Intergenerational Earnings Inequality in Italy: New Evidences and Main Mechanisms

Teresa Barbieri, Francesco Bloise and Michele Raitano[1]

Abstract

This chapter provides new and detailed estimates of intergenerational earnings inequality in Italy and sheds light on mechanisms behind association of gross and net earnings between fathers and sons.

Being not available panel data following subsequent generations in Italy, we make use of a recently built dataset that merges information provided by IT-SILC 2005 (i.e., the Italian component of EU-SILC 2005) with detailed information about the whole working life of those interviewed in IT-SILC recorded in the administrative archives managed by the Italian National Social Security Institute (INPS). This dataset allows us to rely on the two-sample two-stage least squares method (TSTSLS) to predict father earnings and, then, compute point in time intergenerational elasticities (IGE) and imputed rank to rank slopes. Furthermore, the characteristics of the dataset allow us to extend point in time estimates considering, for both sons and "pseudo-fathers", average earnings in a 5-year period and observing sons at various ages, thus assessing the robustness of our estimates to attenuation and life cycle biases.

Confirming previous evidence (Mocetti 2007; Piraino 2007), we find that Italy is characterized by a relatively high earnings inequality in cross country comparison – the size of the estimated β is usually over 0.40 – and the size of the intergenerational association increases when older sons and multi-annual averages are considered.

We also investigate mechanisms behind this association both: i) including a set of possible mediating factors of the parental influence (e.g., sons' education, occupation, labour market experience) among the control variables when regressing sons' earnings on fathers' earnings and ii) following the sequential decomposition approach suggested by Blanden et al. (2007). Results show that a limited share of the intergenerational association is attributable to sons' educational and occupational attainment, while the largest part of the association is mediated by sons' employability along the career, i.e., by their effective experience since the entry in the labour market.

[1] Sapienza University of Rome

Introduction

In the last few decades, a growing body of international literature has focused on intergenerational transmission of social and economic advantages (and disadvantages). Economists have focused their attention on measuring the degree of income persistence across two generations and, more specifically, on the estimation of the intergenerational elasticity coefficient β that captures how much of the income difference between two parents still is preserved between their children (see Blanden, 2013). Due to constraints that effect women participation in the labour market, literature on intergenerational mobility usually focuses on the association between fathers and sons' earnings.

Even if reliable estimates of the intergenerational earnings elasticity (IGE) are available only for few nations, a generally accepted ranking has emerged from cross-country empirical studies on economic mobility as concerns developed countries (Solon, 2002; Corak, 2013; D'Addio, 2007; Bjorklund & Jantti, 2009; Blanden, 2013): Nordic European countries emerge as the most mobile, while the US, the UK and Southern European countries are reported to be the most unequal. According to the few available estimates (Mocetti, 2007, Piraino, 2007), Italy belongs to the low-mobility group.

Actually, due to the unavailability of datasets jointly recording information on children and parents' earnings or income, Italy has received a limited attention in the intergenerational mobility literature. Nonetheless, the IGE in Italy has been estimated in recent years by means of the two-sample two-stage least squares (TSTSLS) method, which allows researchers to overcome the lack of data regarding actual fathers' incomes (Mocetti 2007; Piraino, 2007). More specifically, when long panel data recording income information for both generations of parents and children observed at middle ages are not available, the TSTSLS empirical approach exploits two independent samples of children and pseudo-fathers and some child-reported retrospective information about fathers to obtain a prediction of fathers' earnings in the first stage and the IGE in the second. Mocetti (2007) and Piraino

(2007), following this approach by using various cross-sections of the Survey of Household Income and Wealth (SHIW) carried out by the Bank of Italy, computed point in time measures (i.e. concerning a single year) of intergenerational net earnings elasticity. They obtained estimated values amounting, respectively, to 0.50 and 0.44.

An alternative measure of mobility, i.e. the rank to rank slope, has recently proved to be more robust across samples and specifications (Dahl & Delaire 2008; Chetty et al. 2014) and with respect to both life-cycle and attenuation bias (Gregg et al. 2014) than the IGE[2]. However, to the best of our knowledge, no studies have so far estimated rank to rank slopes for Italy

Therefore, the aim of this chapter is to provide new estimates of the earnings mobility in Italy by means of a proper dataset, that, compared to the SHIW, allows us to observe children and pseudo-fathers earnings for more than one-year. To this aim, we use the AD-SILC dataset, that has been developed merging information provided in IT-SILC (2005) – where a retrospective section on parental characteristics is recorded – with information on the whole working history since the entry in the labour market of all individuals interviewed in IT-SILC provided by the administrative archives managed by the Italian National Social Security Institute (INPS). Thus, we can exploit the longitudinal nature of these data to apply the TSTSLS procedure, computing multi-year earnings values for both generations. INPS data record earnings gross of taxes and contribution paid by the worker, thus they allow use to compute indexes of intergenerational inequality of gross earnings, thus computing the size of income persistence produced by the labour market, before the redistributive effect of taxes and transfers. In order to compare our results with those provided by Mocetti (2007) and Piraino (2007), who focused on net

[2] More specifically, intergenerational elasticity capture both the re-ranking across generations and the differences in the amount of inequality within each generation (due to changes in income distribution across generations). Thus, the IGE is very sensitive to changes in inequality and it may not capture changes in positional income mobility only, but also the evolution of cross-sectional earnings inequality (Lefranc, 2011). This may be problematic when we want to compare the degree of intergenerational mobility across countries with different level of cross-sectional inequality.

earnings, we also reconstructed net earnings for both generations to estimate the intergenerational elasticity of net earnings.

However, the aim of this chapter does not limit to compute summary measures of intergenerational inequality (as the IGE or the rank to rank slope). Indeed, we also aim at assessing which are the main mechanisms behind the correlation between parents' and children' earnings. In particular, we aim at analysing whether the bulk of intergenerational inequality is explained by educational and occupational attainments of children coming from different backgrounds or a significant association between parents' and children' earnings emerges also controlling for these children's outcomes.

