Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento Economia e Statistica Cognetti de Martiis

Working Paper Series

13/20

PARENTAL EDUCATION AND SONS' EARNINGS: A "BEYOND THE MEAN" APPROACH ALONG THE SONS' EARNINGS DISTRIBUTIONS

ROBERTO ZOTTI

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Parental education and sons' earnings: a "Beyond the Mean" approach along the sons' earnings distributions

Roberto Zotti[1]

[1]Department of Economics and Statistics "Cognetti de Martiis", University of Torino

Preliminary version

Abstract

This paper offers evidence on the parental education and sons' earnings relationship by analyzing the father and mother education gradient across the full distribution of sons' earnings. It uses an unconditional quantile approach based on recentered influence function regressions and applies an Oaxaca-Blinder decomposition at various quantiles of the earning distribution to explain time, gender and geographical differentials in earnings. Using six waves of the Survey of Income and Wealth (from 2004 to 2014) for Italy, I find evidence of higher returns to family education in the upper percentiles of the distribution of son's earnings with the probability of ending up in high deciles being significantly correlated with the education level of the father. Results show an important heterogeneity in the association of parental education as well as of individual covariates to sons' earnings across time, gender and geographical areas of the country which varies significantly along the earning distribution and accounts for a substantial percentage of the differentials in observed earnings.

Keywords— Intergenerational inequality; Education; Unconditional quantile regressions; Social mobility

JEL codes— J31; J62; C21

1 Introduction

The link between an individual's income and her family background is a topic of great interest and it is an issue of academic, social and political concern. Many are the reasons to be interested in the association between income and family background. Indeed, in front of strong income persistence across generations, a higher inequality in the parental generation may be the cause of inequality also in the future generation; in a society characterized by a strong association in income among parents and children there is less equality of opportunity then a society in which the correlation between family members' income is weak. The presence of such mechanism calls into question policy interventions that weaken the transfer of inequality from one generation to the next (Björklund and Jäntti, 2012). In other words, the extent to which a family's economic advantage or disadvantage persists across generations is widely seen as a key indicator of equality of opportunity (Schnitzlein, 2015). As underlined by Raitano and Vona (2015*a*), economic theory generally studies intergenerational inequality through the lens of human capital theory (Solon, 2004) and the literature acknowledges different channels through which family background affects skill formation, including financial constraints (Becker and Tomes, 1979, 1986), peer effects (Benabou, 1996), educational policies (Schutz et al., 2008), soft skills (Bowles and Gintis, 2002) and the cumulative effects of educational investments since early childhood (Cunha and Heckman, 2007). As a result, family background not only influences the years of education attained by the child but also its quality (Bratsberg et al., 2007).

More specifically, the association between parental characteristics and children earning is an interest issue also from the empirical point of view as it has been usually assessed at the mean of the distribution, thus not providing insights on the mechanism lying behind such association; indeed, there is no a-priori reason to believe that the intergenerational transmission of economic resources is the same in all parts of these distributions (Black and Devereux, 1993). There have been a small number of studies that have considered whether intergenerational mobility differs at different parts of the distribution of parental income (Bratsberg et al., 2007), seeking to understand whether the association between family background and sons' life chances are greater for those from the poorest or richest families.

This study firstly highlights the association between parental education and sons' earning and how it varies along the sons' earnings distribution. Secondly, the study aims to explore whether the mechanism generating the intergenerational transmission has changed over the years, gender and geographical areas of the country. In order to address the first issue, the paper applies the Unconditional Quantile Regression approach (Firpo et al., 2009) (UQR) which allows to explore the presence of non-linearities in the intergenerational transmission along the different percentiles of the sons' earnings distribution. With regard to the second point, instead, the paper uses the Oaxaca-Blinder (Blinder, 1973, Oaxaca, 1973) (OB) decomposition method at various quantiles of the sons' earnings distributions, in order to analyse year, gender and geographical differentials in sons' earnings and to measure the contribution of parental education (and other covariates) to these differentials.

According to the sociological literature (Erikson and Goldthorpe, 1992, Ganzeboom and Treiman, 2007, Granovetter, 1995), parental occupation is a good proxy for the influence of the family on sons' outcomes and it encompasses unobservable aspects of human capital, socio-economic status and family networks such as the individual's position in the social scale, its capacity to influence economic decisions or of being a part of certain social networks. However, as a proxy for family background, I use, instead, another available variable such as the level of education achieved by both parents. It is a measure of social origin widely used by economists (Ermisch and Del Bono, 2012a, Bradbury et al., 2012) and sociologists (Bukodi and Goldthorpe, 2012), and has been shown to influence child development (Dickson et al., 2016, Chevalier et al., 2013), access to higher education (Cunha et al., 2006, Jerrim, 2017) and other aspects of the intergenerational transmission process (Lampard, 2007). It has also been widely used in international comparisons of intergenerational inequalities (Ermisch et al., 2012a, Jackson, 2013); according to Ermisch et al. (2012b), parental education is correlated with the financial resources available to parents for investing in their children's development and it is a good indicator of family background in terms of being made comparable across countries.

I find evidence of higher returns to family education in the upper percentiles of the distribution of son's earnings with the probability of ending up in high deciles being significantly correlated with the education level of the father. Results show an important heterogeneity in the association of parental education as well as of individual covariates to sons' earnings across time, gender and geographical areas of the country which varies significantly along the earning distribution and accounts for a substantial percentage of the differentials in observed earnings.

The rest of the paper is organized as follows. The next section presents a brief view of the literature. Section 3 discusses the empirical framework while Section 4 presents the data and descriptives. Section 5 presents the results. The final section summarizes and concludes.

2 Literature review

The majority of the existing research on intergenerational mobility has investigated the transmission of incomes or earnings across generations at the mean of the distribution (see for instance Corak (2013) for a review of the existing literature) focusing the attention on the link between the income of fathers and the income of their sons. More specifically, it focuses mainly on non-linearities, by either including higher order polynomials of father's income or sons' earnings or using nonparametric regression techniques. Bratsberg et al. (2007) present a very comprehensive study of the non-linearity of the relationship between family income across a range of countries such as Denmark, Finland, Norway, the UK and the US, and reveal some important differences. While in the US and the UK the relationship turns out to be rather linear, the Nordic countries show a convex pattern, where the relationship starts out flat and is increasingly positive in the middle and upper segments of the distribution of incomes. This suggests that sons growing up in the poorest families have similar life chances to those born in moderately poor families, while coming from a well off family increases the chances of keeping the same economic advantage in the future. Likewise in Canada, Corak and Heisz (1999) find that the intergenerational elasticity is almost equal to zero in lower parts of the distribution of father's earnings and increases along the father's earnings distribution up to 0.4, using a non-parametric method.

An alternative way of characterizing the distribution of sons' earnings is to compute its quantiles (capturing the impact of an explanatory variable on the conditional distribution of the dependent variable at a given point in the distribution). This is to explore non-linearities along the distribution of son's earnings, changing the focus to the outcome of the intergenerational mobility process rather than the origin. This literature has mainly relied upon quantile regression estimates (Koenker and Bassett, 1978) and the intergenerational setting allows the researcher to capture differences in the persistence of fathers' earnings (parental income) across the conditional distribution of sons' earnings. Findings based on Conditional Quantile Regressions (CQR) for the US, Norway and Canada all show that intergenerational persistence is higher at the lower end of the sons' earnings distribution than at the upper end. This empirical evidence suggests that where you come from sorts people more for those who end up at the bottom of the earnings distribution among top jobs, where it is more equitable in these countries. In the US, Eide and Showalter (1999) show that the estimated intergenerational elasticity at the mean of the distribution is 0.45, while the conditional quantile regression estimates reveal a greater intergenerational elasticity at the bottom of the son's conditional distribution rather than the top (0.67 for the 10th percentile against 0.26 for the 90th percentile). Bratberg et al. (2005) and Grawe (2004) demonstrate similar findings for Norway and Canada.

Recently, part of the literature extends the analysis of non-linearities by applying the Unconditional Quantile Regression (UQR) method developed by Firpo et al. (2009). The main idea is that this approach provides information on the marginal effect of parental earning at a given percentile of the unconditional distribution of the child's earnings. This literature still focuses the attention on the link between the income of fathers and the income of their sons (i.e. looks at the intergenerational elasticity). Schnitzlein (2015), using the German Socio Economic Panel, finds more intergenerational persistence at the top of the son's distribution in Germany as well as the US using both standard conditional and unconditional quantile rgegression techniques. He finds no evidence for non-linearities along the fathers' earnings distribution and shows that mobilty is higher for the sons at the lowest quartile of the sons' earning distribution in both countries. Specifically, in Germany, this result is mainly driven by a high downward mobility of sons with fathers in the upper middle part of the earnings distribution (this pattern is less pronounced in the United States). Gregg et al. (2018), using the UQR and the British Cohort Study, investigate whether parental income is a more potential predictor of opportunities at the top and the botton of the sons' earnings compared to those in the middle. They find a J-shaped relationship between parental income and sons' earnings and that parental income is a strong predictor of labour market sources for those at bottom, and to a greater extent, for those at the top of the distribution. They also find that returns to family background are increasing across the son's earning distribution. According to the authors, parental income dominates education at the top of the distribution of earnings.

