of educational attainment among adults, using a dataset that collects information on individual education and parental education for a sample of portuguese adults in 1999. We show that there is strong intergenerational persistence in educational attainment: differences in adult education generate differences in educational opportunity for their children and, in consequence, persistence of educational inequality from generation to generation. For example, more than 90% of offspring of fathers with an incomplete primary education or less never finish high school, while 0% of offspring of fathers with a university degree complete less than high school.

This study has relevance for education policy in Portugal, but also elsewhere. Within the EU, Portugal is a country with an unusually high level of inequality. However, the mechanisms by which inequality of opportunity translates into inequality of outcomes are likely to be similar across countries. For example, it is interesting to see that the main conclusions of a study done 40 years ago for the US hold today for a country as different as Portugal. Carneiro and Spaltro (2006) replicate this study for several other OECD countries, and relate the observed patterns with observable

[2] It may also be that our model (linear regression) is too restrictive to fit the data. One other concern is measurement error: if test scores are a noisy measure of achievement then it is not surprising that we cannot explain their variance. We attempt to use a less noisy measure of achievement by extractive the first principal component of reading, math and science scores and using it in our analysis instead of the individual test scores. The idea is that scores in these different tests are manifestations of the same underlying ability that we can capture in this first principal component. Our sample size is greatly reduced by this procedure since we only have scores on the three tests for 25% of the sample. Still, when we redo our analysis using this new measure of achievement our basic results do not change significantly.

indicators of education policy across countries.

Our paper proceeds as follows. In section 2 we examine the relationship between inequality in education and inequality in income. In section 3 we study the determinants of inequality in student achievement. Section 4 documents differences in home and school environments between children from different family backgrounds, and section 5 presents estimates of educational mobility in Portugal. Section 6 concludes.

# 2 Education and Labor Market Outcomes

It is striking how much of wage dispersion in Portugal is due to schooling (and age). The first two columns of table 1 present the coefficients of a regression of log monthly wages on age, age squared and years of schooling for males and females aged 25 to 65 (using the Portuguese Labor Force Survey, or LFS,[3] for the fourth quarter of 2004). Schooling accounts for roughly 40% of the total variance of log wages for males and 50% of the variance of log wages for females, as shown by the R-Squared of the wage regressions. Age also plays a role but a much smaller one, which we can ignore for the purpose of the paper (age alone explains only 1-3% of the total variance of log wages). The return to one year of schooling is about 7% for males and 9% for females.[4]

Such large values for R-Squared of wage regressions are unusual, especially in countries with large levels of inequality. A similar regression estimated on US data has an R-squared of only 15%. Across different countries in Europe (largely with lower levels of inequality), a set of studies assembled by Harmon, Walker and Westergaard-Nielsen (2001) show that the R-Squared for similar regressions is below 40% in most cases, and very often it is below 30%.

Table 1 - Returns to Education in Portugal
Inquerito ao Emprego, 4th Quarter of 2004

| | All | Males | Females |
|---|---|---|---|
| Years of Schooling | 0.0789 | 0.0704 | 0.0932 |
| | (0.0010) | (0.0013) | (0.0013) |
| Age | 0.0532 | 0.0585 | 0.0498 |
| | (0.0035) | (0.0043) | (0.0049) |
| $\text{Age}^2$ | -0.0004 | -0.0005 | -0.0004 |
| | (0.0001) | (0.0001) | (0.0001) |
| N | 9729 | 4879 | 4850 |
| $R^2$ | 0.3965 | 0.3864 | 0.5075 |

Note: The regression is estimated by OLS. Individuals reporting now wage or reporting wage in intervals are excluded. Years of Schooling is constructed from the schooling categories in the survey. Standard errors in parenthesis.

Most low wage workers in Portugal have very low education levels. The first column of table 2 shows educational attainment in Portugal for individuals aged 25 to 65. Less than 10% of the working age population was ever enrolled in university and 64% of the population has 6 years of schooling or less. The second column of table 2 presents the education composition of the set of workers earning a net wage smaller than 300 euros per month in 2004 (roughly the net minimum wage), who account for 4.3% of the working population (the margin of people below the legal minimum wage). It shows that 90% of these workers have at most 6 years of schooling, and 97% have at most 12 years of

[3] The name of this dataset is Inquerito ao Emprego. The regression we run is the following:

$$\ln Y_i = \alpha + \beta S_i + \gamma A_i + \theta A_i^2 + \varepsilon_i$$

where $Y_i$ is monthly wage for individual $i$, $S_i$ is years of schooling and $A_i$ is age.

[4] However, this estimate does not correct for the heterogeneity and self-selection into schooling (e.g., Card, 1999, Carneiro, Heckman and Vytlacil, 2005, Carneiro and Lee, 2005). We exclude from the regression individuals reporting wages in brackets. Including them in the regression and running an interval regression instead of a standard linear regression yields very similar results.

schooling. Furthermore, among those workers earning less than 600 euros per month (roughly the median wage), 74% have 6 years of schooling or less and 98% have 12 years of schooling or less. There is a clear link between poverty and lack of skills.

Table 2 - Educational Composition of Individuals
Earning Less than 300 and 600 Euros per Month
Inquerito ao Emprego, 4th Quarter of 2004

| | Population | | <300 | | <600 | |
|---|---|---|---|---|---|---|
| | % | Cumulative | % | Cumulative | % | Cumulative |
| No Schooling | 0.0824 | 0.0824 | 0.1680 | 0.1680 | 0.0596 | 0.0596 |
| 4th Grade | 0.3821 | 0.4645 | 0.5186 | 0.6866 | 0.4260 | 0.4856 |
| 6th Grade | 0.1763 | 0.6408 | 0.2084 | 0.8950 | 0.2557 | 0.7413 |
| 9th Grade | 0.1327 | 0.7735 | 0.0436 | 0.9386 | 0.1402 | 0.8815 |
| 12th Grade | 0.1157 | 0.8892 | 0.0355 | 0.9741 | 0.0975 | 0.9790 |
| Some Post-Secondary | 0.0237 | 0.9129 | 0.0017 | 0.9758 | 0.0064 | 0.9854 |
| University | 0.0797 | 0.9926 | 0.0242 | 1.0000 | 0.0143 | 0.9997 |
| Master | 0.0033 | 0.9959 | 0.0000 | 1.0000 | 0.0003 | 1.0000 |
| Doctorate | 0.0041 | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 |

Most of the portuguese labor force is relatively low skilled when compared to the rest of Europe (e.g., OECD, 2002a), although there have been some improvements over time. Among those individuals in the labor market with 55 or more years of age in 2004, more than 85% have less than a high school education, and little more than 5% have a college education. Among those who are 25-35 in 2004, 32% have a high school education or above, and only 12% have ever enrolled in college. At the same time, the returns to education are much smaller for the younger cohorts than for the older cohorts, coming down from 11% for the older cohorts, to 6% for the younger cohorts (although these differences may also hide age effects in the return to schooling). These numbers are computed from the Inquerito ao Emprego, 4th Quarter of 2004. Overall, education is an important determinant of inequality across all cohorts, and education levels are low across all cohorts.