To this end, we first compare our baseline results adding in our estimates of the link between parents' and children's earnings further variables, that can mediate the relationship between fathers' circumstances and children earnings, namely children's, education, contractual arrangement and experience in the labour market.

Moreover, we apply a sequential decomposition approach (Blanden et al. 2007; Hirvonen, 2010; Buchner at al., 2012; Macmillan, 2013; Blanden et al, 2014) to disentangle the share of the IGE explained by various children's characteristics that might be affected by parental circumstances. According to the Becker and Tomes theoretical framework (1979 and 1986), when capitals markets are not perfect and public investment in education does not fully compensate for them, investment in children human capital by parents coming from disadvantaged background are limited, since parents face liquidity constraints. Literature on intergenerational mobility based on this theoretical framework recognizes education as the main transmission mechanism of persistence across generations: children coming from a more disadvantaged background receive a lower level of investments education and, consequently, later in life they will have less job opportunity and lower earnings.

However, this "human capital view" has been challenged by some scholars that recognize the importance of a "direct" effect of family background on earnings, not mediated by "formal" educational attainments. (e.g., Breen & Goldthorpe, 2001;

Goldthorpe & Jackson, 2008; Franzini & Raitano, 2009; Franzini et al, 2013; Hudson and Sessions 2011, Raitano & Vona, 2015).

The decomposition approach, as mentioned, measures to which extent the IGE is explained by sons' characteristics (e.g. education or occupation). The explained part is measured accounting for both the relationship between parent's earnings and children's characteristics and the return to those characteristics in the labour market. Thus, the intergenerational elasticity can be decomposed in two parts: the indirect effect of parental background acting through children's endowment of different characteristics and a residual direct effect not explained by these characteristics.

The reminder of this chapter is structured as follows. Section 1.1 presents the main findings of the empirical literature on intergenerational mobility, focusing on the differences between the empirical approaches that have been proposed to tackle with the issue of intergenerational inequality when information of parents' earnings are not available. Section 1.2 describes the dataset and the sample selection used to run our estimates. Section 1.3 describes the methodology and the empirical strategy that we follow in this chapter. Section 1.4 presents results of the TSTSLS estimates of the IGE comparing results obtained observing parents and children for different time spans. Section 1.5 presents results of the estimates of imputed rank-rank slopes comparing again results obtained observing parents and children for different time spans. Section 1.6 shows how IGE and rank to rank slopes change when we add some children outcomes among the control variables, while Section 1.7 shows results of the decomposition of the IGE for Italy into different mediating variables that may account for the transmission of earnings between parents and children. Section 1.8 concludes, summarizing our main results.

1.1. Intergenerational earnings mobility: OLS estimates

Over the last two decades economists have broadly analysed to what extent economic advantages are transmitted from one generation to the next. The ideal way to evaluate the degree of intergenerational economic mobility is to use permanent earnings (or permanent incomes, when also information on not labour incomes are available) as a measure of economic welfare of individuals and to estimate the following equation:

$$y_i^s = \alpha + \beta \ y_i^f + \varepsilon_i \tag{1}$$

where y_i^s and y_i^f are respectively the logarithm of permanent children's' and parents' earnings and β is the IGE[3]. According to this measure of economic association between generations, a country is completely mobile when the estimated β equals 0, while the higher the earnings elasticity is, the lower the degree of economic mobility across generations will be.

Unfortunately, several methodological issues arise when trying to estimate equation 1. Firstly, also the few datasets covering two generations usually report short-term rather than permanent measures of earnings. This implies that, under classical measurement errors assumptions, estimated elasticities obtained using yearly instead of permanent fathers' earnings are likely to be downward biased due to the so-called attenuation bias (Solon, 1992; Zimmerman, 1992). A usual way to reduce this kind of bias is to average fathers' earnings over a period as large as possible. The greater the number of years available when averaging fathers' earnings is, the closer to the true β the estimated IGE will be (Mazudmer, 2005).

Secondly, the lack of permanent measures of earnings might cause the so-called lifecycle bias if too young children are considered. More specifically, estimated elasticities are influenced by the amount of earnings dispersion which tends to become higher as individuals get older, since earnings profiles are steeper for those with higher long-run earnings. Therefore, Haider and Solon (2006) suggest choosing both parents and children at median age to minimise the lifecycle bias when permanent measures of earnings are not available.

Table 1.1 summarizes estimated earnings elasticities from different empirical studies on 8 developed countries which use a 4/5 year-time average of parental

[3] For a review of the studies on intergenerational earnings mobility, see Solon (1999), Black and Deveroux (2010) and Blanden (2013).

earnings on the right-hand side of equation 1. Reported elasticities identify the Unites States as the less mobile society among those considered with an estimated β of 0.54. Conversely, Denmark is reported to be the most mobile country with an estimated earnings elasticity of 0.14.

Table 1.1: Intergenerational earnings elasticity: OLS estimates (4/5yrs averaged fathers'earnings)

Country	Source	Earnings Elasticty
U.S.	Zimmerman (1992)	0.54
Norway	Nilsen et al. (2012)	0.27 (on average)
Germany	Vogel (2008)	0.25
Sweden	Björklund & Chadwick (2003)	0.24
Canada	Corak & Heisz (1999)	0.23
Finland	Pekkarinen et al. (2009)	0.23-0.30
Denmark	Hussain et al. (2008)	0.14

As described in chapter 1, information about parents' earnings and/or income are usually absent in most of developed countries and in almost all less developed countries. For this reason, it is extremely hard to rank countries in terms of economic mobility by considering only those for which an OLS estimate on effective parents' and children's incomes is available. A way to overcome this issue was first proposed by Björklund and Jäntti (1997) that make use of the two-sample instrumental variable methodology (TSIV), originally described by Angrist and Krueger (1992) and Arellano and Meghir (1992), to estimate intergenerational elasticities in Sweden and the United States. This approach exploits two independent samples and some information about some socio-economic characteristics of actual parents (usually of the father) reported by their children (usually the sons) to predict earnings of the older generation.