A different approach in the literaturelooks, instead, at the link between the income of fathers and the income of their sons, investigates the presence of non-linearities in the son's earning distribution by using parental occupation or parental education to measure family background. Raitano and Vona (2015*b*), using the EU-SILC data, find that parental occupation is a stronger predictor of earnings for those that end up at the top of the earnings distribution rather than at the bottom. This result is consistent across some European countries with the exception of Spain and particularly strong for the UK. Jerrim (2017) finds an increasing relationship between parental education and sons' earnings for France, Germany and Switzerland but a stable or slightly declining relationship across the distribution of sons' earnings for the UK. More specifically, he finds that the gap between the "least successful" (i.e. lowest earnings) individual from high parental education backgrounds and the "least successful" (i.e. lowest earning) individual from low parental background is particularly pronounced in UK. Raitano et al. (2016) analyse the effect of parental background (father and mother occupation and education) along the sons' earnings distribution. Using an UQR and EU-SILC data, they find that returns to background are higher at the top of the distribution. More specifically, returns to parental background increases along the son's earning distribution and the probability of ending up in high deciles is significantly correlated with parental background. Finally, Naticchioni et al. (2016) explore, using the UQR, whether the deterioration of earnings dynamics in the early phase of the career is homogeneous across skill levels. They find that for unskilled workers there are no differences across cohorts. Differences begin to emerge at median earnings and deterioration across cohorts is much more marked for skilled individuals. Indeed, they show that the high skilled workers of the most recent cohorts have suffered, compared to the previous cohorts, an earnings penalty much more severe than that experienced by unskilled workers.

3 The empirical framework

The empirical strategy is based on the unconditional quantile regression approach developed by Firpo et al. (2009) to estimate the impact of a marginal change in parental education on the entire distribution of sons' earnings. Then, an Oaxaca-Blinder decomposition (Blinder, 1973, Oaxaca, 1973) at various quantiles of sons' earnings distribution is applied in order to explore whether the mechanism generating the intergenerational transmission has changed over the years, across genders and geographical areas in Italy.

3.1 Unconditional quantile regression

Firstly, I use an unconditional quantile approach (UQR) based on Recentered Influence Function (RIF) regressions (Fortin et al., 2011) to estimate the impact of a marginal change in parental education on the entire distribution of son's earnings. More specifically, this method allows to properly assess the impact of parental education at different points of the unconditional distribution of sons' earnings. This is desirable in order to check for non-linearities in the relationship between parents' education and sons' earnings and to assess the role of parental background at extreme sons' earnings levels which often indicate the presence of a mechanism generating intergenerational inequality. This possibility is essentially given by the fact that RIF method works by providing a linear approximation of the unconditional quantiles of the dependent variable. The law of iterated expectations can be applied to the quantile being approximated and used to estimate the marginal effect of a covariate through a simple regression of a function of the outcome variable, the Recentered Influence Function, on the covariates X.

In my setting, the RIF of sons' earnings is estimated directly from the data by first computing the sample quantile and then estimating the density of the distribution of sons' earnings at that quantile using kernel density methods. Then, for a given observed quantile q, a RIF is generated which can take one of two values depending upon whether or not the observation's value of the outcome variable is less than or equal to the observed quantile:

$$RIF(Earnings; q_\tau) = q_\tau + \frac{\tau - 1[Earnings \leq q_\tau]}{f_{Earnings}(q_\tau)} \tag{1}$$

Where q_τ is the observed sample quantile, $1[Earnings \leq q_\tau]$ is an indicator variable equal to one if the observation's value of the sons' earnings is less than or equal to the observed quantile and zero otherwise. $f_{Earnings}(q_\tau)$is the estimated kernel density of the sons' earnings at the τ_{th} quantile.

The RIF defined in equation (1) is then used as a dependent variable in a OLS regression on the covariates X. In practice, this amounts to estimate a rescaled linear probability model; indeed, the unconditional quantile of the sons' earnings q_τ, may be obtained as follows:

$$q_\tau = E_x[E[RIF(\widehat{Earnings}; q_\tau)|X]] \tag{2}$$

Where $RIF(\widehat{Income}; q_\tau)|X$ is the estimate of RIF as defined in equation (1) conditional on covariates X. Thanks to this linear approximation, it is now possible to apply the law of iterated expectations. Thus, q_τ can be written as :

$$q_\tau = E[X]\hat{\delta}_\tau \tag{3}$$

Where $\hat{\delta}_\tau$ is the coefficient of the unconditional quantile regression. This linearization allows estimation of the marginal effect of a change in distribution of covariates X (including parental education) on the unconditional quantile of sons' earnings, measured by the parameter $\hat{\delta}_\tau$.

Summarizing, the contribution of the UQR and the differences with respect to the other standard regression methods such as OLS or quantile regressions, are based ont the fact that the UQR method is similar to a standard linear regression but the dependent variable is replaced by the RIF of a distributional parameter of interest, in our case a given percentile. In addition to the standard quantile regression, the UQR allows us to retrieve the marginal impact of any explanatory variable and the percentiles of the unconditional distribution of the dependent variable while controlling for other covariates. More specifically, differently to the standard (conditional) quantile regression, UQR estimates provide information on the marginal effect of parental earnings at a given percentile of the unconditional distribution of the sons' earnings. Thus, this method allows us to determine whether the effect of parental education differs along the unconditional child's earnings distribution. In this way the focus is on the outcome of the intergenerational transmission process such as the position of the sons in their own earnings distribution. More specifically, in a standard OLS regression, β can both be interpreted as the association between an explanatory variable on the conditional mean of the dependent variable (conditional mean interpretation) as well as the effect of increasing the mean value of an explanatory variable on the unconditional mean value of the dependent variable (unconditional mean interpretation). By contrast, when using Conditional Quantile Regression models for the τ_{th} conditional quantile, only the conditional quantile interpretation can be applied such as the effect of an explanatory variable on the conditional quantile τ_{th} of Y given $X : q_t(x) = X\beta_t$. This is because the law of iterated expectation does not apply in the case of quantiles. In other words, β_t cannot be interpreted as the effect of increasing the mean value of X on the unconditional outcome variable at quantile q_t. Therefore estimates based on conditional quantile regressions may lead to confusing results and need to be interpreted with caution (see Fournier and Koske (2012) for a discussion).

On the other hand, instead, UQR allows us to estimate the association between an explanatory variables and quantiles q_t (or other distributional parameters) of the unconditional (marginal) distribution of the outcome variable using the RIF technique. This method builds upon the concept of the influence function which is a tool used to obtain robust estimates of statistical and econometric models, measuring the influence of an individual observation on a distributional statistic of interest (Monti, 1991). This RIF-regression is similar to a standard regression except that the dependent variable is replaced by the RIF of the statistic of interest, ν (i.e. quantile), RIF(y, ν). The RIF is obtained by adding the distributional parameter concerned to the influence function IF(y, ν). In its simplest form the conditional expectation of the RIF can be modeled as a linear function of the explanatory variables and the parameters can simply be estimated using standard OLS regressions. The expected value of the RIF is equivalent to its statistic of interest (i.e quantile) and by applying the law of iterated expectations, it is possible to write:

$$q_\tau = E[RIF(y; q_\tau)] = E_x\{E[RIF(y; q_\tau)|X]\} = X'\gamma_\tau \tag{4}$$

If the statistic of interest is the quantile ($\nu = q_\tau$), Firpo et al. (2009) refer to this RIF-regression also as an unconditional quantile regression. Basically, what we are going to estimate at the end is the following:

$$RIF(Y_i^{sons' earnings}; q_\tau) = \alpha^\tau + \beta'^\tau y_i^{parents' education} + X'\theta^\tau + \epsilon_i \tag{5}$$

using a RIF regression at different quantiles q_τ where τ=0.25, 0.50, 0.75, 0.90, 0.95 (25th, 50th, 75th, 90th and 95th) and X is a vector of sons' characteristics such as gender, age, marital status, education, residence, professional status.

Therefore, the paper explores the relationship between family education and sons' earnings across the distribution of sons' earnings (β'^τ), conditional on some individual characteristics of the sons; in other words, the paper assesses how $\hat{\beta}$ varies at different parts of the distribution of earnings of the second generation. This allows us to understand whether family background (in childhood) has a strong association with later earnings for those who end up being rich compared to those who end up being poor. Moreover, I can also explore the heterogeneous returns to education and other individual characteristics of the sons across the distribution of the sons' earnings.