Finally, education is strongly associated with employment. Table 3 presents the average derivatives of a regression (probit) of employment on years of schooling, age and age squared, using individuals aged 25 to 65 from the LFS.[5] An increase in one year of schooling is associated with an increase in the probability of being employed of 1% for males and 2.7% for females. In 2004, more than 75% of the nonemployed (34% of the population aged 25 to 65 in this year) have 6 years of education or less and 94% of the nonemployed have 12 or less years of schooling.[6] Education not only explains a large fraction of the variance of wages, but an overwhelming proportion of low wage and nonemployed individuals have very low levels of education.[7] This suggests that investing in education is an important mechanism for escaping poverty.

[5] The regression model is the following:

$$\begin{aligned} E^* &= \alpha + \beta S + \gamma A + \theta A^2 + \varepsilon \\ E &= 1 \text{ if } E^* > 0 \end{aligned}$$

where $E$ is employment, $S$ is years of schooling and $A$ is age.

[6] Results available on request.

[7] Obviously, this does not mean that all low skilled individuals have low wages are are not employed. For example, 91% of those individuals with 6 years of schooling earn more than 300 euros per month, and 33% of them earn more than 600 euros per month. Similarly, 75% of the individuals in this group are employed in 2004. What this says is that almost no highly educated individuals are either low wage or nonemployed. Among those individuals with 12 years of schooling, 98.6% have wages higher than 300 euros per month, and 65 % have wages above 600 euros per month. Among those with a university degree, these percentages are 98.8% and 93.5% respectively. The fraction of nonemployed individuals is 22% among those with 12 years of schooling and 15% among those with a university degree.

Table 3 - Education and Employment
Inquerito ao Emprego, 4th Quarter of 2004

| | All | Males | Females |
|---|---|---|---|
| Years of Schooling | 0.0195 | 0.0105 | 0.0273 |
| | (0.0008) | (0.0009) | (0.0011) |
| Age | 0.0597 | 0.0625 | 0.0565 |
| | (0.0022) | (0.0028) | (0.0033) |
| $Age^2$ | -0.0008 | -0.0008 | -0.0007 |
| | (0.0001) | (0.0001) | (0.0001) |
| N | 28643 | 12713 | 13770 |

Note: This table presents average marginal derivatives of a probit of employment on years of schooling and age. Years of Schooling is constructed from the schooling variable in the survey. Standard errors are in parenthesis.

Furthermore, to paraphrase Cardoso (1998), inequality in Portugal is “high and rising”. Several researchers such as Cardoso (1998), Machado and Mata (2001), Pereira and Martins (2000) and Hartog, Pereira and Vieira (2001), document an important increase in inequality during the 1980s and early 1990s (Parente and d’Uva, 2002, suggest that there was some stabilization after that). This research also shows that most of the recent increase in inequality has been at the top of the income distribution. Over the same period there has been a substantial increase in the returns to schooling, which is also documented in the papers cited above. The increase in the returns to schooling generates an increase in inequality between individuals with different levels of education. However, there has also been an increase in inequality within education groups, which suggests that there was an increase in the return to (observed and unobserved) skill. The most standard explanation for the increase in inequality in the western world is skill biased technical change.

This increase in inequality occurs at a time of rapid economic growth. The period comprising the 1980s and the early 1990s was a period of high growth for Portugal. Rich and poor individuals have benefited very differently from the overall improvement in the performance of the portuguese economy, both in absolute and in relative terms. Machado and Mata (2001) document that between 1982 and 1994 wages increased by 20% for the 10th percentile of the wage distribution and by 52% at the top.

In other countries, such as for example the UK, wages at the bottom of the wage distribution have stagnated in recent years while wages at the top continue to increase, and in the US low wage individuals have experienced a decline in their real wages in the last 30 years. This is especially worrisome once we think about the problem of poverty and of poverty alleviation mechanisms. Rebecca Blank (1996) argues that the most effective poverty alleviation program available to governments is economic growth. In fact, in the years after the second world war, the poor were doing relatively well across the industrialized world. Even though most of them were unskilled there were good jobs available for them, mainly in manufacturing. However, modern economic growth is driven by technological growth, and access to the benefits of economic growth is restricted to those individuals who have invested in skills. Due to the overall growth of the economy, as the demand for services increases, jobs are still available for the poor in these countries but these are mostly low value added jobs.

Surprisingly, the relative performance of low wage and low skilled individuals in Portugal has not been as poor as in countries with similar inequality levels, such as the US and the UK. One reason has been that the low skilled jobs in manufacturing are still available in this country. The other is probably the social safety net which is stronger in continental Europe than in anglo-saxon countries. However, in the future as the economy grows, reconverts and adapts to the competitive pressures of the modern world, the sizeable low skilled population that exists in Portugal is likely to not only prevent the development of the economy, but also to see their situation deteriorate, not only in relative terms, but possibly in absolute terms as well.

Education is not only important for labor market outcomes, but also for many other dimensions of an individual’s life, and therefore the focus on labor market outcomes is too narrow. For example, education affects criminal behavior (e.g., Lochner and Moretti, 2004), health (e.g., Grossman, 2005, Smith, 2005), family formation and child development (e.g., the studies summarized in Carneiro and Heckman, 2003), among many other things. Of particular interest in this paper will be the role of education on parental investments and child development, an important source of intergenerational mobility. Inequality in education generates inequality of opportunity among children growing up in different families, not only because family resources are smaller in families with lower levels of education, but probably more importantly, because education affects parental behavior (e.g., Carneiro, Meghir and Parey, 2005, among others).

# 3 Sources of Inequality in Educational Achievement

The overall picture described above illustrates the crucial (but certainly not exclusive) role of skill for individual outcomes in a modern economy (and for the aggregate success of the economy), and the importance of human capital policy in the modern environment. There are large skill disparities in the portuguese society that dramatically affect an individual's life chances. However, these skill disparities are likely to emerge well before individuals enter the labor market. Using US data, Carneiro and Heckman (2003) document that skill differences among children from different socio-economic groups emerge very early in life, often as early as ages 1 and 2. These early skill gaps are substantial and if anything they tend to expand as children grow older.