As time goes by, the TSTSLS method becomes gradually more used because computationally more convenient and asymptotically more efficient than the TSIV (Inoue and Solon, 2010). The TSTSLS method is implemented by exploiting a sample of adult sons, who report some socio-economic characteristics of their actual fathers, and an independent sample of individuals different from the actual fathers observed during the childhood of the adult sons to obtain intergenerational elasticities in a two-stage approach.

The number of retrospective variables available may influence obtained IGE in two ways. Firstly, as $0 \leq R^2 \leq 1$, the variance of father's predicted earnings is less than or equal to the variance of actual father's earnings and $\hat{\beta}_{TSTSLS} = \beta_{TRUE}$ when $R^2 = 1$. This means that the higher the number of good auxiliary variables is, the higher the explained variance of pseudo-fathers' earnings will be and the lower the bias of the TSTSLS estimator is expected to be. This is mainly due to the fact that the estimated elasticities converges in probability to the following expression:

$$\beta \cong \rho_{sf} \frac{sd_s}{sd_f} \quad (2)$$

where ρ_{sf} is the correlation between sons' and fathers' earnings and sd_s and sd_f are the two standard deviations.

1.2. Data and sample selection

Our estimates of intergenerational earnings mobility are obtained by relying on AD-SILC, a very rich panel dataset built merging the 2005 wave of the Italian sample of the Survey on Income and Living Condition (IT-SILC) conducted by Istat (the National Italian Statistical Institute) with information collected from administrative archives managed by the Italian Social Security Institute (INPS) that cover individual earnings histories from the moment they enter labour market earnings up to the end of 2013. The administrative archives provide records of every job relationship that individuals experienced during the year such as the duration (measured in weeks), the fund where the worker pays contributions (allowing us to distinguish private and public employees and the various groups of self-employed), gross earnings (including personal income taxes and pension contributions paid by the worker). Furthermore, we can distinguish weeks spent working from weeks spent receiving maternity, sickness and CIG allowances or unemployment benefits. The panel structure of our data allows us to exactly measure the time of entry in the labour market and the effective labour market experience since the entry. Note that administrative archives record information on all types of workers in Italy, thus they are free from attrition; furthermore, earnings measured in administrative archives are less affected by measurement errors than survey data.

As in most of studies on intergenerational mobility we analyse the relationship between parents and children focusing on fathers and sons. In order to carry out our empirical strategy, the IT-SILC 2005 survey contains a specific section about intergenerational mobility and thus, information about fathers' characteristics when sons were aged around 14, e.g. father's educational attainments, occupations and activity status. The 2005 wave of IT-SILC has then been merged with the INPS archives in order to obtain retrospective information on fathers through the IT-SILC and sons' earnings from the administrative archives.

We select two subsamples of sons and pseudo-father according to the following rules.We consider sons born in the period 1970-1974 and follow these individuals since they are aged 35 up to age 39. Thus, according to their birth year, sons are followed in the period 2005-2013. Earnings since age 35 to age 39 are also averaged. Pseudo-fathers are selected among those individuals observed in the period 1980-1988 and aged between 40 and 44 in INPS archives (and their earnings over the period are also averaged): thus we consider pseud-fathers born in the period 1940-1944. The two generations are thus observed at middle ages according to the selection rules proposed by Haider and Solon (2006) to minimize the amount of lifecycle bias.

Our main variable of interest, annual gross earnings, includes both employment and self-employment labour income and is considered in real terms (it has been deflated according to the 2012 Consumer Price Index). Thus, considering gross incomes, we are able to first evaluate the extent of intergenerational mobility in the labour market before the effect of taxes and transfers tooks place. Then, in order to compare our results to previous estimates of the IGE for Italy (Piraino, 2007; Mocetti, 2007), we reconstruct net earnings and estimate intergenerational mobility measures after the redistributive intervention on earnings exerted by the State through social contributions and income taxes[4].

Descriptive statistics presented in table 1.2 show that the two final samples count 1445 sons and 2742 pseudo-fathers. Gross earnings are slightly more dispersed in the sample of sons than in the sample of pseudo-fathers. As expected, income taxation reduces earnings dispersion in both generations.

Table 1.2: Two-Sample Descriptive Statistics

	Sons	Pseudo-Fathers
Age (Mean)	38.80	41.97
	(0.69)	(0.44)
Log Gross Earnings (Mean)	10.00	9.90
	(0.66)	(0.49)
Log Net Earnings (Mean)	9.65	9.63
	(0.59)	(0.46)
Observations	1445	2742

Author's elaboration based on the AD_SILC dataset.
Standard deviations in parenthesis. All economic variables are deflated by using the 2012 consumer price index

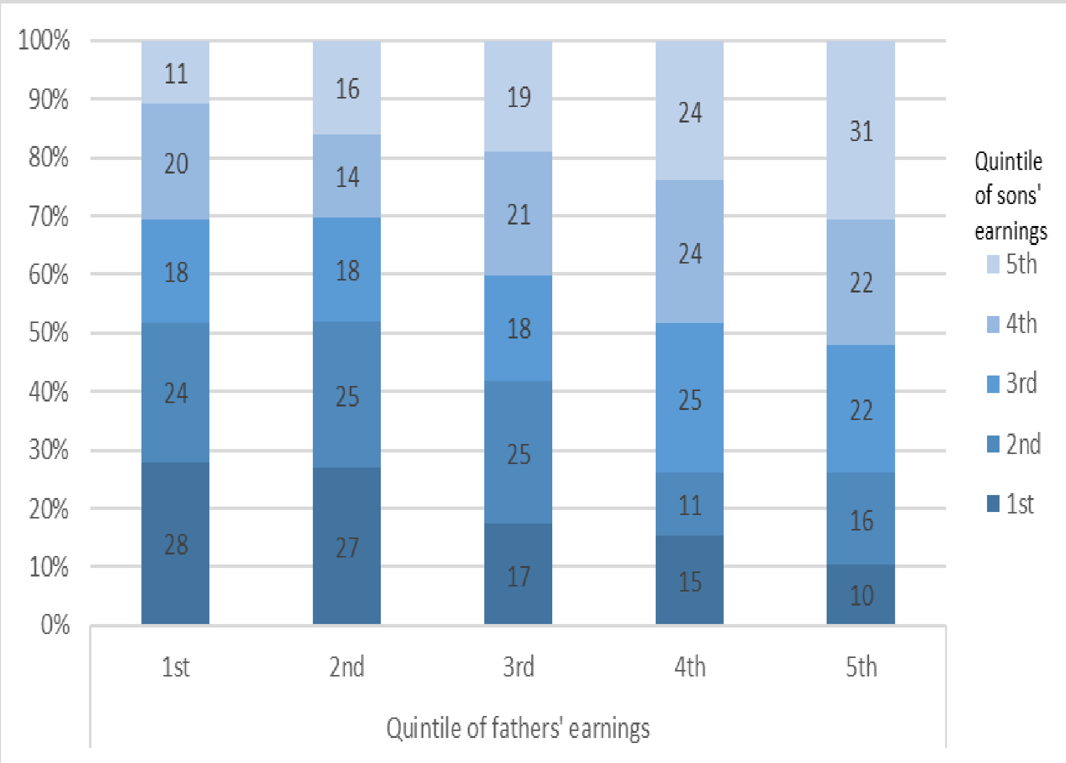
A first way to describe the extent of intergenerational mobility in Italy is to present sons' probabilities of ending up in a specific quintile of the earnings distribution given the quintile of their fathers' (figure 1.1). This kind of descriptive