3.2 Oaxaca-Blinder decomposition

To explore whether the mechanism generating the intergenerational transmission has changed over the years, gender and geographical areas, I use Oaxaca-Blinder (Blinder, 1973, Oaxaca, 1973) decomposition method using the RIF regression in equation (2) as a basis for the decomposition. A similar logic to the OB decomposition at the mean applies also here (see Fortin et al. (2011) for a review). Basically, I apply an Oaxaca-Blinder decomposition at various quantiles of the sons' earnings distributions to analyse year, gender and geographic differentials in sons' earnings and to measure the contribution of parental education (and other individual covariates) to these differentials. More specifically, Oaxaca-Blinder decomposition assesses to what extent differentials in sons' earnings are explained by compositional differences, ie. differences in observed covariates across years, gender and areas, or by differences in the association of sons' earnings across years, gender and areas.

Formally, differences in estimated sons' earnings between years (2014 versus 2004), gender (men versus women) and geographical areas (North versus South), at each quantile, can be decomposed as follows (for simplicity equations are written as a measure of year differential, but can be easily extended and modified to account for gender and geographical differences)[1]:

$$\Delta^{\tau}_{Earnings} = [RIF(\widehat{Earnings}2014\tau, q_{2014})] - [RIF(\widehat{Earnings}2004\tau, q_{2004})] \tag{6}$$

$$\Delta^{\tau}_{Earnings} = (\bar{X}_{2014} - \bar{X}_{2004}]\delta_{pooled} + X_{pooled}(\bar{\delta}_{2014} - \bar{\delta}_{2004}) \tag{7}$$

where $\bar{X}_{2014}$ and $\bar{X}_{2004}$ represent the sample means of covariates X for the subsample of individuals in the 2014 wave and 2004 wave, and $\bar{\delta}_{2014}$ and $\bar{\delta}_{2004}$ represent the coefficients of the unconditional quantile regression as in equation (2) for the subsample of individuals in the 2014 wave and 2004 wave, respectively.

The first term in equation (7) is the part of differential in sons' earnings that is "explained" by differences in observed covariates between the subsample of individuals in 2014 wave and those in 2004 wave. This is often called as a "composition effect". Differences in covariates across years are weighted by the coefficients of the unconditional quantile regression from a model estimated on the pooled sample ($\bar{\delta}_{pooled}$). The decomposition is not formulated from the viewpoint of individuals in 2014 wave or 2002 wave because the choice of the "discriminated" group is complicated by the fact that we are not aware of any reason according to which individuals in 2014 might have differences respect to those in 2002. Therefore, our reference group will be the pooled sample of individuals.

The second term in equation (7) measures, instead, the "unexplained" of the differential in sons' earnings. This is often called also as "structural" part and it accounts for differences in earnings across years (2014 and 2004) which is due to differences in the impact of the covariates and it also captures all potential effects of differences in unobserved variables.

The explained and unexplained part can be further decomposed into contributions of each covariate at each quantile. It is particularly useful to derive both the total contribution and the detailed contribution of the parental education to the years differentials in sons' earnings. This allows us to understand to what extent differences in sons' earnings are driven by differences in parental education between those in 2014 and those in 2004 ("explained part") and/or by differences in the association of sons'

[1]In order to consider gender differentials, substitute 2014 with men and 2004 with women. In order to consider geographical area differentials, substitute 2014 with north and 2004 with south

earnings to parental education across years ("unexplained part"). Thanks to the additivity assumption of the OB decomposition, this is possible because the "explained" and "unexplained" part in equation (7) are simply given by the sum of the contribution of individual covariates. To draw inference on the contributions of each covariate to the explained and unexplained part, standard errors are computed using the delta method (See Jann (2008) for more details)

4 Data

I use microdata provided by the Survey on Household Income and Wealth (SHIW) which has been carried out every two years by the Bank of Italy. The dataset contains information regarding the family education (illiterate, primary, secondary and university). Individual information of the sons have also been included in the analysis such as their age, gender, marital status, education, residence and professional status. In what follows, I provide a description of the variables used in the analysis. A complete list of the variables along with some descriptive statistics is then presented in Table 1.

The main outcome of interest is individual earnings (*Wage*) measured in log in all regressions.

The main independent variable is parental education. More specifically, four variables are used to take into account formal education obtained both by the father and mother of the individuals interviewed. *Father illiterate*, *Father primary*, *Father secondary* and *Father university* are dummy variables taking the value of 1 in case the father obtained no education, primary, secondary, or university education, and 0 otherwise. *Mother illiterate*, *Mother primary*, *Mother secondary* and *Mother university* are dummy variables taking the value of 1 in case the Mother obtained no education, primary, secondary, or university education, and 0 otherwise. In alternative, I also use a continuous variable measuring the number of years of schooling obtained by the father and mother (*Father schooling; Mother schooling*).

As control variables, I include age of individuals (*Age*), a dummy variable taking the value of 1 if the individual is a male and 0 a female (*Gender*), four dummies controlling for marital status (*Single, Married, Divorced, Widow*), four dummies controlling for educational status (*Illiterate, Primary, Secondary, University*) and 5 dummies controlling for professional status (*Blue collar, Office worker, Manager, Member of profession, Self-Employed*). I also include controls for the geographical area of residence of individuals (*North, Centre, South*). Omitted categories in our analysis are females, single, and individuals in the South, with no qualification, having father and mother with no qualification and being blue collar.

Table 1 shows the descriptive statistics. With respect to parental education, we observe that around 50% of individuals have a father or a mother with primary education and around 30% of individuals have a father or a mother with secondary education. Only 5% of individuals have a father with university education while even less individuals have a mother with university education (2%). Around 70% of individuals are married. The share of individuals having a formal secondary education is substantial (approximately 76%) while only 17% have a university degree.

5 Results

5.1 Parental education-sons' earnings relationship

Table 2 shows the results of RIF regressions described in equation (2) for individual earnings. Column 1 of the table includes OLS regression at the mean for comparison, while columns 2-6 include results of the RIF regressions at the 25th, 50th, 75th, 90th and 95th percentile of earnings. In order to make the interpretation of the coefficients of interest easier, I also plotted the main covariates coefficients (father and mother education) at all points of the earnings distribution (every 5 percentiles) in Figures 1-2.

The first consideration is that the relationship between parent's education and son's earnings is complex such that a standard analysis at the mean misses important information. Indeed, OLS estimates suggest only a positive education gradient (column 1), while the RIF regression indicates that the education (parents) – earnings (son) relationship varies at different points of the earning distribution.

Table 2 shows that the father's education – sons' earnings relationship is in general positive and statistically significant. More specifically, the relationship between the education (primary) of the father and the earnings of the son is positive and statistically significant up to the 75th percentile of earnings (i.e. not significant at the highest quantiles of the distribution), while is statistically significant at the lowest quantiles of the earnings distribution (25th, 50th and 75th percentiles). With regard to the other levels of father education, instead, the relationship between the education (secondary and university) of the father and the earnings of the son is positive and statistically significant at all quintiles of the earning distribution. Having a father with a university degree is particularly relevant as also Figure 2 shows being the 60th percentile a sort of cut-off point. Indeed, after this threshold the father tertiary education gradient increases in magnitude and reaches a peak at the 95th percentile with a highly statistically significant coefficient of 0.522.

The situation changes when the mother education – son's earnings relationship has been taken into account. Indeed, interestingly, now it is positive and statistical significant only when secondary education is taken into account. Indeed, the relationship between the education (secondary) of the mother and the earnings of the son is positive and statistically significant only at the medium-high quantiles of the distribution (50th, 75th, 90th and 95th percentiles), while the relationship turns to be not statistically significant at the lowest percentiles of the earning distribution. See also Figure 2 for more graphical details where it is clear that plotted coefficients for the RIF regression related to the mother secondary education (rd line) almost always dominates primary and tertiary education (blu and green line, respectively).

With respect to the other variables, instead, we find higher earnings levels among males, older and married individuals, and also a high association with education at all levels of the earning distribution (especially when individuals with university degree are considered).

Results are confirmed also when father and mother years of schooling have been used instead of education categorical variables. Results, summarized in Table A1 in Appendix, confirm a positive education gradient at all points of the distribution when father years of schooling are taken into account. The years of schooling coefficient at the 90th and 95th percentile is almost three times higher than the coefficient at the 25th percentile. The coefficients, although mainly positive, become not statistically significant for the mother years of schooling. This pattern is more clearly depicted in Figures A1 and A2 in the Appendix.