Carneiro and Heckman (2003) and Cunha, Heckman, Lochner and Masterov (2005) suggest that complementarity between human capital investments taking place at different periods in time is a key feature of the technology of skill formation. If early and late investments are complementary in the production function of human capital then the productivity of late investments increases with the amount of early investments. Intuitively, the more I invest early the better equiped I am to learn in the future, because learning builds on accumulated skill. This also implies that late investments are unlikely to be productive if they are not preceded by early investments. Therefore, compensatory education programs targeted to young disadvantaged adults may not be very effective because they cannot remedy early neglect. Finally, complementarity also implies that the productivity of early investments increases with late investments. If early investments are not followed up they do not amount to anything. A consequence of complementarity is that childhood skill disparities translate into even larger adult skill disparities. Furthermore, it may also be difficult and costly to design compensatory programs in adulthood that correct such skill gaps.

Carneiro and Heckman (2002, 2003) argue that skill disparities are mostly a consequence of home environments, and present a large body of supporting evidence. In fact, a similar claim was made in what became known as the Coleman Report. Coleman and his colleagues investigated the determinants of school performance among children in the US and concluded that the family background of ones peers was the most important factor affecting individual school performance, while school resources only played a limited role. Cognitive achievement in the adolescent years is good predictor of final educational attainment, and it also affects wages in several developed countries (e.g., Carneiro and Heckman, 2002, Heckman and Vytlacil, 2001, Blundell, Dearden and Sianesi, 2005, Currie and Thomas, 2003), which makes it an interesting and relevant adolescent outcome to look at. Although some of the conclusions of this report have been questioned in the last 40 years, the core of Coleman's argument is still thought to be correct. Not much is known for Portugal, but the patterns are likely to be similar (as shown in the remaining of this section).

In this section I examine the sources of inequality in educational achievement in Portugal using a sample of 15 year old individuals surveyed by the Programme for International Student Assessment (PISA). I focus on the 2000 wave of this survey. The PISA is an international assessment of literacy in reading, mathematics and science for 15 year old adolescent in a large set of (mostly developed) countries.[8] Its findings have received wide attention in the media and in the academic world, and were summarized in OECD (2001, 2003). Portuguese students rank relatively poorly relatively to

[8] For a detailed description see OECD (2002b).

their counterparts in other OECD countries (OECD, 2001, 2003). The PISA collects rich information on cognitive skills, family environments and school environments which allows us to do an analysis of these different variables on student achievement. We would like to start by examining skill inequality in a younger sample of individuals but no comparable data exists (in Portugal) for the early childhood years. Even though the spirit of our work is that of the original Coleman report, our methodology is slightly different.

We start by estimating the following model:

$$T_{ij} = \alpha + F_i\beta + S_j\gamma + \varepsilon_{ij} \tag{1}$$

where $T_{ij}$ is a test score (in reading, mathematics or science) for individual $i$ in school $j$, $F_i$ is a vector of family and home characteristics, $S_j$ is a vector of school dummies and $\varepsilon_{ij}$ is the error term (orthogonal to $F_i$ and $S_j$). Let $\Phi_i = F_i\beta$ represent family effects and $\Psi_j = S_j\gamma$ represent school effects. The variables in $F_i$ are indices of parental socio-economic background, cultural communication with parents, social communication with parents, home educational resources, activities related to "classical" culture, possessions related to "classical" culture in the family home and time spent on homework. These indices were constructed by the PISA staff from answers to the student questionnaire. A description of the construction of these indices is provided in the manual for the PISA 2000 database (OECD, 2002b).[9]

Then we can decompose test scores in family, school and orthogonal residual effects:

$$T_{ij} = \alpha + \Phi_i + \Psi_j + \varepsilon_{ij}$$

(this is a restrictive model - it does not allow for interactions between $\Phi_i$ and $\Psi_j$, although these could be included). Finally:

$$\begin{aligned} Var(T_{ij}) &= Var(\Phi_i) + Var(\Psi_j) + Var(\varepsilon_{ij}) + 2 * Cov(\Phi_i, \Psi_j) \\ &\text{or} \\ 1 &= \frac{Var(\Phi_i)}{Var(T_{ij})} + \frac{Var(\Psi_j)}{Var(T_{ij})} + \frac{Var(\varepsilon_{ij})}{Var(T_{ij})} + \frac{2 * Cov(\Phi_i, \Psi_j)}{Var(T_{ij})}. \end{aligned} \tag{2}$$

This model allows us to assess the relative contribution of family background/home environment and school characteristics to inequality in educational achievement. The R-Squared of regression (1) is given by $\frac{Var(\Phi_i)}{Var(T_{ij})} + \frac{Var(\Psi_j)}{Var(T_{ij})} + \frac{2*Cov(\Phi_i,\Psi_j)}{Var(T_{ij})}$.

The Coleman Report argued that the main determinant of a child's school success was the family background of his peers. Therefore, it is instructive to separate school effects ($\Psi_j$) into two components: (observed) school quality effects ($\Psi_j^S$) and family background effects ($\Psi_j^F$). In order to do that I estimate the following model (using one observation per school, independently of school size):

$$\Psi_j = \theta + S_j^S\sigma + S_j^F\eta + \upsilon_j \tag{3}$$

where $\Psi_j$ is estimated from equation (1), the variables in $S_j^S$ are school size, hours of schooling per year, number of computers per student per school, student-teaching staff ratio and the proportion

[9] The PISA Index of Socio-Economic Status was derived from students' responses on parental occupation. The Index of Cultural Communication with parents was derived from students' reports on the frequency with which their parents engaged with them in the following activities: discussing political or social issues; discussing books, films or television programmes; and listening to classical music. The index of Social Communication with parents was derived from students' reports on the frequency with which their parents engaged with them in the following activities: discussing how well they are doing at school, eating with them around a table, spending time simply talking to them. The index of Home Education Resources was derived from students' reports on: i) the availability, in their home, of a dictionary, a quiet place to study, a desk for study and textbooks; and ii) the number of calculators at home. The index of Activities Related with Classical Culture was derived from the students' reports on how often they had participated in the following activities during the preceding year: visited a museum or art gallery; attended an opera, ballet or classical symphony concert; and watched live theatre. The index of Possessions Related to Classical Culture was derived from the students' reports on the availability of the following items in the home: classical literature; and books of poetry and works of art. The index of Time Spent on Homework was derived from the students' reports on the amount of time they devote to homework per week in reading, mathematics and science.