[4] We first subtract employee and self-employed mandatory social contributions and then apply to all individuals tax rules (i.e. tax rates and related deductions).

analysis may be also useful to evaluate the pattern of mobility along the distribution, as previous literature reports higher levels of intergenerational economic correlation at the top of the earnings' distribution (see Björklund et al. 2010).

Figure 1.1 shows that for most gross earnings quintiles[5], sons are more likely to ends up in the same quintile as their fathers (diagonal probabilities are all greater than 20 percent expect the third quintile).

Figure 1.1: Probability of ending up in a specific quintile of the gross earnings distribution given the quantile of their fathers'



Author's elaboration based on the AD_SILC dataset.

The degree of persistence of earnings across generations is particularly high at the top and at the bottom of the distribution: 31 percent of sons whose pseudo-

[5] We do not report also mobility matrix for net earnings since results are the same, as taxation does not re-rank individuals.

fathers were collocated in the highest quintile of the distribution remains in the same quintile and more than 50 percent in one of the highest two quintiles. Conversely, only about 10 percent of sons from the best economic background end up in the worst quintile. The degree of persistence is also high at the bottom-end of the earnings distribution: about 52 percent of sons coming from the lowest quintile of the fathers' earnings distribution remains in one the bottom two quintiles.

1.3. Empirical strategy

As in previous studies on intergenerational economic mobility in Italy, we exploit the TSTSLS method to obtain measures of intergenerational associations for both gross and net annual earnings. We perform the method by exploiting a set of fathers' socio-economic characteristics reported by sons that can be used to predict their fathers' earnings. More formally, in the first stage we estimate the following equation by exploiting the sample of pseudo-fathers:

$$Y_{i,t}^{pf} = \alpha + \theta_1 Z_i^{pf} + v_{i,t} \tag{3}$$

where $Y_{i,t}^{pf}$ is is the logarithm of pseudo-fathers' earnings, Z_i^{pf} is the vector of socio-economic characteristics of pseudo-fathers and $v_{i,t}$ is the usual disturbance. The set of auxiliary variables contained in Z_i^{pf} includes 4 educational categories (primary or lower, lower secondary, upper secondary and tertiary degree), 27 occupational categories (according to the 2 digits ISCO-88 classification), 20 dummies on the region of residence[6] and a dummy for self-employment.

Then, we obtain the IGE in the second stage, by regressing the logarithm of sons' earnings on that of pseudo-fathers':

$$y_{i,t}^{S} = \alpha + \beta \hat{y}_i^F + \mu B_i^s + \epsilon_{i,t} \tag{4}$$

[6] We link sons' region of birth to parents' region of residence to avoid biases related to a possible mobility across regions of children during their adult age.

where $y_{i,t}^{S}$ is the logarithm of sons' earnings, $\hat{y}_i^F = \hat{\theta}_1 Z_i^F$ is the prediction of the logarithm of fathers', B_i^s is year of birth of the son and β is the IGE.

Even though we are exploiting an instrumental variable approach based on two independent samples, we do not aim to identify the causal effect of fathers' earnings on sons' earnings. The object is to merely predict the former in the best possible way. This is the reason why we do not require the set of auxiliary variables - used to predict fathers' earnings in the first stage - to satisfy any exclusion restriction. However, we are not able to obtain a perfect prediction of fathers' lifetime earnings by using the set of socio-economic characteristics at our disposal. This is why a TSTSLS estimator could be affected by three different kind of potential biases compared to the OLS estimator obtained by using fathers' earnings averaged over a multi-year period.

Firstly, an attenuation bias deriving from the fact that we are using an imputed value instead of an actual value as a regressor. We are thus introducing measurement error.

Secondly, if socio-economic characteristics of fathers are positively correlated with the error term in equation 4 (if auxiliary variables are not exogenous), we are introducing an upward bias in our estimates as the predicted variance of the earnings of the first generation is lower than actual variance.

A further source of potential bias can derive from the fact that there could be other unobservables included in $v_{i,t}$ (e.g. soft skills, social networks, cultural factors, cognitive and non-cognitive abilities) not totally captured by the set of auxiliary variables used in the first stage. In this case, estimates of earnings mobility could be upward biased (downward biased) if these unobservables are negatively (positively) correlated across generations[7].