Overall, results show higher returns to family education in the upper percentiles of the distribution of son's earnings. This is particularly evident considering the relationship between the education (university) of the father and the earnings of the son.

5.2 Oaxaca-Blinder decomposition of year differentials in earnings

In order to perform the Oaxaca-Blinder decomposition and to explore whether the mechanism generating the intergenerational transmission has changed over the years, I use data from two surveys which have been carried out in 2004 (first observed year in the data) and 2014 (last observed year in the data).

I first estimate the parent's education – son's earnings relationship in years 2004 and 2014. Results are summarized in Table A2 and Table A3 in Appendix, for year 2004 and year 2014, respectively. The empirical evidence confirms the positive and statistically significant relationship between father's education and son's earnings along the whole distribution of earnings both in 2004 (see Table A2) and in 2014 (see Table A3). Having a father with a university degree is particularly relevant reaching its peak at the 90th and 95th percentiles with a highly statistically significant coefficient of 0.483 and 0.553 for year 2004 (Table A2, Columns 5 and 6) and of 0.495 and 0.794 for year 2014 (Table A3, Columns 5 and 6). Again, the relationship between the education of the mother and the earnings of the son turns out to be not statistically significant in all the percentiles of the earnings distribution. See Figures A3 and A4 (for Year 2004) as well as Figures A5 and A6 (for Year 2014) for more graphical details.

The results of the Oaxaca-Blinder detailed decomposition of the 25th, 50th, 75th, 90th, and 95th percentile distribution of earnings are shown in Table 3. Decomposition is explained as a difference between levels for year 2014 minus levels for year 2004. Table 3 also includes total differences, the explained and the unexplained part and their respective standard errors.

Results show that Year 2014 have generally slightly higher values of earnings at all quantiles of earning distribution (with the exception of the 90th percentile). More importantly, year differentials are not the same along the entire distribution of earnings. Indeed, they are negligible and not statistically significant at the 90th and 95th percentile of the earnings distribution while they are statistically significant at the 25th (0.004 points), 50th (0.086) and 75th (around 0.11 points) percentile.

The last two rows of Table 3 show that the difference among years is explained both by compositional differences in covariates, which is the explained part (i.e. endowments) and by a difference in the impact of covariates on earnings, which is the unexplained part (i.e. coefficients). More specifically, the explained part represents the mean increase in earnings in 2014, respect to the average earnings between 2004 and 2014, if the 2004 year had the same characteristics of 2014. The unexplained part, instead, represents the change in 2004 earnings, with respect to the average earnings between 2004 and 2014, when applying the 2014 coefficients to the 2004 year.

The main evidence is that at the 50th and 75th percentile of the earnings distribution, the unexplained part is more important to explain "year" (generational) differences. In other words compositional differences are much less important in explaining year differentials in earnings, which are instead largely due to different association of the covariates. This means that "year" differences in earnings are not related to the (higher) number of individuals with certain characteristics (ie. better familiar background), but are related to the fact that in 2014 having certain characteristics counts more. Detailed decomposition of year differentials is shown in Figure 3, where the contribution of the main factors (parents' education, sons' education, profession and demographics) to the explained and unexplained part at 25th, 50th, 75th, 90th and 95th percentile distribution of earnings is highlighted, showing an important role played by family education at the 25th percentile as well as a strong role of sons' education in determining the outcomes of individuals.

5.3 Oaxaca-Blinder decomposition of gender differentials in earnings

I then turn to the examination of possible heterogeneity across genders.

The results of OB decomposition at the 25th, 50th, 75th, 90th and 95th percentile distribution of earnings, by gender, are shown in Table 4. Decomposition is expressed as a difference between levels for males "minus" levels for females. Thus a positive (negative) difference means that a given earning value is higher (lower) among males. Table 4 also includes total differences, the explained and the unexplained part and their respective standard errors. Detailed decomposition is shown in Figure 4, where the contribution of the main factors (parents' education, sons' education, profession and demographics) to the explained and unexplained part at 25th, 50th, 75th, 90th and 95th percentile distribution of earnings is highlighted.

Table 4 shows that males have generally higher values of earnings at all quantiles of the earning distribution being gender differentials not the same along the entire distribution. Indeed, although being not negligible and statistically significant at all percentiles, they are increasing and almost doubled as we move towards the 95th percentile.

The second part of Table 4 shows that this is explained both by compositional differences in covariates ("explained part") and by a difference in the impact of covariates on earnings ("unexplained part"). More specifically, the explained part represents the mean increase in earnings in women, respect to the average earnings between women and men, if the women had the same characteristics of men. The coefficient part, instead, represents the change in women's earnings, with respect to the average earnings between women and men, when applying the men's coefficients to the women's characteristics. Especially at high levels of earnings (ie. at the 75th, 90th and 95th percentile) both compositional differences and differences in the impact of covariates on earnings are important, being however the unexplained part more important to explain gender differentials. In other words compositional differences are much less important in explaining gender differentials in earnings, which are instead largely due to different association of earnings to the covariates.

5.4 Oaxaca-Blinder decomposition of territorial differentials in earnings

Finally, I take into account possible heterogeneity across areas of the country. Italy is, indeed, characterized by sharp and widening economic disparities between northern and southern macro areas, with a long history of south-to-north labor migration.

The results of OB decomposition at the 25th, 50th, 75th, 90th and 95th percentile distribution of earnings, by areas of the country, are shown in Table 5. Decomposition is expressed as a difference between levels for Northern regions "minus" levels for Central-Southern regions. Thus a positive (negative) difference means that a given earning value is higher (lower) in the North. Table 5 also includes total differences, the explained and the unexplained part and their respective standard errors. Detailed decomposition is shown in Figure 5, where the contribution of the main factors (parents' education, sons' education, profession and demographics) to the explained and unexplained part at 25th, 50th, 75th, 90th and 95th percentile distribution of earnings is highlighted.

Table 5 shows that Northern regions have generally higher values of earnings at all quantiles of the earning distribution. Differentials by areas are not the same along the entire distribution and although being not negligible and statistically significant at all percentiles, they are increasing as we move towards the 95th percentile. The highest differential is registered in the upper part of the earning distribution such as at the 90th (0.139 points) and 95th percentile (around 0.124 points) and at the lowest level of earnings such as the 25th percentile (0.126).

The second part of Table 5 shows that this is explained both by compositional differences in covariates ("explained part") and by a difference in the impact of covariates on earnings ("unexplained part"). Despite the fact the both compositional differences and differences in the impact of covariates on earnings are important, again, as in the case of gender, at all levels of the earnings distribution, the unexplained part is much more important to explain areas differentials which are, therefore, largely due to different association of earnings to the covariates, while compositional differences are less important.

6 Conclusions

The relationship between family background such as education as well occupational status (or income) of parents and sons' outcomes is probably one of the most explored topics in the academic literature on the intergenerational trasmission. A large literature documents the existence of a positive relationship between income of parents (or their educational and occupational status) and the income of their sons (ie. coming from a well off family increases the chances of keeping the same economic advantage in the future). Despite this large interest, less is known about this relationship at different points of the income distribution.

Indeed, studies of the influence of economic conditions on income typically measure the effect of the former on the conditional mean of the income status variable through regression analysis. Analysis based solely on the mean while offering useful information, misses potentially important information in other parts of the distribution. For instance, it does not check for non-linearities in the relationship between different covariates and income across the full conditional distribution. Moreover, it does not permit analysis of the role of economic conditions at the tails of the income distribution which may be associated with large welfare losses for individuals and high costs for the society.

This study highlights the association between parental background (education of the father and mother) and sons' earning and how it varies along the sons' earnings distribution and aims to explore whether the mechanism generating the intergenerational transmission has changed over the years, gender and geographical areas of the country. I use a distributional method proposed in the recent literature, the recentered influence function approach by Firpo et al. (2009), to estimate the parental education gradient across the full distribution for continuous measures of sons' earnings. Furthermore, I apply Oaxaca-Blinder decompositions at various quantiles of the earnings distributions to explain years, gender, and geographic differentials in earnings in Italy. I use data from the Survey on Household Income and Wealth for Italy on the 2004-2014 time-spam.

I find evidence of higher returns to family education in the upper percentiles of the distribution of son's earnings and that the probability of ending up in high deciles is significantly correlated with the education level of the father. Results show an important heterogeneity in the association of parental education as well as of individual covariates to earnings across time, gender and areas of the country which varies significantly along the earning distribution and accounts for a substantial percentage of the differentials in observed earnings.

References

Becker, G. and Tomes, N. (1979), 'An equilibrium theory of the distribution of income and intergenerational mobility', *Journal of Political Economy* **87**(6), 1153–1189.

Becker, G. and Tomes, N. (1986), 'Human capital and the rise and fall of families', *Journal of Labor Economics* **4**(3), S1–S39.