of teachers with a university degree in pedagogy. The variables in $S_j^F$ are the average indices of social communication, cultural communication, home education resources, cultural possessions of the family, and parental socio-economic background in the school (described above). Again, we can write:

$$
\begin{aligned}
\Psi_j &= \theta + \Psi_j^S + \Psi_j^F + \upsilon_{ij} \\
Var\left(\Psi_j\right) &= Var\left(\Psi_j^S\right) + Var\left(\Psi_j^F\right) + Var\left(\upsilon_{ij}\right) + 2 * Cov\left(\Psi_j^S, \Psi_j^F\right) \\
&\text{or} \\
1 &= \frac{Var\left(\Psi_j^S\right)}{Var\left(\Psi_j\right)} + \frac{Var\left(\Psi_j^F\right)}{Var\left(\Psi_j\right)} + \frac{Var\left(\upsilon_{ij}\right)}{Var\left(\Psi_j\right)} + \frac{2 * Cov\left(\Psi_j^S, \Psi_j^F\right)}{Var\left(\Psi_j\right)} \qquad (4)
\end{aligned}
$$

The R-Squared of this equation is given by $\frac{Var\left(\Psi_j^S\right)}{Var\left(\Psi_j\right)} + \frac{Var\left(\Psi_j^F\right)}{Var\left(\Psi_j\right)} + \frac{2*Cov\left(\Psi_j^S, \Psi_j^F\right)}{Var\left(\Psi_j\right)}$.

Since families choose schools, and since school resources are not randomly allocated in the population, the interpretation of some of the results of this model may be unclear. Below we make these concerns explicit and we comment on them in more detail. We first present our main results.

Table 4a shows the coefficients estimated from equation (1) and table 4b presents the estimates from equation (3) for three sets of tests: reading, math and science. All home variables appear with the expected sign and are statistically important, especially in the reading equation. The effect of observed school resources is often quite weak. Table 5 presents the variance decompositions corresponding to equations (2) (panel A) and (4) (panel B). Panel A shows that we can only explain about 45% of the variance in reading scores (less for the other scores) using school and observable family variables. School effects alone account for 22% of the total variance of these scores, which is about than 50% of the explained variance of the model. As shown in panel B, most of the variance in school effects is due to the family background of the students in the school (for reading, 60% of the total variance or almost 90% of the explained variance; similar, or even more extreme, results hold for math and science). Notice also that most inequality is within schools, not between schools, which means that schools by themselves cannot explain a large portion of inequality.

In summary, the main conclusion of the Coleman report seems to hold for Portugal: the average family characteristics of the students in an individual's school are the main observable determinant of inequality in educational achievement.

Table 4a - Regression of Test Scores on School Dummies and Family Characteristics
Data from PISA 2000

| | Reading | Math | Science |
|---|---|---|---|
| Parental Cultural Communication | 13.820 | 9.185 | 11.951 |
| | (1.286) | (1.692) | (1.726) |
| Parental Social Communication | 4.168 | 3.411 | 3.437 |
| | (1.273) | (1.676) | (1.727) |
| Home Educational Resources | 10.552 | 12.528 | 6.498 |
| | (1.381) | (1.834) | (1.866) |
| Cultural Possessions of the Family | 6.372 | 3.205 | 2.432 |
| | (1.304) | (1.732) | (1.752) |
| Socio Economic Index of Parents | 0.801 | 0.882 | 0.753 |
| | (0.082) | (0.107) | (0.109) |
| N | 4379 | 2421 | 2430 |
| $R^2$ | 0.43 | 0.41 | 0.39 |

Note: Standard errors in parenthesis.

Table 4b - Regression of School Dummies on School and Family Characteristics
Data from PISA 2000

| | Reading | Math | Science |
|---|---|---|---|
| Number of Students in the School | 0.012 | 0.004 | -0.002 |
| | (0.005) | (0.007) | (0.007) |
| School Hours per Year | -0.003 | -0.018 | -0.004 |
| | (0.021) | (0.031) | (0.028) |
| Number of Computers / School Size | 8.564 | 5.978 | 7.144 |
| | (7.152) | (14.290) | (7.524) |
| School Size / Number of Teachers | -1.812 | 0.687 | -0.926 |
| | (0.872) | (1.238) | (1.096) |
| Proportion of Teachers with Degree in Pedagogy | 12.398 | 8.028 | 21.419 |
| | (8.610) | (12.790) | (19.154) |
| Average Parental Cutural Communication in the School | 47.951 | 51.094 | 56.451 |
| | (14.804) | (21.879) | (16.364) |
| Average Parental Social Communication in the School | 50.623 | 43.466 | 39.718 |
| | (12.098) | (20.923) | (18.126) |
| Average Home Education Resources in the School | 6.999 | 5.498 | 23.367 |
| | (14.905) | (20.974) | (19.420) |
| Average Socio Economic Index in the School | 0.429 | 0.315 | 0.179 |
| | (0.575) | (0.782) | (0.768) |
| N | 132 | 70 | 75 |
| $R^2$ | 0.53 | 0.60 | 0.65 |

Note: Standard errors in parenthesis.

Table 5 - Variance Decomposition of Test Scores on Family and School Variables
Percentage of Variance Explained by Different Components
Data from PISA 2000

| Panel A - Regress Test Scores on Family Variables and School Dummies | | | |
|---|---|---|---|
| | Reading | Math | Science |
| V(F) | 0.1030 | 0.0954 | 0.0710 |
| V(S) | 0.2291 | 0.2180 | 0.2316 |
| COV(F,S) | 0.1132 | 0.0982 | 0.0880 |
| V(Residual) | 0.5547 | 0.5884 | 0.6094 |
| Panel B - Regress School Dummies on School and Family Characteristics | | | |
| | Reading | Math | Science |
| V(F) | 0.6023 | 0.5870 | 0.6495 |
| V(S) | 0.0286 | 0.0164 | 0.0253 |
| COV(F,S) | 0.0592 | 0.0006 | -0.0207 |
| V(Residual) | 0.3099 | 0.3960 | 0.3459 |

This may be due to several reasons. Variance decompositions depend not only on the coefficients in the regressor of interest but also on the variance of the regressors themselves. For example, a measure of school quality such as the student-teacher ratio may be an important determinant of student achievement (large coefficient). However, if the variance of the student-teacher ratio is small then this variable will not explain very much of the total variance in achievement, even if its coefficient is large. As we will see in detail in section 4, school quality variables do not differ widely between individuals from different family backgrounds and therefore cannot explain differences in achievement due to differences in family background. However, school quality variables do vary within family background groups and their variance is quite considerable in some cases (table A1 in the appendix shows mean and dispersion measures for five different school quality variables). Therefore, the reason for the low contribution of this type of variables to inequality in achievement does not seem to be a low level of variability of these variables in the population, but a relatively unimportant association between the variation in these variables and school performance.

Unfortunately, as we emphasized above, our results cannot be given a straightforward causal interpretation. There are unobserved family, school and student factors which affect test scores, and which are likely to be correlated with our measures of family background and school resources.