However, the R^2 of the first stage regression may be considered as a good measure of the fraction of the variance of averaged pseudo-fathers' earnings predicted from auxiliary variables only if averaged rather than yearly pseudo-fathers' earnings were available in the first stage. Unfortunately, empirical works

[7] See Olivetti and Paserman (2015) for a more detailed and formalised discussion of the different potential sources of bias deriving from imputing fathers' earnings.

that use the TSTSLS approach are often not able to use averaged earnings as a dependent variable in the first stage. More formally, at any point of time, earnings of pseudo-father i may be expressed according to the following expression:

$$Y_{i,t}^{pf} = Y_i^{pf} + \varphi_{i,t} + \omega_t \tag{5}$$

where $Y_{i,t}^{pf}$ and Y_i^{pf} are respectively yearly and averaged pseudo-fathers' earnings, $\varphi_{i,t}$ are transitory individual shocks (or measurement errors) and ω_t are aggregate transitory shocks. This means that:

$$\sigma^2\left(Y_{i,t}^{pf}\right) > \sigma^2\left(Y_i^{pf}\right) \tag{6}$$

$$R^2\left(Y_{i,t}^{pf}\right) < R^2\left(Y_i^{pf}\right) \tag{7}$$

where $\sigma^2\left(Y_{i,t}^{f}\right)$ and $\sigma^2\left(Y_i^{f}\right)$ are the two variances and $R^2\left(Y_{i,t}^{f}\right)$ and $R^2\left(Y_i^{f}\right)$ are the proportion of the two variances that is predictable from the set of auxiliary variables in the first stage. According to this framework, it is plausible to say that R^2 in the first stage depends on three factors: 1. The number of auxiliary variables exploited to predict earnings (and their predictive power); 2. The number of years on which pseudo-fathers' earnings are averaged; 3. The amount of transitory shocks occurred to individuals over the period of analysis[8].

In this chapter, we try to partially reduce some of these sources of biases with respect to previous evidence for Italy. To do that we exploit a set of auxiliary variables which allow us to explain about 40% of the variance of pseudo-fathers' earnings (about 10% higher than those obtained in the first stage by Piraino, 2007 and Mocetti, 2007). This means that we are partially reducing unexplained variance due to unobservables included in $v_{i,t}$. Moreover, we are able to reduce measurement

[8] For a more detailed discussion of the downward bias derived from using yearly instead of averaged earnings for the first generation, see Jerrim et al. (2016)

error also by averaging earnings of pseudo-fathers over a 5-year period in order to get a better prediction of lifetime earnings.

1.4. Estimated intergenerational earnings elasticities

Estimated IGE for gross earnings is presented in table 1.3. The first column reports an estimated IGE of 0.496, obtained when earnings of both pseudo-fathers and sons are observed in a 5-year period. Such a result means that a 10 percent variation in fathers' earnings is associated with a 4.96 percent variation in son's earnings. This estimated IGE is slightly higher than that of 0.44 reported by Piraino (2007) and close to that of 0.50 reported by Mocetti (2007) using net instead of gross incomes.

However, when we use a single year measure for both the two generations – observing fathers and sons, respectively, in 1985 and 2009 only – the estimated IGE becomes lower than that obtained by Piraino (2007) and Mocetti (2007) that used point time estimates for both generations. This lower value is probably related to a reduction in the unexplained variance of pseudo-fathers' earnings. On the contrary, the use of time-averages increases our estimated elasticity. This means that, although the two generations are taken at middle ages as suggested by Haider and Solon (2006), using a single year measure of earnings may cause a downward bias in estimated IGE due to both left-hand and right-hand side measurement errors. These results are consistent with previous evidence which show that both TSTSLS and OLS estimates of intergenerational earnings elasticities are likely to be downward biased using point in time measures of fathers' earnings, even when commonly used selection rules for both generations are exploited (Gregg et al. 2014, Jerrim et al., 2016).

The last column of table 1.3 show that the IGE increases when zero earnings observations are not excluded from the analysis, i.e. when individuals that are not present in INPS archives in a year in the observed period are considered in the estimates considering a zero-earning value for that year. This suggests that sons of poorer fathers are likely to have more unstable careers (i.e. to spend a year without earnings) than workers coming from a better background.

Table 1.3: Association between son's and father's gross earnings. Prime age sons[a]. OLS estimates in the second stage[b]

	Observation span of fathers' and sons' earnings				
	5 years-5 years	5 years-1 year	1 year-5 years	1 year-1 year	5 years-5 years Imputing zeros[c]
Father's earnings	0.496***	0.441***	0.402***	0.382***	0.623***
s.e.	[0.056]	[0.056]	[0.054]	[0.054]	[0.081]
Obs	1445	1365	1445	1365	1481
R^2	0.059	0.048	0.043	0.040	0.044
R^2 first stage	0.409	0.409	0.404	0.404	0.409

[a] when observed in a 5 year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 2009. b TSTSLS are carried out: in the first stage father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage sons' log annual gross earnings are regressed on predicted fathers' log earnings, also controlling for sons' year of birth. c 5-year average of sons' earnings in age class 35-39 are computed assigning a zero value to sons who do not report earnings in administrative archives in a certain year. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset

The presence of the life cycle bias is confirmed when we replace the sons' generation by considering a sample of younger sons aged 25-29 born from 1970 1974 (Table 1.4). In this case, the estimated IGE obtained by measuring over a 5-year period falls to 0.270 (0.365 when also zero earnings observations are included in the analysis).

Some sensitivity tests are performed to evaluate the robustness of our estimated elasticity. First, we check whether the estimated IGE changes if we exclude a single predictor from the first stage regression[9]. More specifically, the TSTSLS estimate should be considered upward (downward) biased if auxiliary variables used in the first stage have positive (negative) direct effect on sons' earnings. Results presented in table A.1 in the appendix show that the estimated elasticity tends to be extremely stable if either educational or occupational categories are excluded from the set of auxiliary variables exploited as predictors in the first stage and becomes higher (lower) if we exclude the dummy for self-employment (region of residence) from the first-stage regression of pseudo-fathers' earnings.

[9] We perform the Sargan test to evaluate if the full set of instruments used in the first stage in uncorrelated with the error term of the second stage regression. The test rejects the null hypothesis, which means that at least one instrument is not exogenous.

