Intergenerational persistence of educational attainment in Italy[☆]

Daniele Checchi [a,b], Carlo V. Fiorio [a,c,*], Marco Leonardi [a,b]

[a] *University of Milan, Department of Economics, Management and Quantitative Methods, Italy*
[b] *IZA, Germany*
[c] *Econpubblica, Italy*

ABSTRACT

In this paper we suggest a simple decomposition of the correlation coefficient of education to account for the different intergenerational mobility of subgroups of the population, which is of key importance from a policy perspective.

Focusing on the interesting Italian case, we show that the high persistence of educational attainment found in the data is due to a much larger probability of obtaining a college degree of children of highly educated fathers.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The literature on the temporal evolution of educational attainment by parental background is vast, and the correlation coefficient between parents' and children's years of education is the main descriptive statistic used therein.

Blossfeld and Shavit (1993) produced one of the first comparative studies of intergenerational persistence in education by studying the correlation of children's attainment with parental background by age cohorts and claimed that the expansion of higher education gave no contribution to improving intergenerational mobility. More recently, Hertz et al. (2007) provide a large cross-country analysis of intergenerational correlations in educational attainment, documenting large regional differences in educational persistence. Their main conclusion is that global average educational persistence, measured as the correlation between parents' and children's schooling, has remained substantially stable over the last 50 years, despite the increased participation in school of recent cohorts.

In our opinion, the use of the regression coefficient of fathers' and children's years of education has two main shortcomings. It does not allow one to account for differences in average schooling across generations and, above all, to disentangle differential intergenerational mobility of subgroups of the population, which is of key importance from a policy perspective. For instance, the correlation coefficient could decrease because compulsory education increased the amount of education for recent cohorts, or because the upward mobility of children of educated fathers dominates the immobility of children of uneducated fathers. This limitation remains also using the standardized correlation coefficient, where parents' and children's years of education are divided by their respective standard deviations to account for different dispersion of education of different cohorts. In Section 2 we suggest a straightforward decomposition of the correlation coefficient of education, focusing on the probability of one's educational attainment given that of one's father.

Hence we focus on Italy as an interesting case study. Hertz et al. (2007) also document the decreasing intergenerational persistence of educational attainment in Italy, whose absolute levels remain high compared with similarly developed countries. This result is consistent also with Checchi et al. (1999) and d'Addio (2007). In

Section 3, using the Bank of Italy data set on Household Income and Wealth, we confirm evidence of the declining intergenerational persistence of education in Italy across different age cohorts and investigate why it has decreased so slowly over time by decomposing the standardized correlation coefficient. We find that the high level of intergenerational persistence of education is largely due to the fact that higher degrees are disproportionately more likely to be attained by children with highly educated fathers.

This decomposition could be easily replicated for better understanding of the trend of educational persistence in other countries.

2. A conceptual framework

The analysis of the intergenerational transmission of education over time is often performed by a univariate regression to be estimated separately for each cohort, such as

$$c_i = \alpha + \rho f_i + \varepsilon_i \quad \text{for } i = 1, \dots, N, \tag{1}$$

where $c_i := C_i/\sigma_c$ and $f_i := F_i/\sigma_f$ are the number of years of education of child i (C_i) and of father (F_i) normalized by their corresponding standard deviations (σ_c, σ_f, respectively), ε_i is an error term and ρ is the correlation coefficient.[1] When the number of years of education is not readily available, a common estimation strategy is to replace the level of education attained with the number of regular years needed to obtain it Black and Devereux (2010).

The coefficient ρ could be interpreted as a measure of the inequality of opportunities due to circumstances, which are independent of a child's effort. However, changes in ρ capture not only changes in the child–father education transmission, but different phenomena such as the secular rise in schooling and changes in compulsory education.