We would expect observed family factors to be positively correlated with unobserved family factors and unobserved student ability, in which case we should interpret our estimates of the effect of family background as being inclusive of unobserved family variables and unobserved student ability.

Unobserved school resources may be positively or negatively related to observed family background. On one hand, children from richer families tend to sort into better schools, in which case this correlation would be positive. On the other hand, if governments seek to compensate inequalities in family background and provide extra support to failing schools, there may be a negative correlation between unobserved school resources and family background. The same arguments could apply to the correlations between observed school resources and unobserved family factors and student ability.

Finally, we would expect observed and unobserved school resources to be positively correlated, although the compensatory policy argument could again reverse this correlation. If compensatory policies are very strong we may end up underestimating the importance of school resources.

Unfortunately, our data is not rich enough to address this problem adequately. However, there are still suggestive checks to the data that we can do. Even though we do not observe school quality directly, we do have good measures of family background. If students from high quality families sort into high quality schools, then the average family background of the students in the school will be a proxy for school quality. In fact, we observe that individuals with richer family backgrounds enrol in schools with larged fixed effects as measured in our first stage regressions: the correlation between $\Phi_i$ and $\Psi_j$ is larger than 0.3 and statistically significant for all three test scores, suggesting that this is not a bad assumption. Therefore, we computed the correlation between observed school resources and observed family background at the school level, or $\Psi_j^S$ and $\Psi_j^F$: for reading test scores we obtained a correlation of 0.23, significant at the 1% level; for math scores the correlation is zero and insignificant; for science scores the correlation is -0.08, and insignificant. In summary, the correlation between measured school resources and measured family background at the school level, which for this purpose we assume to be a proxy for underlying school quality, is either zero or positive, suggesting that the compensatory policy argument may not be very strong. However, we recognize that these arguments are merely suggestive, and that given our data we cannot definitely solve this question.

One thing we cannot really do is to put a finger on what are the attributes of schools that are most important or what are the attributes of families that are most important. Since our observed family background and school resource variables may not vary exogenously and may be correlated with several other unobserved family background or school resource variables, we do not know if the effect we are capturing is that of the variable we observe or of some other variable correlated with it. Nevertheless, even though this would certainly make this exercise more complete and useful, this is not the major point of this paper. Due to the nature of the data, we have limited ourselves to be mostly interested in understanding the relative role of school and family factors in explaining inequality in student achievement, whatever factors these may be.

The results from this section are consistent with a large literature in the economics of education which shows very strong family background effects on educational achievement, and less strong effects of measurable school resources. There are a few school programs for which significant effects have been found, but they are the exception rather than the rule (although not many evaluations exist for this type of programs), and their effects are generally quantitatively small. One important exception has to do with teacher quality. Teacher quality systematically shows up in studies of achievement as an important determinant of educational success. However, teacher quality is a "black box", generally measured as a teacher fixed effect in a regression of individual test scores on teacher effects and other school and family variables. These teacher effects cannot be explained by observable measures of teacher quality, such as qualification or experience. Researchers agree that teachers matter very much, but it is not clear what makes a good teacher (e.g., Carneiro and Heckman, 2003, Hanushek, 2001, Hanushek, Kain and Rivkin, 2005).

Another potentially useful set of alternative school policies involves a better use of incentives in schools. The increase in school choice and school competition may be one way to do it, but unfortunately the evidence on this is still controversial and relatively scarce. Providing direct incentives for students to exert effort, or even to stay in school, or for teachers to improve their performance is also a possibility. Therefore, our conclusion and that of a large literature on the topic is not that schools are unimportant. On the contrary, schools are extremely important, but innovative education policy is needed so that the resources which are poured into schools are better used.

Unfortunately, we cannot explain a large fraction of the variance of test scores neither with family background variables nor with school variables. For example, for reading scores we can get an R-squared of 40% for equation (1) and an R-squared of 53% for equation (3). Given that we measure school effects in a flexible way (school dummies), we can conclude that most of the variance of test scores is within schools and it is likely to be caused by unobserved family background variables and their correlation with school variables, which if true will only strengthen our basic result. However, between school variance (equation (3)) may be explained by several unobservable factors, and which can be either school or family factors.[10]

Finally, the almost exclusive emphasis that is usually given to cognitive test scores as a diagnostic of educational success and as a guide to educational policy may be unbalanced because many other skills matter. Academic achievement is relatively easy to measure through test scores, but many other unmeasured skills can be equally or more important for future success in several dimensions of life. For example, Heckman, Hsee and Rubinstein (2000), Heckman Sixtrud and Urzua (2005) and Carneiro, Crawford and Goodman (2005) show that several measures of "noncognitive skills" (which aim to capture traits such as sociability, patience or discipline) are strong determinants of educational success, labor market outcomes and a variety anti-social behaviors (such as criminal activity, teenage pregnancy or drug use). Therefore, we replicated the analysis of table 5 for five measures of noncognitive ability: sense of belonging, engagement in reading, effort and perseverance, instrumental motivation and interest in reading.[11] Table A2 shows our results. The explained portion of the variance of these variables is even lower than when we used test scores. Even though the role of an individual's family is relatively more important than before, surprisingly, the role of the peer's families seems to be relatively weaker than before.

[10] There are additional school and family variables in the PISA dataset which we have not used. Their inclusion does not change the conclusions of this paper. More work needs to be done although it may be very difficult to explain a larger fraction of inequality in achievement than we already do. Our plan is to briefly examine other databases, namely the TIMSS95 (Trends in Mathematics and Science Study of 1995) and the 2003 wave of the PISA.

[11] These are five indices described in detail in OECD (2002b). The belonging index was derived from students' reports on their level of agreement with the following statements concerning their school: I feel like an outsider (or left out of things); I make friends easily; I feel like I belong; I feel awkward and out of place; other students seem to like me; and, I feel lonely. The engagement in reading index was derived from students' level of agreement with the following statements: I read only if I have to; reading is one of my favourite hobbies; I like talking about books with other people; I find it hard to finish books; I feel happy if I receive a book as a present; for me, reading is a waste of time; I enjoy going to a bookstore or a library; I read only to get information that I need; and, I cannot sit still and read for more than a few minutes. The effort and perseverance index was derived from the frequency with which students used the following strategies when studying: I work as hard as possible; I keep working even if the material is difficult; I try to do my best to acquire the knowledge and skills taught; and, I put forth my best effort. The index of instrumental motivation was derived from the frequency with which students study for the following reasons: to increase my job opportunities; to ensure that my future will be financially secure; and, to get a good job. The index of interest in reading was derived from students' level of agreement with the following statements: because reading is fun, I wouldn't want to give it up; I read in my spare time; and, when I read, I sometimes get totally absorbed.