Benabou, R. (1996), 'Equity and effectiveness in human capital investment: the local connection', *Review of Economic Studies* **63**, 37–64.

Björklund, A. and Jäntti, M. (2012), *Intergenerational Income Mobility and the Role of Family Background*, In B. Nolan, W. Salverda, and T. M. Smeeding, The Oxford Handbook of Economic Inequality, Oxford University Press, Oxford.

Black, S. E. and Devereux, P. J. (1993), *Recent Developments in Intergenerational Mobility*, in O. Ashenfelter and D. Card, Handbook of Labor Economics, Elsevier, Massachusetts.

Blinder, A. S. (1973), 'Wage discrimination: Reduced form and structural estimates', *Journal of Human Resources* **8**, 436–455.

Bowles, S. and Gintis, H. (2002), 'The inheritance of inequality', *Journal of Economic Perspectives* **16**(3), 3–30.

Bradbury, B., Corak, M., Waldfogel, J. and Washbrook, E. (2012), *Recent Developments in Intergenerational Mobility*, In J. Ermisch, M. Jantti and T. Smeeding, Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage, Russell Sage Foundation, New York.

Bratberg, E., Nilsen, O. A. and Vaage, K. (2005), 'Intergenerational earnings mobility in norway: Levels and trends', *The Scandinavian Journal of Economics* **107**(3), 419–435.

Bratsberg, B., Røed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T. and Österbacka, E. (2007), 'Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons', *Economic Journal* **117**, C72–C92.

Bukodi, E. and Goldthorpe, J. (2012), 'Decomposing "social origins": the effects of parents' class, status, and education on the educational attainment of their children', *European Sociological Review* **29**, 1024–39.

Chevalier, A., Harmon, C., O' Sullivan, V. and Walker, I. (2013), 'The impact of parental income and education on the schooling of their children', *IZA J Labor Econ* **2**(8).

Corak, M. (2013), 'Income inequality, equality of opportunity, and intergenerational mobility', *Journal of Economic Perspectives* **27**(3), 79–102.

Corak, M. and Heisz, S. (1999), 'The intergenerational earnings and income mobility of canadian men: evidence from longitudinal income tax data', *Journal of Human Resources* **34**(3), 504–33.

Cunha, F. and Heckman, J. (2007), 'The technology of skill formation', *American Economic Review* **97**(2), 31–47.

Cunha, F., Heckman, J., Lochner, L. and Masterov, D. (2006), *Interpreting the evidence on life cycle skill formation*, In E. Hanushek and F. Welch, Handbook of the Economics of Education, North Holland, Amsterdam.

Dickson, M., Gregg, P. and Robinson, H. (2016), 'Early, late or never? when does parental education impact child outcomes?', *The Economic Journal* **126**(596), F184–F231.

Eide, E. R. and Showalter, M. H. (1999), 'Factors affecting the transmission of earnings across generations: A quantile regression approach', *The Journal of Human Resources* **34**(2), 253–267.

Erikson, R. and Goldthorpe, J. H. (1992), *The Constant Flux: a Study of Class Mobility in Industrial Societies*, Oxford University Press, Oxford.

Ermisch, J. and Del Bono, E. (2012a), *Inequality in achievements during adolescence*, in J. Ermisch, M. Jantti and T. Smeeding, From Parents to Children: The Intergenerational Transmission of Advantage, Russell Sage Foundation, New York.

Ermisch, J., Jantti, M. and Smeeding, T. (2012a), *Socioeconomic gradients in children's outcomes*, in J. Ermisch, M. Jantti and T. Smeeding, From Parents to Children: The Intergenerational Transmission of Advantage, Russell Sage Foundation, New York.

Ermisch, J., Jantti, M. and Smeeding, T. (2012b), *From Parents to Children: The Intergenerational Transmission of Advantage*, Russell Sage Foundation, New York.

Firpo, S., Fortin, N. M. and Lemieux, T. (2009), 'Unconditional quantile regressions', *Econometrica* **77**, 953–973.

Fortin, N., Lemieux, T. and Firpo, S. (2011), *Decomposition methods in economics*, in O. Ashenfelter and D. Card, Part A of Handbook of Labor Economics, Elsevier, Amsterdam.

Fournier, J. and Koske, I. (2012), 'The determinants of earnings inequality – evidence from quantile regressions', *OECD Journal: Economic Studies* pp. 7–36.

Ganzeboom, H. and Treiman, D. (2007), *Ascription and Achievement in Comparative Perspective*, University Working Group on Social Inequality, Russell-Sage, Los Angeles.

Granovetter, M. (1995), *Afterword*, in M. Granovetter, Getting a Job: a Study of Contacts and Careers, Chicago University Press, Chicago.

Grawe, N. D. (2004), 'Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints', *Journal of Human Resources* **39**, 813–27.

Gregg, P., Macmillan, L. and Vittori, V. (2018), 'Intergenerational income mobility: access to top jobs, the low-pay no-pay cycle and the role of education in a common framework', *Journal of Population Economics* **32**, 501–528.

Jackson, M. (2013), *Determined to Succeed? Performance versus Choice in Educational Attainment*, Stanford University Press, Redwood City, CA.

Jann, B. (2008), 'The blinder-oaxaca decomposition for linear regression models', *The Stata Journal* **8**(4), 453–79.

Jerrim, J. (2017), 'The link between family background and later lifetime income: How does the uk compare with other countries?', *Fiscal studies* **38**(1), 49–79.

Koenker, R. W. and Bassett, G. (1978), 'Regression quantiles', *Econometrica* **46**, 33–50.

Lampard, R. (2007), 'Is social mobility an echo of educational mobility? parents' educations and occupations and their children's occupational attainment', *Sociological Research Online* **12**(5), 44–66.

Monti, A. (1991), 'The study of gini concentration ratio by means of influence function', *Statistica* **51**, 561–577.

Naticchioni, P., Raitano, M. and Vittori, C. (2016), 'La meglio gioventu: earnings gaps across generations and skills in italy', *Economia Politica* **33**, 233–264.

Oaxaca, R. (1973), 'Male-female wage differentials in urban labor markets', *International Economic Review* **14**, 693–709.

Raitano, M., Vittori, C. and Vona, F. (2016), 'The effect of parental background along the sons' earnings distribution: does one pattern fit for all?', *Applied Economic Letters* **23**(15), 1073–1078.

Raitano, M. and Vona, F. (2015*a*), 'Direct and indirect influences of parental background on children's earnings: a comparison across countries and genders', *The Manchester School* **83**(4).

Raitano, M. and Vona, F. (2015*b*), 'Measuring the link between intergenerational occupational mobility and earnings: evidence from eight european countries', *Journal of Economic Inequality* **13**, 83–102.

Schnitzlein, D. D. (2015), 'A new look at intergenerational mobility in germany compared to u.s.', *Review of Income and Wealth* **62**(4), 650–657.

Schutz, G., Ursprung, H. and Woessmann, L. (2008), 'Education policy and equality of opportunity', *Kyklos* **61**(2), 279–308.

Solon, G. (2004), 'A model of intergenerational mobility variation over time and place'.

Tables

Table 1: Descriptive statistics - Whole sample

Variable	Obs	Mean	Std. Dev.	Min	Max
Wage (log)	18,356	9.684	0.691	0	13.815
Father illiterate	18,356	0.123	0.328	0	1
Father primary	18,356	0.499	0.500	0	1
Father secondary	18,356	0.334	0.471	0	1
Father university	18,356	0.042	0.202	0	1
Mother illiterate	18,356	0.163	0.369	0	1
Mother primary	18,356	0.530	0.499	0	1
Mother secondary	18,356	0.286	0.451	0	1
Mother university	18,356	0.020	0.141	0	1
Father schooling	18,356	6.532	4.180	0	18
Mother schooling	18,356	5.827	3.890	0	18
Gender	18,356	0.681	0.465	0	1
Age	18,356	47.008	9.982	19	92
Married	18,356	0.694	0.460	0	1
Single	18,356	0.162	0.369	0	1
Divorced	18,356	0.114	0.318	0	1
Widow	18,356	0.028	0.165	0	1
North	18,356	0.506	0.499	0	1
Centre	18,356	0.208	0.406	0	1
South	18,356	0.285	0.451	0	1
Illiterate	18,356	0.003	0.058	0	1
Primary	18,356	0.055	0.229	0	1
Secondary	18,356	0.763	0.424	0	1
University	18,356	0.176	0.381	0	1
Blue collar	18,356	0.330	0.470	0	1
Office worker	18,356	0.329	0.470	0	1
Manager	18,356	0.085	0.280	0	1
Member of profession	18,356	0.101	0.301	0	1
Self-employed (other)	18,356	0.122	0.328	0	1