Table 1.4: Association between son's and father's earnings. Young sons[a]. OLS estimates in the second stage[b]

	Observation span of fathers' and sons' earnings				
	5 years-5 years	5 years-1 year	1 year-5 years	1 year-1 year	5 years-5 years Imputing zeros[c]
Father's earnings	0.270***	0.225***	0.235***	0.237***	0.365***
s.e.	[0.058]	[0.079]	[0.057]	[0.076]	[0.097]
Obs	1395	1147	1395	1147	1410
R^2	0.016	0.020	0.013	0.022	0.018
R^2 first stage	0.409	0.409	0.404	0.404	0.409

[a] When observed in a 5 year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 25-29 in the period 1995-2003. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 1999. b TSTSLS are carried out: in the first stage father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage sons' log annual gross earnings are regressed on predicted fathers' log earnings, also controlling for sons' year of birth. c 5-year average of sons' earnings in age class 25-29 are computed assigning a zero value to sons who do not report earnings in administrative archives in a certain year. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset

As mentioned, previous estimates for Italy (Piraino, 2007; Mocetti, 2007) are computed on net earnings. In order to better compare our estimates, we derive also measures of net earnings for both generations. In table 1.5 we present the IGE computed on net earnings. The estimated IGE of 0.428 is obtained when earnings of both sons and pseudo-fathers are observed over a 5-year period. This estimated IGE is lower than the one previously obtained when using gross earnings. It is possible to notice that, with respect to our previous estimates, the R^2 for the first stage equation – thus, the explained variance of pseudo-fathers' earnings – increases: the same auxiliary variables seem to better predict net earnings than gross earnings. Moreover, the income taxation system has reduced more earnings dispersion in the sons' generation than in the first generation.

When comparing our estimates using net earnings, we can see that estimated IGE now is lower compared to those reported by both Mocetti (2007) and Piraino (2007), even if we use earnings of both pseudo-fathers and sons observed in a 5-year period. Moreover, when we use a 1-year measure of net earnings for both generations, we obtain an even lower estimated of the IGE with respect to previous estimates for Italy (Mocetti, 2007; Piraino, 2007) that use point in time measures.

As in the case of the estimated IGE obtained using gross earnings, the use of time-averages increases our estimated IGE.

Table 1.5: Association between son's and father's net earnings. Prime age sons[c]. OLS estimates in the second stage[b]

	Observation span of fathers' and sons' earnings				
	5 years-5 years	5 years-1 year	1 year-5 years	1 year-1 year	5 years-5 years Imputing zeros[c]
Father's earnings	0.428***	0.383***	0.350***	0.333***	0.540***
s.e.	[0.048]	[0.048]	[0.049]	[0.046]	[0.073]
Obs	1445	1365	1445	1365	1481
R^2	0.056	0.047	0.042	0.040	0.038
R^2 first stage	0.463	0.463	0.472	0.472	0.463

[a] when observed in a 5 year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 2009. [b] TSTSLS are carried out: in the first stage father's earnings are imputed regressing log annual net earnings on dummies on education, occupation, self-employment, region of work; in the second stage sons' log annual net earnings are regressed on predicted fathers' log earnings, also controlling for sons' year of birth. [c] 5-year average of sons' earnings in age class 35-39 are computed assigning a zero value to sons who do not report earnings in administrative archives in a certain year. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset.

Comparing estimates of intergenerational elasticity coefficient is not a trivial exercise, because differences in estimates must be interpreted considering many factors such as the measure of earnings or income used, the sample selection and the applied methodology. For example, when the TSTSLS method is applied, the set of auxiliary variables used in the first stage to predict fathers' earnings vary across different studies, depending on the availability of retrospective socio-economic information about fathers reported by sons. In the best-case scenario, a large number of socio-economic characteristics of the father such as his education, occupational qualification, sector of activity, geographic area and age (or year of birth) are all exploited as predictors; in the worst, estimates are obtained by using only one retrospective variable such as in the case of Brazil (see for example Dunn, 2007).

Some results from empirical studies estimating earnings elasticities for different countries using either the TSIV of the TSTSLS method are presented in table 1.6.

Table 1.6: Intergenerational earnings elasticity for developed and less developed countries: TSTSLS or TSIV estimates

Country	Source	Earnings Elasticity
Ecuador	Grawe (2004)	1.13
Brazil	Dunn (2007)	0.69
Chile	Nunez and Miranda (2011)	0.66
South Africa	Piraino (2015)	0.62-0.68
China	Gong et al. (2012)	0.63*
Peru	Grawe (2001)	0.60
Brazil	Ferreira & Veloso (2006)	0.58
U.S.	Björklund and Jäntii (1997)	0.52
Italy	Mocetti (2007)	0.50
Pakistan	Grawe (2001)	0.46
Italy	Piraino (2007)	0.44
Nepal	Grawe (2001)	0.44
Spain	Cervini-Plà (2015)	0.42
France	Lefranc & Trannoy (2005)	0.40
South Korea	Ueda (2013)	0.35
Japan	Ueda (2013)	0.35
U.K.	Bidisha et al. (2013)	0.33
Germany	Cavaglià (2015)	0.30
U.K.	Nicoletti and Ermish (2007)	0.29
Sweden	Björklund and Jäntii (1997)	0.28
Australia	Leigh (2007A)	0.20-0.30
Taiwan	Kan et al. (2015)	0.18*

* Income elasticity is reported since an estimate of the earnings elasticity is not available

Unlike OLS estimates of earnings mobility, which are only available for a small number of developed countries, elasticities obtained by means of either the TSTSLS or TSIV approach are provided for many developed and less developed countries. The latter are reported to be less mobile societies as estimated earnings elasticities are basically greater than 0.60. On the contrary, Taiwan, Sweden and Australia, among those considered, are countries with high levels of intergenerational mobility with estimates IGE below the value of 0.30. If we consider our estimates of the IGE of 0.43, when we use net earnings or 0.50 when we use gross earnings, we can consider Italy as a medium-mobility country, when compared to both developed and less developed countries, and a low-mobility country when restricting the analysis to the subsample of developed countries. For instance, IGE estimates in other developed countries such as Germany, UK, Sweden, Australia and Japan are below the value of 0.40. It is interesting to compare our estimates with the elasticity of another Mediterranean country, Spain (Cervini, 2015), obtained using TSTSLS method and gross earnings. According to her results, the IGE for Spain is of 0.42, lower than our estimated IGE of 0.50 obtained using gross earnings and very close to the one obtained when using net earnings (0.43).