# 4 Equality of Opportunity: Families and Schools

In the previous section we showed how schools and families strongly influence the academic achievement of 15 year old individuals in Portugal. Different individuals are born in different families, and therefore experience unequal environments at home and in the school throughout their childhood and adolescence. In this section we document how inequality in paternal education is associated with differences in home and school environments. In short, we show how unequal paternal education translates into different education opportunities, which give rise to inequality in school achievement.

Table 6 shows that there are large differences in the level of schooling of fathers of different individuals. These are just a consequence of inequality in education in the previous generation, which generates inequality of opportunities for the current generation. Table 7 shows that highly educated fathers show more interest for their children's progress in school. Table 8 shows that children of highly educated fathers are more likely to be taken to museums, and table 9 shows that they spend longer hours studying. Table 10 and 11 refer to school environments. Table 10 shows that there are no strong differences in standard school quality variables between children of fathers with high and low levels of education. Therefore, differences in school quality variables cannot explain how unequal parental background translates into unequal test scores. However, table 11 shows that the education of the father of the child and the education of the fathers of students in the same school are strongly correlated.

Table 6 - Inequality in Paternal Schooling
Data from PISA 2000

| | % |
|---|---|
| Primary | 2.25 |
| Lower Secondary | 37.65 |
| Upper Secondary | 39.90 |
| Lower Tertiary | 13.19 |
| Upper Tertiary | 17.01 |

Table 7 - Paternal Schooling and Percentage at Each Level of Parental Interest
"In general, how often do your parents discuss how well you are doing at school?"
Data from PISA 2000

| % | Never or hardly ever | A few times a year | About once a month | Several times a month | Several times a week |
|---|---|---|---|---|---|
| Primary | 0.0306 | 0.1122 | 0.0714 | 0.1836 | 0.6020 |
| Lower Secondary | 0.0263 | 0.0422 | 0.0759 | 0.2181 | 0.6372 |
| Upper Secondary | 0.0177 | 0.0431 | 0.0501 | 0.2004 | 0.6885 |
| Lower Tertiary | 0.0122 | 0.0228 | 0.0385 | 0.1578 | 0.7684 |
| Upper Tertiary | 0.0094 | 0.0296 | 0.0336 | 0.1549 | 0.7722 |
| Total | 0.0191 | 0.0394 | 0.0560 | 0.1933 | 0.6920 |

Table 8 - Paternal Schooling and Percentage at Each Level of of Museum Visits
"During the past year, how often have you visited a museum or art gallery?"
Data from PISA 2000

| % | Never or hardly ever | Once or Twice a year | About 3 or 4 times a year | More than 4 times a year |
|---|---|---|---|---|
| Primary | 0.2755 | 0.2959 | 0.1122 | 0.3163 |
| Lower Secondary | 0.1579 | 0.2029 | 0.1433 | 0.4957 |
| Upper Secondary | 0.1013 | 0.1459 | 0.1374 | 0.6152 |
| Lower Tertiary | 0.0743 | 0.1072 | 0.1159 | 0.7024 |
| Upper Tertiary | 0.0563 | 0.1006 | 0.0993 | 0.7436 |
| Total | 0.1153 | 0.1579 | 0.1297 | 0.5969 |

Table 9 - Paternal Schooling and Time on Home Work
"On average, how much time do you spend each week on homework or studying portuguese?"
Data from PISA 2000

| % | No time | Less than 1 hour a week | Between 1 and 3 hours a week | 3 or more hours a week |
|---|---|---|---|---|
| Primary | 0.0938 | 0.4583 | 0.2604 | 0.1875 |
| Lower Secondary | 0.0446 | 0.3534 | 0.4590 | 0.1428 |
| Upper Secondary | 0.0382 | 0.3608 | 0.4778 | 0.1230 |
| Lower Tertiary | 0.0505 | 0.3292 | 0.4756 | 0.1445 |
| Upper Tertiary | 0.0577 | 0.3463 | 0.4174 | 0.1785 |
| Total | 0.0468 | 0.3535 | 0.4554 | 0.1442 |

Table 10 - Paternal Schooling and School Resources
Data from PISA 2000

| % | Hours of Schooling per Year | Number of Computers per Student in the School | Student-teaching Staff Ratio | Proportion of Teacher Degree in Pedagogy |
|---|---|---|---|---|
| Primary | 893 | 0.1170 | 8.752 | 0.5032 |
| Lower Secondary | 899 | 0.0886 | 9.019 | 0.4439 |
| Upper Secondary | 896 | 0.0799 | 8.961 | 0.4346 |
| Lower Tertiary | 896 | 0.0680 | 8.666 | 0.4185 |
| Upper Tertiary | 889 | 0.0872 | 8.689 | 0.3975 |
| Total | 896 | 0.0837 | 8.893 | 0.4311 |

Table 11 - Paternal Schooling and Average Maternal Schooling in the School
Data from PISA 2000

| % | Proportion Primary | Proportion Lower Secondary | Proportion Upper Secondary | Proportion Lower Tertiary | Proportion Upper Tertiary |
|---|---|---|---|---|---|
| Primary | 0.0753 | 0.4314 | 0.2868 | 0.0928 | 0.1081 |
| Lower Secondary | 0.0257 | 0.4451 | 0.2880 | 0.1097 | 0.1312 |
| Upper Secondary | 0.0215 | 0.3626 | 0.3270 | 0.1289 | 0.1598 |
| Lower Tertiary | 0.0167 | 0.3131 | 0.2922 | 0.1799 | 0.1978 |
| Upper Tertiary | 0.0142 | 0.2905 | 0.2809 | 0.1535 | 0.2606 |
| Total | 0.0224 | 0.3764 | 0.2990 | 0.1319 | 0.1700 |

In summary, families with better educated fathers provide better home and school environments for their children, which then translate into better test scores. We have examined only a few dimensions of such environments. It is interesting that children from better educated fathers do not attend schools with more resources. The most important (and maybe only significant) difference between schools attended by children from low and high educated fathers is in the family background of the students in the school. Better educated fathers enrol their children in schools with similar peers (children from highly educated parents). This type of segregation is prevalent in a variety of situations. Individuals tend to associate with others who are alike, for a variety of reasons. Unfortunately, this may result in strong inequality in outcomes if peer effects are important.

# 5 Intergenerational Mobility in Education in Portugal

Inequality in school achievement is likely to generate inequality in school attainment and labor market outcomes. In this section we examine the relationship between parental education and an individual's final educational attainment and labor market outcomes. We use the Social Inequalities II dataset collected by the Instituto Nacional de Estatistica and the Instituto de Ciencias Sociais of the University of Lisbon (as part of a European network).[12]

[12] A detailed description of this dataset can be found in ISSP (1999).