Own calculation

Table 2: Results - OLS and RIF regressions - Years 2004-2014

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	RIF	RIF	RIF	RIF	RIF
VARIABLES		Q25	Q50	Q75	Q90	Q95
Father primary	0.063***	0.099***	0.028**	0.026*	0.009	0.012
	(0.018)	(0.022)	(0.013)	(0.015)	(0.022)	(0.026)
Father secondary	0.065***	0.113***	0.038**	0.045**	0.051*	0.094***
	(0.021)	(0.024)	(0.015)	(0.019)	(0.029)	(0.036)
Father university	0.178***	0.092***	0.062***	0.152***	0.396***	0.522***
	(0.032)	(0.034)	(0.022)	(0.032)	(0.067)	(0.100)
Mother primary	0.012	-0.002	0.024**	0.020	0.019	0.008
	(0.016)	(0.019)	(0.012)	(0.014)	(0.021)	(0.025)
Mother secondary	0.027	-0.002	0.036**	0.053***	0.091***	0.073*
	(0.020)	(0.022)	(0.015)	(0.019)	(0.030)	(0.038)
Mother university	-0.024	-0.051	-0.008	-0.015	0.017	-0.032
	(0.039)	(0.042)	(0.028)	(0.039)	(0.084)	(0.127)
Gender	0.334***	0.343***	0.256***	0.249***	0.203***	0.213***
	(0.010)	(0.012)	(0.007)	(0.009)	(0.015)	(0.019)
Age	0.001	-0.000	0.003***	0.004***	0.008***	0.010***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Married	0.083***	0.054***	0.072***	0.091***	0.128***	0.113***
	(0.013)	(0.014)	(0.009)	(0.012)	(0.018)	(0.024)
Divorced	0.117***	0.085***	0.088***	0.082***	0.096***	0.116***
	(0.017)	(0.019)	(0.013)	(0.016)	(0.026)	(0.035)
Widow	0.042	0.001	0.026	0.051**	0.076**	0.103**
	(0.029)	(0.036)	(0.020)	(0.023)	(0.036)	(0.049)
Centre	0.135***	0.125***	0.090***	0.106***	0.105***	0.124***
	(0.013)	(0.014)	(0.009)	(0.012)	(0.020)	(0.028)
North	0.201***	0.196***	0.133***	0.153***	0.182***	0.145***
	(0.011)	(0.012)	(0.008)	(0.010)	(0.017)	(0.022)
Primary	0.083	0.066	0.004	-0.094*	-0.005	-0.027
	(0.077)	(0.099)	(0.047)	(0.048)	(0.035)	(0.039)
Secondary	0.330***	0.320***	0.150***	0.016	0.083**	0.088**
	(0.076)	(0.097)	(0.046)	(0.047)	(0.033)	(0.038)
University	0.446***	0.339***	0.232***	0.194***	0.466***	0.578***
	(0.077)	(0.098)	(0.047)	(0.049)	(0.044)	(0.056)
Office worker	0.363***	0.428***	0.315***	0.251***	0.035***	-0.091***
	(0.011)	(0.012)	(0.008)	(0.010)	(0.013)	(0.014)
Manager	0.747***	0.483***	0.504***	0.837***	1.312***	1.057***
	(0.018)	(0.014)	(0.010)	(0.018)	(0.047)	(0.066)
Member of profession	0.493***	0.185***	0.279***	0.533***	0.933***	1.040***
	(0.016)	(0.019)	(0.013)	(0.019)	(0.041)	(0.060)
Self-employed (other)	0.142***	-0.035*	0.123***	0.245***	0.305***	0.234***
	(0.014)	(0.019)	(0.012)	(0.015)	(0.025)	(0.031)
Constant	8.525***	8.463***	8.808***	9.079***	9.150***	9.391***
	(0.079)	(0.101)	(0.049)	(0.053)	(0.056)	(0.074)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,352	18,352	18,352	18,352	18,352	18,352
R-squared	0.250	0.176	0.240	0.263	0.252	0.168

Standard Errors in parenthesis; ***, **, * significance at 1%, 5% and 10%

Table 3: Oaxaca-Blinder Decomposition of year differentials in earnings

	Q25	Q50	Q75	Q90	Q95
Mean prediction 2014	9.443	9.754	10.028	10.344	10.642
Mean prediction 2004	9.438	9.667	9.919	10.368	10.635
$\Delta_{2014-2004}$	0.004***	0.086***	0.108***	-0.023	0.006
	(0.020)	(0.011)	(0.014)	(0.030)	(0.042)
Explained	-0.019*	-0.007	-0.002	0.024	0.052**
	(0.011)	(0.007)	(0.009)	(0.018)	(0.022)
Unexplained	0.023	0.093***	0.111***	-0.047	-0.045
	(0.026)	(0.014)	(0.021)	(0.030)	(0.272)

Standard Errors in parenthesis; ***, **, * significance at 1% 5% and 10%

Table 4: Oaxaca-Blinder Decomposition of gender differentials in earnings

	Q25	Q50	Q75	Q90	Q95
Mean prediction Men	9.576	9.838	10.104	10.496	10.726
Mean prediction Women	9.220	9.629	9.849	10.107	10.341
$\Delta_{Men-Women}$	0.355***	0.208***	0.255***	0.388***	0.384***
	(0.014)	(0.007)	(0.008)	(0.016)	(0.024)
Explained	-0.068***	-0.027***	0.016**	0.098***	0.104***
	(0.006)	(0.004)	(0.006)	(0.019)	(0.012)
Unexplained	0.424***	0.236***	0.239***	0.289***	0.279***
	(0.018)	(0.014)	(0.013)	(0.020)	(0.024)

Standard Errors in parenthesis; ***, **, * significance at 1% 5% and 10%

Table 5: Oaxaca-Blinder Decomposition of territorial differentials in earnings

	Q25	Q50	Q75	Q90	Q95
Mean prediction North	9.547	9.807	10.053	10.444	10.703
Mean prediction South	9.421	9.713	9.959	10.305	10.5579
$\Delta_{North-South}$	0.126***	0.093***	0.094***	0.139***	0.124***
	(0.010)	(0.007)	(0.009)	(0.019)	(0.026)
Explained	0.012***	0.004	0.008*	0.019**	0.023*
	(0.004)	(0.003)	(0.004)	(0.008)	(0.009)
Unexplained	0.114***	0.089***	0.086***	0.119***	0.101***
	(0.016)	(0.009)	(0.011)	(0.028)	(0.033)

Standard Errors in parenthesis; ***, **, * significance at 1% 5% and 10%

Figures

Figure 1: Coefficient of RIF regression - Father education - Years 2004-2014

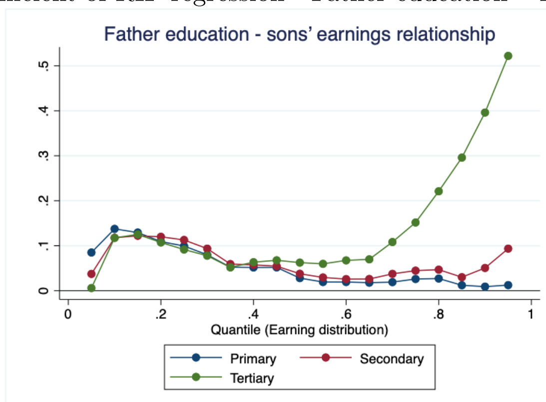


Figure 2: Coefficient of RIF regression - Mother education - Years 2004-2014

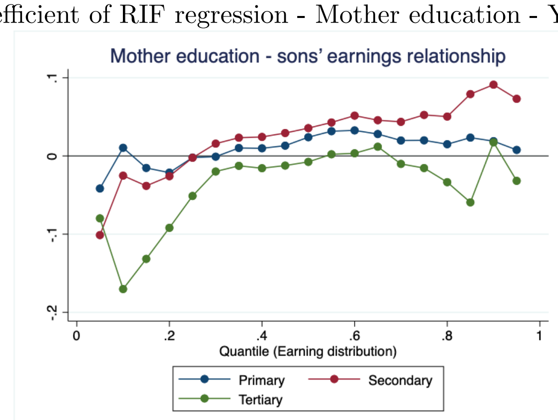


Figure 3: Detailed O-B decomposition of year differentials in earnings - Year 2014 versus 2004

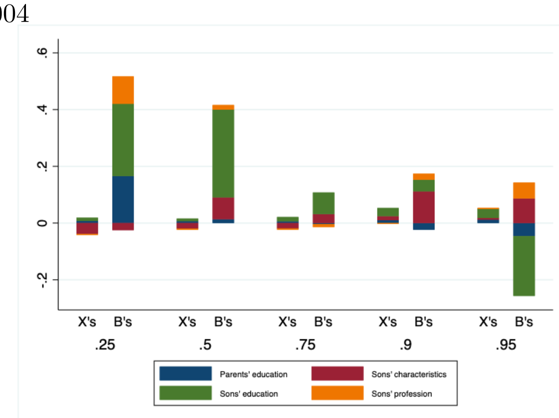


Figure 4: Detailed O-B decomposition of gender differentials in earnings - Years 2004-2014