1.5. Estimated rank-rank slopes

Since the size of the intergenerational elasticity coefficient depends on the income dispersion in the two generations, we also estimate an alternative measure of intergenerational mobility: the rank-rank slope, a measure of the association between fathers' relative position in their respective earnings distributions (Dahl & DeLeire, 2008). From a statistical point of view, rank-rank slopes are usually intended to be more robust across samples and specifications (Chetty et al. 2014; Gregg et al. 2014).

The intergenerational elasticity coefficient converges in probability to the correlation coefficient of log earnings times the ratio between the standard deviation in sons' generation and that in fathers' generation. Hence, given the connection between the intergenerational elasticity coefficient and the correlation coefficient and given that the correlation coefficient and the rank-rank slope are both scaleinvariant measures of relative mobility, we can easily see how the rank-rank slope is closely related to the intergenerational elasticity coefficient, but has the advantage to be "independent" from inequality within generations. In other words, the intergenerational elasticity coefficient may be affected by a change in inequality across the two generations, whereas rank to rank slope is not.

If our aim is to provide estimates of intergenerational mobility for Italy that can be compared with those of other countries, estimating rank to rank slopes may be a more suitable strategy. Since the level of inequality is not the same across countries, the rank to rank slope may provide a better picture of differences in intergenerational mobility.

Rank to rank slopes are also more robust with respect to both the two key measurement issues, namely life-cycle and attenuation bias (Gregg et al. 2014). Life-cycle bias is mainly driven by mismeasurement of earnings gaps between individuals rather than positional inaccuracy along the earnings distribution. When using rank to rank slope we have also to deal with an attenuation bias smaller in magnitude since measurement errors and transitory shocks cause scale mismeasurement rather than positional inaccuracy in the earnings distribution.

Rank-rank slopes are usually obtained by estimating the following equation:

$$p_s = \alpha + \gamma p_f + \varepsilon \tag{8}$$

where p_c is the percentile of sons' earnings in their own distribution and p_f is the percentile of fathers'. In this framework, an estimated γ of 0.5 means that the expected difference in ranks between sons would be about 5 percentiles if the difference in ranks among their fathers was 10 percentiles. However, we are not able to estimate rank-rank slopes by simply re-categorizing earnings of the two generations since data on actual fathers' earnings are not available. For this reason, we exploit a different approach consisting in three different steps. Firstly, we obtain a prediction of the logarithm of fathers' earnings by exploiting the sample of pseudo-fathers and the same set of auxiliary variables used for obtaining TSTSLS estimates of the IGE. Secondly, predicted fathers' earnings are percentile ranked so that we can estimate in the last step the following equation:

$$p_s = \alpha + \gamma \hat{p}_f + \varepsilon \tag{9}$$

where p_c is the percentile of sons' earnings in their own distribution and $\hat{p}_f$ is the imputed percentile of fathers' earnings distribution. This approach, apart for the set of auxiliary variables exploited in the first step, is very close to that used by Olivetti and al. (2016) to obtain intergenerational and multigenerational imputed rank-rank slopes for the US[10].

From a statistical point of view, it is not easy to understand to what extent our imputed rank-rank slope can be compared to rank-rank slopes obtained by percentile ranking actual fathers' earnings. Obviously, when we impute the percentile of the father from a predicted variable we are likely to make some errors in placing all fathers in the right percentile of their earnings distribution. For this reason, our estimates are likely to be affected by attenuation bias. However, this kind of positional measurement errors cannot be intended as "classical" (see Nybom and Stuhler, 2016) since both our dependent variable and the regressor in equation 8 and 9 are uniformly distributed. This means that all statistical properties based on the assumption of normally distributed variables do not hold in our case. This is why we should exercise caution in comparing our imputed rank-rank slope to estimates obtained in previous studies for other countries.

Table 1.7 presents our findings and shows, in the case of multi-year averages for both generations, an estimated value of 0.254 which, interestingly is characterized by a lower reduction, compared to the IGE in table 1.5, when different yearly values of annual earnings are considered. Moreover, in table 1.8 we report imputed rank to rank slopes for net earnings. Since income taxation should affect earnings dispersion but not the position on the ladder of the income distribution, estimates

[10] As in a previous article by Olivetti and Paserman (2015) they impute father's income, which is unobserved, using the average income of fathers of children with a given first name.

of imputed rank to rank slopes do not change. These findings further confirm rank to rank slope robustness.

Table 1.7: Association between son's and father's gross earnings. Prime age sons[a]. Rank to rank estimates[b]

	Observation span of fathers' and sons' earnings				
	5 years-5 years	5 years-1 year	1 year-5 years	1 year-1 year	5 years-5 years Imputing zeros[c]
Father's earnings	0.254***	0.237***	0.228***	0.217***	0.249***
s.e.	[0.025]	[0.026]	[0.026]	[0.027]	[0.025]
Obs	1445	1365	1445	1365	1481
R^2	0.071	0.062	0.058	0.053	0.071
R^2 first stage	0.409	0.409	0.404	0.404	0.409

[a] When observed in a 5 year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 2009. [b] TSTSLS are carried out: in the first stage percentiles of father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage percentiles of sons' log annual gross earnings are regressed on predicted percentiles of fathers' log earnings, also controlling for sons' year of birth. [c] 5-year average of sons' earnings in age class 35-39 are computed assigning a zero value to sons who do not report earnings in administrative archives in a certain year. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset

However, it is not easy to compare our estimate to those obtained for other countries as this alternative measure of intergenerational association has a shorter history with respect to IGE and moreover, to the best of our knowledge, there is no evidence of rank to rank estimates obtained by computing fathers percentiles according to parental earnings obtained through an imputation procedure.