To illustrate this issue, let us rewrite the correlation coefficient as

$$\rho = \int \underbrace{(c - E(c))(f - E(f))}_{(a)} \underbrace{\Pr(c|f)}_{(b)} \underbrace{\Pr(f)}_{(c)}. \tag{2}$$

This shows that ρ may change over time because of changes in the dispersion of children's or of fathers' (standardized) education around their respective means (term (a)), because of changes in children education conditional on that of their fathers (term (b)) or because of changes in the unconditional distribution of fathers' education (term (c)).

Changes in term (a) can be due to a uniform convergence towards higher levels of education. Term (c) could vary because of institutional framework changes that often go along with the development of a country and increase the level of compulsory education of fathers across generations. Here we suggest to focus on term (b), i.e. on the distribution of children's education conditional on that of their fathers, as the policy-relevant indicator of intergenerational persistence in educational attainment. In fact, by decomposing the conditional probability $\Pr(c|f)$ and denoting with $t, j \in \{0, 1, \dots, E\}$ the educational degree attained, one can easily compute the following.

An immobility index,

$$\sum_t \Pr(c = t|f = t). \tag{3}$$

An upward mobility index,

$$\sum_{j>t} \Pr(c = j|f = t). \tag{4}$$

A downward mobility index,

$$\sum_{j<t} \Pr(c = j|f = t). \tag{5}$$

A "family premium" for education level t,

$$d \Pr(c = t|f = j) \equiv \Pr(c = t|f = j) - \Pr(c = t). \tag{6}$$

A positive (negative) family premium for education level t means that the family background increases (decreases) one's probability of achieving it.

3. Data and empirical analysis

The empirical analysis of intergenerational transmission of education requires a data set providing information on the education of children and their parents over time. Here we use the Survey on Household Income and Wealth Historical Archive (SHIW) produced by the Bank of Italy based on biannual surveys, which provides a representative sample of the Italian population in each survey year. Starting from 1993, the SHIW contains a section asking information on the householder's and spouse's parents when they were of the same age as the interviewees, including their education, and we use this information extensively. We pool SHIW waves from 1993 to 2008, selecting only the householders and, when present, their partners. We name "children" the householders and their spouse and analyse their educational achievement as opposed to that of their respective fathers. As for parents, we retain only fathers' educational attainment[2] and we select only individuals whose age is over 30 at the time of the interview, to reduce the selectivity due to early marriages and/or yet uncompleted educational careers. Finally, the data set is organized by five-year cohorts of children's birth years.

Table 1 reports the unconditional distribution of the highest educational attainment of children and fathers organized by five-year birth cohorts of children, where educational attainment has been replaced with the legal duration of the degrees considered (i.e. 0, 5, 8, 13, and 18 for no education, primary, lower secondary, upper secondary, and college education, respectively). The table is divided into two parts, one referring to fathers and one to children. The former reports term (c) of Eq. (2), showing that indeed the marginal distribution of educational attainment of fathers changed over time, increasing their average years of education. The latter shows that also in the children generation the marginal distribution has changed, the percentage of children with no degree decreasing constantly over time. An increasing proportion of children nowadays attains a high school or a college degree: in the most recent cohorts, slightly less than 50% have lower secondary degree and about 10% have a college degree, with an average differential in years of education with respect to their fathers of about 4 years.

The correlation coefficient between standardized children's and fathers' numbers of years of education attained exhibits a constant reduction from 0.63 to 0.50 over the period of investigation. We would like to know whether the relatively high level of the correlation coefficient even in the most recent cohorts, decades after the 1963 reform that raised the amount of compulsory

---

[1] Alternatively, one could estimate the model $C_i = \gamma + \beta F_i + \upsilon_i$ and interpret β as the elasticity coefficient, provided that the number of years of education variables is measured in logs. The advantage of estimating the correlation coefficient ρ is that it factors out the difference in the variance of educational attainment across generations.

[2] In previous versions of the paper we considered also the education of the mother but the results are substantially unaffected, most likely due to assortative mating.