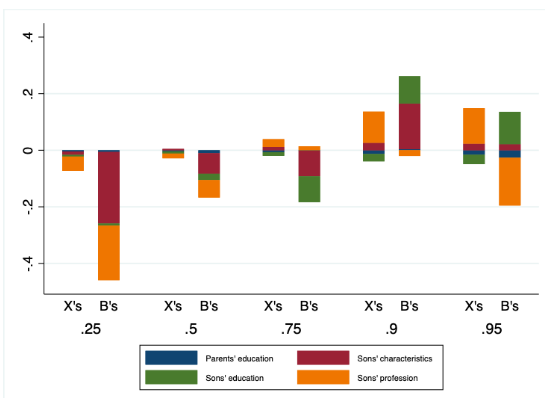
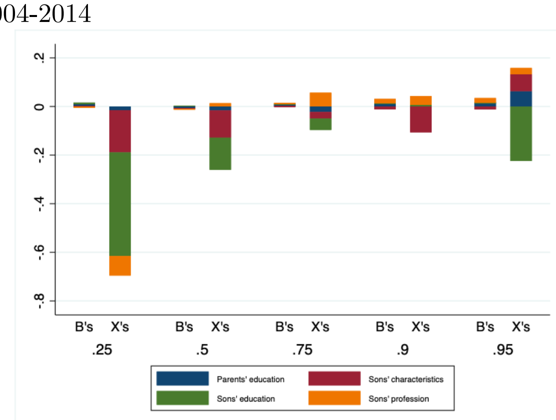


Figure 5: Detailed O-B decomposition of geographical area differentials in earnings - Years 2004-2014



Appendix: Tables and Figures

Table A1: Results - OLS and RIF regressions - Years 2004-2014

VARIABLES	(1) OLS	(2) RIF Q25	(3) RIF Q50	(4) RIF Q75	(5) RIF Q90	(6) RIF Q95
Father schooling	0.006***	0.005***	0.003**	0.007***	0.018***	0.023***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)
Mother schooling	0.000	-0.001	0.001	0.003*	0.004	0.004
	(0.002)	(0.002)	(0.001)	(0.002)	(0.003)	(0.004)
Gender	0.333***	0.343***	0.255***	0.249***	0.203***	0.213***
	(0.010)	(0.012)	(0.007)	(0.009)	(0.015)	(0.019)
Age	0.001	-0.000	0.003***	0.004***	0.008***	0.010***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Married	0.085***	0.056***	0.073***	0.092***	0.127***	0.111***
	(0.013)	(0.014)	(0.009)	(0.012)	(0.018)	(0.024)
Divorced	0.119***	0.087***	0.090***	0.082***	0.095***	0.114***
	(0.017)	(0.019)	(0.013)	(0.016)	(0.026)	(0.035)
Widow	0.040	-0.002	0.025	0.051**	0.077**	0.105**
	(0.029)	(0.036)	(0.020)	(0.023)	(0.037)	(0.049)
Centre	0.138***	0.132***	0.093***	0.106***	0.097***	0.113***
	(0.013)	(0.014)	(0.009)	(0.012)	(0.020)	(0.028)
North	0.205***	0.204***	0.138***	0.153***	0.175***	0.134***
	(0.010)	(0.012)	(0.008)	(0.010)	(0.016)	(0.022)
Primary	0.099	0.100	0.016	-0.096**	-0.041	-0.074*
	(0.077)	(0.099)	(0.047)	(0.048)	(0.035)	(0.039)
Secondary	0.355***	0.373***	0.170***	0.015	0.031	0.019
	(0.075)	(0.096)	(0.046)	(0.047)	(0.033)	(0.038)
University	0.469***	0.382***	0.248***	0.186***	0.418***	0.516***
	(0.077)	(0.097)	(0.047)	(0.049)	(0.045)	(0.057)
Office worker	0.362***	0.434***	0.318***	0.249***	0.023*	-0.104***
	(0.011)	(0.012)	(0.008)	(0.010)	(0.013)	(0.014)
Manager	0.747***	0.487***	0.507***	0.836***	1.306***	1.050***
	(0.018)	(0.014)	(0.010)	(0.018)	(0.047)	(0.066)
Member of profession	0.495***	0.187***	0.281***	0.531***	0.930***	1.036***
	(0.016)	(0.019)	(0.013)	(0.019)	(0.041)	(0.060)
Self-employed (other)	0.143***	-0.032*	0.124***	0.244***	0.301***	0.229***
	(0.014)	(0.019)	(0.012)	(0.015)	(0.025)	(0.031)
Constant	8.524***	8.475***	8.814***	9.076***	9.139***	9.379***
	(0.079)	(0.101)	(0.049)	(0.052)	(0.056)	(0.073)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,352	18,352	18,352	18,352	18,352	18,352
R-squared	0.249	0.175	0.239	0.263	0.251	0.166

Standard Errors in parenthesis; ***, **, * significance at 1%, 5% and 10%

Table A2: Results - OLS and RIF regressions - Year 2004

VARIABLES	(1) OLS	(2) RIF Q25	(3) RIF Q50	(4) RIF Q75	(5) RIF Q90	(6) RIF Q95
Father primary	0.073*	0.034	-0.006	-0.027	-0.015	0.038
	(0.042)	(0.041)	(0.028)	(0.034)	(0.063)	(0.065)
Father secondary	0.084*	0.054	0.009	0.022	0.015	0.051
	(0.049)	(0.045)	(0.032)	(0.040)	(0.076)	(0.094)
Father university	0.223***	0.038	-0.003	0.147**	0.483***	0.553*
	(0.079)	(0.065)	(0.049)	(0.070)	(0.179)	(0.295)
Mother primary	0.020	-0.016	0.045*	0.063**	0.065	0.015
	(0.038)	(0.036)	(0.025)	(0.030)	(0.057)	(0.065)
Mother secondary	0.031	-0.025	0.069**	0.055	0.116	0.217**
	(0.047)	(0.042)	(0.031)	(0.039)	(0.077)	(0.106)
Mother university	0.100	0.005	0.077	0.041	0.009	0.495
	(0.098)	(0.077)	(0.059)	(0.084)	(0.201)	(0.369)
Gender	0.369***	0.290***	0.233***	0.193***	0.208***	0.178***
	(0.027)	(0.025)	(0.016)	(0.020)	(0.041)	(0.060)
Age	0.000	0.001	0.002**	0.003***	0.007***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)
Married	0.096***	0.078**	0.081***	0.130***	0.118**	-0.014
	(0.032)	(0.031)	(0.020)	(0.023)	(0.049)	(0.079)
Divorced	0.162***	0.138***	0.143***	0.118***	0.033	-0.092
	(0.044)	(0.040)	(0.028)	(0.035)	(0.069)	(0.110)
Widow	0.161**	0.085	0.081*	0.166***	0.082	-0.006
	(0.069)	(0.068)	(0.042)	(0.051)	(0.104)	(0.168)
Centre	0.148***	0.145***	0.088***	0.078***	0.078	0.173**
	(0.031)	(0.028)	(0.020)	(0.027)	(0.052)	(0.076)
North	0.206***	0.180***	0.119***	0.109***	0.180***	0.222***
	(0.026)	(0.024)	(0.017)	(0.022)	(0.044)	(0.061)
Primary	0.123	-0.115	-0.156	-0.129	-0.194	0.023
	(0.145)	(0.156)	(0.098)	(0.088)	(0.143)	(0.109)
Secondary	0.295**	0.081	-0.075	0.005	-0.042	0.120
	(0.143)	(0.153)	(0.096)	(0.087)	(0.144)	(0.103)
University	0.486***	0.156	0.020	0.197**	0.416**	0.826***
	(0.147)	(0.155)	(0.098)	(0.093)	(0.165)	(0.166)
Office worker	0.290***	0.346***	0.263***	0.168***	0.031	-0.142***
	(0.028)	(0.026)	(0.018)	(0.022)	(0.034)	(0.041)
Manager	0.634***	0.378***	0.435***	0.701***	1.172***	0.781***
	(0.047)	(0.030)	(0.023)	(0.044)	(0.118)	(0.176)
Member of profession	0.552***	0.185***	0.262***	0.534***	1.193***	1.503***
	(0.041)	(0.037)	(0.027)	(0.041)	(0.105)	(0.179)
Self-employed (other)	0.171***	0.004	0.117***	0.257***	0.421***	0.415***
	(0.034)	(0.035)	(0.023)	(0.031)	(0.064)	(0.093)
Constant	8.520***	8.744***	9.126***	9.169***	9.336***	9.356***
	(0.155)	(0.163)	(0.103)	(0.100)	(0.179)	(0.201)
Observations	3,074	3,074	3,074	3,074	3,074	3,074
R-squared	0.248	0.176	0.235	0.267	0.257	0.187