Table 12 is a transition matrix, which gives the probability that an individual completes a certain educational degree given the educational attainment of his father. It is obvious that there is a large increase in educational attainment from the parents to the offspring generation in this dataset, but there is also substantial intergenerational persistence in educational status. Notice that more than 90% of offspring of fathers with an incomplete primary education or less never finish high school, while 0% of offspring of fathers with a university degree complete less than high school. This disparity is striking, especially when we consider that 50% of the individuals in this sample have fathers with less than a complete primary education, while only 1.9% have fathers with a complete university education.

Table 12 - Paternal Schooling and Child's Schooling (transition matrix)
Data from Social Inequalities II

| Father/Child | None | Inc. Prim. | Prim. | Inc. Sec. | Sec. | Inc. Univ. | Univ. |
|---|---|---|---|---|---|---|---|
| None | 18.80 | 20.81 | 45.69 | 13.45 | 0.76 | 0.00 | 0.51 |
| Incomplete Primary | 5.85 | 18.13 | 49.71 | 21.64 | 2.34 | 0.58 | 1.75 |
| Primary | 1.74 | 4.48 | 24.13 | 44.78 | 13.93 | 3.73 | 7.21 |
| Incomplete Secondary | 0.00 | 0.00 | 9.18 | 34.69 | 18.37 | 13.27 | 24.49 |
| Secondary | 0.00 | 0.00 | 0.00 | 20.00 | 40.00 | 20.00 | 20.00 |
| Incomplete University | 0.00 | 0.00 | 0.00 | 66.67 | 33.33 | 0.00 | 0.00 |
| University | 0.00 | 0.00 | 0.00 | 0.00 | 14.29 | 23.81 | 61.90 |

Table 13 reports employment rates for individuals in different "own education - father's education" cells and table 14 reports the proportion of individuals earning more than 500 Euros (net) per month (roughly the median wage in Portugal in 1999) for individuals in different "own education - father's education" cells (we compute the mean of the relevant variable for individuals belonging to each cell). Conditional on own education, the effect of parental education on labor market outcomes is relatively weak. Inequality of opportunity generated from being born from parents with different levels of education seems to translate into unequal labor market outcomes mostly through its effects on educational attainment. However, this is effect is quite strong and contributes to a persistence of inequality between individuals from different backgrounds for several generations to come.

Table 13 - Paternal Schooling, Child's Schooling and Child's Full Time Employment
Data from Social Inequalities II

| Father/Child | None | Inc. Prim. | Prim. | Inc. Sec. | Sec. | Inc. Univ. | Univ. | Total |
|---|---|---|---|---|---|---|---|---|
| None | 0.27 | 0.31 | 0.59 | 0.73 | 1.00 | - | 1.00 | 0.53 |
| Incomplete Primary | 0.00 | 0.50 | 0.58 | 0.80 | 1.00 | - | 1.00 | 0.62 |
| Primary | 0.25 | 0.40 | 0.58 | 0.81 | 0.85 | 0.66 | 0.88 | 0.74 |
| Incomplete Secondary | - | - | 0.50 | 0.65 | 0.87 | 0.71 | 0.95 | 0.78 |
| Secondary | - | - | - | 1.00 | 0.71 | 1.00 | 0.80 | 0.82 |
| Incomplete University | - | - | - | - | 1.00 | - | - | 1.00 |
| University | - | - | - | - | 0.50 | 1.00 | 0.88 | 0.86 |

Table 14 - Paternal Schooling, Child's Schooling and
Proportion of Children Earning more than 500 Euros a Month
Data from Social Inequalities II

| Father/Child | None | Inc. Prim. | Prim. | Inc. Sec. | Sec. | Inc. Univ. | Univ. | Total |
|---|---|---|---|---|---|---|---|---|
| None | 0.00 | 0.10 | 0.17 | 0.34 | 1.00 | - | 1.00 | 0.18 |
| Incomplete Primary | 0.00 | 0.00 | 0.19 | 0.31 | 0.66 | 1.00 | 1.00 | 0.22 |
| Primary | 0.00 | 0.11 | 0.17 | 0.39 | 0.58 | 0.88 | 0.96 | 0.41 |
| Incomplete Secondary | - | - | 0.00 | 0.40 | 0.62 | 0.85 | 0.95 | 0.63 |
| Secondary | - | - | - | 1.00 | 0.85 | 0.33 | 1.00 | 0.82 |
| Incomplete University | - | - | - | - | 1.00 | - | - | 1.00 |
| University | - | - | - | - | 1.00 | 1.00 | 1.00 | 0.93 |

# 6 Conclusion

In this paper we examined the role of educational attainment on earnings inequality. We showed that years of schooling account for 40 to 50% of the variance of log wages in the portuguese labor market.

Inspired by the Coleman report, we then study the sources of inequality in educational achievement. The major observable factor driving inequality in test scores among adolescents in Portugal is family background, especially the family background of one's peers in school. Measured school resources have a very limited role. However, a large proportion of the variance in student achievement remains unexplained by observable variables.

Finally, there is a large degree of persistence in educational status. In such a system, inequality is likely to persist from generation to generation.

Our study, in conjunction with a large literature on similar topics, has important policy implications. First human capital policy is a powerful tool to address the problems of inequality and poverty for future generations. Second, education policy needs to be creative and recognize the families are the fundamental education institution in society and that the role of traditional input-based school policies is very limited. Therefore, improving the life-chances of poor children requires intervening at early ages when family influences are the most dramatic. Third, human capital policy has important intergenerational effects: improving the skills of the current generation not only improves their opportunities to succeed but it also has dramatic effects of the opportunities of their offspring.