Standard Errors in parenthesis; ***, **, * significance at 1%, 5% and 10%

Table A3: Results - OLS and RIF regressions - Year 2014

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	RIF	RIF	RIF	RIF	RIF
VARIABLES		Q25	Q50	Q75	Q90	Q95
Father primary	0.072	0.153*	0.024	0.042	0.052	0.092
	(0.046)	(0.080)	(0.035)	(0.035)	(0.061)	(0.080)
Father secondary	0.133**	0.201**	0.062	0.096**	0.095	0.257**
	(0.053)	(0.089)	(0.041)	(0.043)	(0.081)	(0.109)
Father university	0.292***	0.220*	0.108*	0.205***	0.495***	0.794***
	(0.079)	(0.121)	(0.056)	(0.067)	(0.174)	(0.270)
Mother primary	-0.013	0.047	0.022	-0.021	-0.048	-0.075
	(0.042)	(0.069)	(0.032)	(0.033)	(0.061)	(0.083)
Mother secondary	-0.008	0.041	0.042	-0.002	0.038	-0.079
	(0.050)	(0.079)	(0.038)	(0.041)	(0.084)	(0.117)
Mother university	-0.150*	-0.054	-0.053	-0.102	-0.073	-0.434
	(0.088)	(0.141)	(0.067)	(0.074)	(0.205)	(0.293)
Gender	0.344***	0.456***	0.265***	0.251***	0.272***	0.303***
	(0.023)	(0.037)	(0.017)	(0.018)	(0.040)	(0.050)
Age	0.000	-0.001	0.003***	0.003***	0.005**	0.007**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)
Married	0.090***	0.030	0.070***	0.120***	0.242***	0.169***
	(0.028)	(0.043)	(0.022)	(0.023)	(0.044)	(0.058)
Divorced	0.094**	0.010	0.037	0.054*	0.244***	0.303***
	(0.038)	(0.061)	(0.029)	(0.030)	(0.064)	(0.090)
Widow	0.049	-0.022	0.038	0.094**	0.272**	0.268*
	(0.071)	(0.118)	(0.049)	(0.048)	(0.117)	(0.154)
Centre	0.104***	0.070	0.070***	0.103***	0.109**	0.116
	(0.031)	(0.049)	(0.023)	(0.025)	(0.054)	(0.074)
North	0.191***	0.218***	0.139***	0.151***	0.216***	0.179***
	(0.025)	(0.040)	(0.019)	(0.020)	(0.044)	(0.059)
Primary	-0.068	-0.127	0.073	-0.018	-0.011	-0.087
	(0.193)	(0.385)	(0.053)	(0.050)	(0.089)	(0.109)
Secondary	0.294	0.370	0.252***	0.094**	0.006	-0.041
	(0.189)	(0.377)	(0.046)	(0.047)	(0.065)	(0.070)
University	0.358*	0.361	0.292***	0.216***	0.391***	0.365***
	(0.191)	(0.379)	(0.051)	(0.054)	(0.096)	(0.117)
Office worker	0.405***	0.565***	0.334***	0.244***	0.115***	-0.027
	(0.027)	(0.042)	(0.020)	(0.022)	(0.036)	(0.038)
Manager	0.833***	0.662***	0.527***	0.778***	1.664***	1.559***
	(0.045)	(0.046)	(0.026)	(0.036)	(0.132)	(0.189)
Member of profession	0.427***	0.267***	0.237***	0.327***	0.904***	1.207***
	(0.039)	(0.062)	(0.031)	(0.038)	(0.103)	(0.156)
Self-employed (other)	0.099***	-0.055	0.015	0.079***	0.288***	0.307***
	(0.037)	(0.067)	(0.029)	(0.030)	(0.067)	(0.088)
Constant	8.607***	8.277***	8.803***	9.187***	9.137***	9.426***
	(0.195)	(0.385)	(0.061)	(0.064)	(0.127)	(0.160)
Observations	3,107	3,107	3,107	3,107	3,107	3,107
R-squared	0.274	0.168	0.252	0.286	0.266	0.189

Standard Errors in parenthesis; ***, **, * significance at 1%, 5% and 10%

Figure A1: Coefficient of RIF regression - Father years of schooling - Years 2004-2014

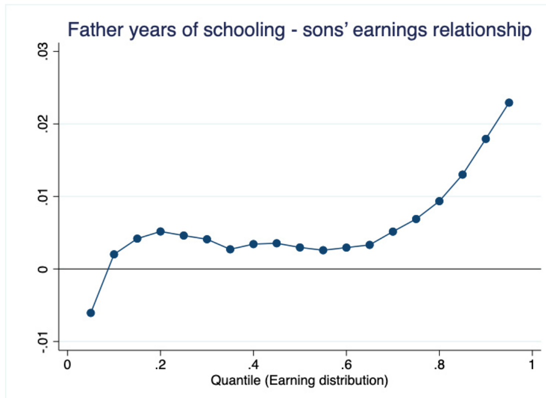


Figure A2: Coefficient of RIF regression - Mother years of schooling - Years 2004-2014

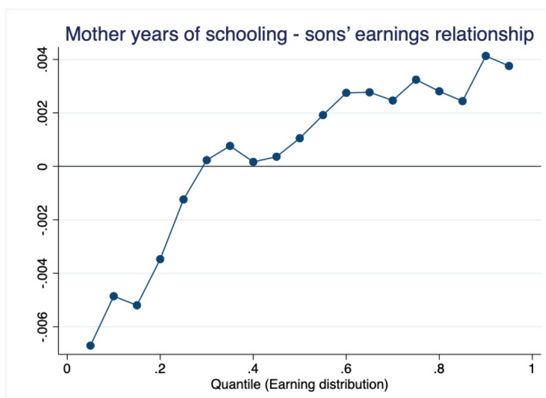


Figure A3: Coefficient of RIF regression - Father education - Year 2004

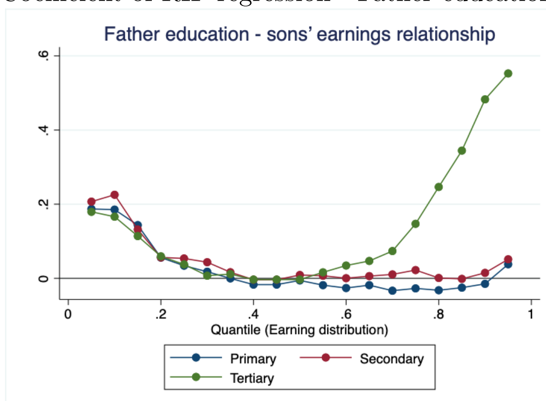


Figure A4: Coefficient of RIF regression - Mother education - Year 2004

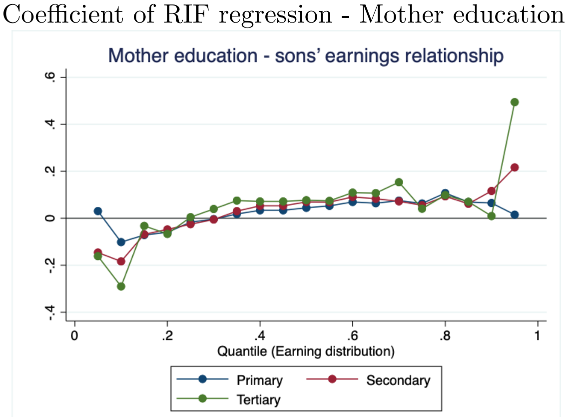


Figure A5: Coefficient of RIF regression - Father education - Year 2014

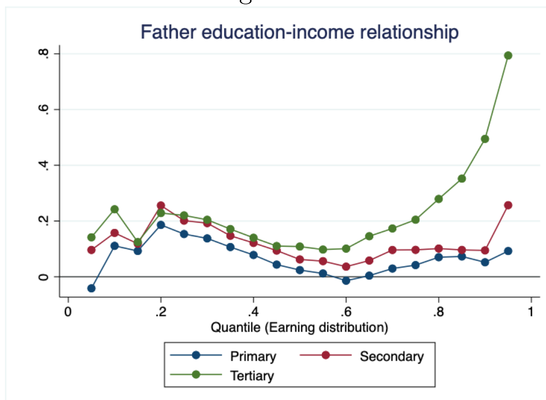


Figure A6: Coefficient of RIF regression - Mother education - Year 2014