# 7 Appendix

Table A1 - Inequality in School Resources
Data from PISA 2000

| | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|
| Number of Students in School | 990 | 571 | 18 | 3724 |
| School Hours per Year | 896 | 116 | 275 | 1387 |
| Number of Computers / School Size | 0.087 | 0.352 | 0.001 | 3.667 |
| School Size / Number of Teachers | 8.86 | 3.35 | 0.10 | 23.64 |
| Proportion of Teachers with Pedagogy Degree | 0.43 | 0.31 | 0.00 | 1.00 |

Table A2 - Variance Decomposition of Noncognitive Outcomes on Family and School Variables
Percentage of Variance Explained by Different Components
Data from PISA 2000

| Panel A - Regress Test Scores on Family Variables and School Dummies | | | | | |
|---|---|---|---|---|---|
| | Sense of Belonging | Engagement in Reading | Effort and Preserverance | Instrumental Motivation | Interest in Reading |
| V(F) | 0.0609 | 0.0967 | 0.0611 | 0.0355 | 0.0813 |
| V(S) | 0.0445 | 0.0401 | 0.0445 | 0.0479 | 0.0395 |
| COV(F,S) | 0.0143 | 0.0096 | 0.0048 | 0.0041 | 0.0048 |
| V(Residual) | 0.8803 | 0.8535 | 0.8895 | 0.9123 | 0.8742 |
| Panel B - Regress School Dummies on School and Family Characteristics | | | | | |
| | Sense of Belonging | Engagement Belonging | Effort and Preserverance | Instrumental Motivation | Interest in Reading |
| V(F) | 0.0730 | 0.0077 | 0.0066 | 0.0044 | 0.0012 |
| V(S) | 0.0010 | 0.0175 | 0.0225 | 0.0662 | 0.0268 |
| COV(F,S) | 0.0021 | 0.0018 | -0.0003 | 0.0036 | -0.0031 |
| V(Residual) | 0.9240 | 0.9729 | 0.9712 | 0.9256 | 0.9750 |

# References


[1] Blank, R. (1996), It Takes a Nation, Princeton University Press.

[2] Blundell, R., L. Dearden and B. Sianesi (2005), "Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey", Journal of the Royal Statistical Society: Series A, 168(3).

[3] Card, D. (1999), "The Causal Effect of Education on Earnings", in O. Ashenfelter and D. Card, Handbook of Labor Economics, Volume 3, North Holland.

[4] Cardoso, A. (1998), "Earnings Inequality in Portugal: High and Rising?", Review of Income and Wealth, 44, 3.

[5] Cameron, S. and J. Heckman (2001), "The Dynamics of Educational Attainment for Black, Hispanic and White Males", The Journal of Political Economy, 109.

[6] Carneiro, P., C. Crawford and A. Goodman (2005), "The Importance of Cognitive and Non-Cognitive Skills: Evidence from the National Child Development Survey", working paper, University College London.

[7] Carneiro, P. and J. Heckman (2002), "The Evidence on Credit Constraints in Post-Secondary Schooling", The Economic Journal, 112.

[8] Carneiro, P. and J. Heckman (2003), "Human Capital Policy", in J. Heckman and A. Krueger (eds.), Inequality in America : What Role for Human Capital Policies, MIT Press.

[9] Carneiro, P., J. Heckman and E. Vytlacil (2005), "Understanding What Instrumental Variables Estimate: Estimating the Average and Marginal Return to Schooling", working paper, The University of Chicago.

[10] Carneiro, P. and S. Lee (2005), "Changes in College Enrollment and Wage Inequality: Distinguishing Price and Composition Effects", working paper, University College London.

[11] Carneiro, P., C. Meghir and M. Parey (2005), "Maternal Education and Child Outcomes", working paper, University College London.

[12] Carneiro, P. and M. Spaltro (2006), "Equality of Opportunity and Educational Achievement in Europe: Evidence from PISA", working paper, University College London.

[13] Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld and R. York (1966), Equality of Educational Opportunity, Washington, D.C.; U.S. Government Printing Office.

[14] Cunha, F., J. Heckman, L. Lochner and D. Masterov (2005), "Interpreting the Evidence on Life-Cycle Skill Formation", NBER working paper no. 11331.

[15] Currie, J. and D. Thomas (2001), "Early Test Scores, Socioeconomic Status, School Quality and Future Outcomes", Research in Labor Economics, 20.

[16] Grossman, M. (2005), "Education and Non-Market Outcomes", NBER working paper no. 11582.

[17] Hanushek, E. (2002), "Publicly Provided Education", in A. Auerbach and M. Feldstein, Handbook of Public Economics, Volume 4, North-Holland.

[18] Hanushek, E., J. Kain and S. Rivkin (2005), "Teachers, Schools, and Academic Achievement", Econometrica, 73(2).

[19] Harmon, C., I. Walker and N. Westergaard-Nielsen (2001), Education and Earnings in Europe, Cheltenham, Edward Elgar.

[20] Hartog, J, P. Pereira and J. Vieira (2001), "Changing Returns to Education in Portugal during the 1980s and Early 1990s: OLS and Quantile Regression Estimators", Applied Economics, 33.

[21] Heckman, J. (2001), "Policies to Foster Human Capital", Research in Economics, 54.

[22] Heckman, J., J. Hsee and Y. Rubinstein (2000), "The GED is a Mixed Signal", working paper, The University of Chicago.

[23] Heckman, J., J. Sixtrud and S. Urzua (2005), "The effect of cognitive and non-cognitive factors in behavioral and labor outcomes", working paper, The University of Chicago.

[24] ISSP (1999), Social Inequalities II, CD-Rom.

[25] Krueger, A. and D. Whitmore (2001), "The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR", The Economic Journal, 11.

[26] Lochner, L. and E. Moretti (2004), "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Reports", American Economic Review, 94(1).

[27] Machado, J. and J. Mata (2001), "Earning Functions in Portugal 1982-1994: Evidence from Quantile Regressions", Empirical Economics, 26.

[28] Machin, S. and S. McNally (2005), "The Literacy Hour", working paper, University College London.

[29] Manski, C. (2000), "Economic Analysis of Social Interactions", The Journal of Economic Perspectives, 14(3).

[30] OECD (2001), Knowledge and Skills for Life, OECD.

[31] OECD (2002a), Education Policy Analysis, OECD.

[32] OECD (2002b), Manual for the PISA 2000 Database, OECD.

[33] OECD (2003), Literacy Skills for the World of Tomorrow, OECD.

[34] OECD (2005), Society at a Glance, OECD.

[35] Parente, P. and T. B. D'Uva (2002), "Evolucao da Pobreza e da Desigualdade em Portugal no periodo 1995 a 1997", Instituto Nacional de Estatistica.

[36] Pereira, P. and P. Martins (2000), "Does Education Reduce Wage Inequality? Quantile Regressions Evidence from Fifteen European Countries", IZA Discussion Paper no 120.

[37] Portugal, P. (2004), "Mitos e Factos sobre o Mercado de Trabalho Portugues: a Tragica Fortuna dos Licenciados", Boletim Trimestral do Banco de Portugal, Marco.

[38] Sacerdote, B. (2000), "Peer Effects with Random Assignment: Results for Dartmouth Roommates", NBER working paper no. 7469.

[39] Smith, J. (2005), "Unravelling the SES Health Connection", Aging, Health and Public Policy, vol. 30.

[40] Solon, G. (1999), "Intergenerational Mobility in the Labor Market", in O. Ashenfelter and D. Card, Handbook of Labor Economics, Volume 3, North Holland.