To the best of our knowledge Dahl & Delaire (2008) were pioneers in the usage of rank to rank: for the USA, they estimate a rank to rank slope of 0.289 taking 34 years old sons and averaging fathers' earnings from age 20 to 55 also including years of zero earnings. Still for the USA, Chetty et al. (2014) estimate a rank to rank slope of 0.34 whereas Mazumder (2015) estimate a rank to rank coefficient of 0.40 when using 15 years of fathers' earnings (0.31 if using a single year of father's earnings). Bratberg et al. (2017) in a recent work on cross country measures of intergenerational mobility display rank to rank estimate for several countries: 0.383

for the US, 0.257 for Germany, 0.233 for Norway and 0.215 for Sweden. For the UK, we have rank to rank estimates from Gregg et al. (2014): 0.34 for sons aged 42 and parental income measured when sons were 16.

Table 1.8: Association between son's and father's earnings percentiles. Young sons[a]. Rank to rank estimates[b]

	Observation span of fathers' and sons' earnings				
	5 years-5 years	5 years-1 year	1 year-5 years	1 year-1 year	5 years-5 years Imputing zeros[c]
Father's earnings	0.181***	0.182***	0.169***	0.177***	0.148***
s.e.	[0.0265]	[0.0282]	[0.0268]	[0.0298]	[0.0283]
Obs	1395	1147	1395	1147	1410
R^2	0.038	0.058	0.034	0.056	0.042
R^2 first stage	0.409	0.409	0.404	0.404	0.409

[a] When observed in a 5 year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 25-29 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 1999. [b] TSTSLS are carried out: in the first stage percentiles of father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage percentiles of sons' log annual gross earnings are regressed on predicted percentiles of fathers' log earnings, also controlling for sons' year of birth. [c] 5-year average of sons' earnings in age class 25-29 are computed assigning a zero value to sons who do not report earnings in administrative archives in a certain year. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset.

It is quite striking how these results differ from those related to cross-country rankings based on the IGE. Even if rank-rank slope estimates are available only for few countries, and thus we cannot insert our results in a widely accepted cross-country ranking based on the rank-to rank slope, now distances between countries are reduced and Italy results to be not very distant from Nordic countries, with a level of mobility very close to that of Germany.

As in the case of estimated IGE, also rank to rank estimates appear to be affected by the life cycle bias. Indeed, when we consider sons aged 25-29 instead than 35-39 (table 1.8), the estimated rank to rank coefficient falls to 0.181 when earnings of both generations are measured over a 5-year period.

1.6. Intergenerational mechanisms

A classical way to examine empirically intergenerational mechanisms behind the intergenerational transmission of earnings is to is to re-estimate the equation 4 with some additional controls included in the vector $X_{i,t}^{s}$ (e.g. Raitano and Vona 2015):

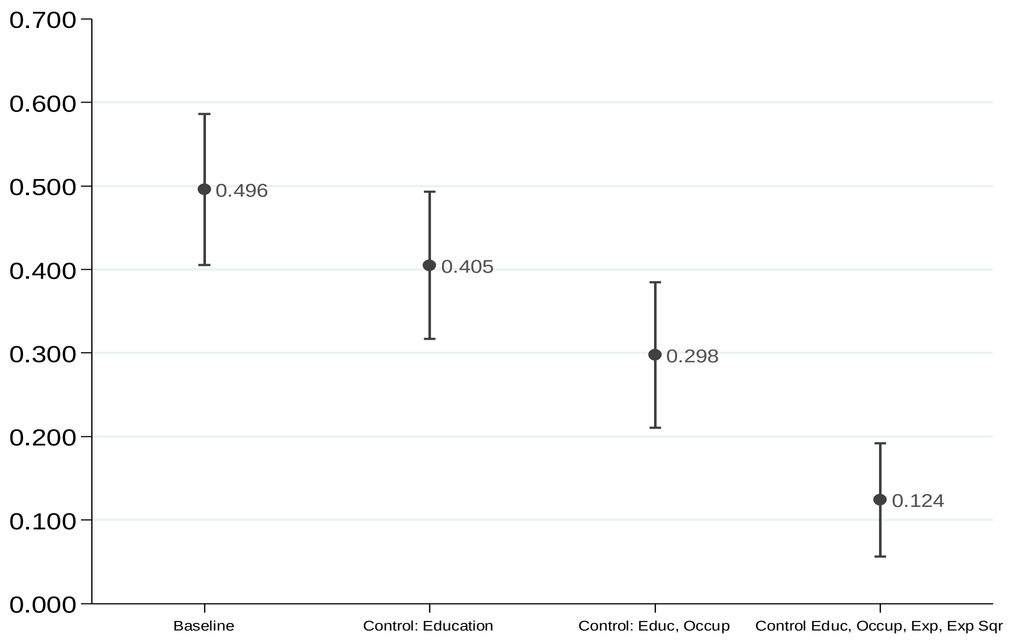
$$y_{i,t}^{S} = \alpha + \beta_2 \hat{y}_i^F + \delta X_{i,t}^{s} + \theta B_i^s + \epsilon_{i,t} \tag{10}$$

where $y_{i,t}^{S}$ is the logarithm of sons' earnings (a time average is used for sons with two or more time observations), $\hat{y}_i^F = \hat{\theta}_1 Z_i^F$ is the prediction of the logarithm of fathers', $X_{i,t}^{s}$ is the vector of control variables, B_i^s is year of birth of son and β_2 is the new estimated IGE.

Among all possible channels of influence, we consider 8 categories of sons' educational level, 27 categories of occupations (according to 2 digits ISCO) and the working status (private employee, public employee, self-employed, professional or parasubordinate worker) and work experience measured as the number of working weeks since they entry into activity. We consider three different models, where 5-year average earnings for both generations are considered: in the first only son's educational attainment is included in the vector $X_{i,t}^{s}$; the second one includes both sons' educational levels and occupation and working status; the last adds experience and thus considers all mediating variables.

The assumption is that if a mediating variable is positively correlated with both fathers' and sons' earnings, the estimated elasticity will fall once this control is included in the regression. Therefore, the difference between $\hat{\beta}$ obtained by estimating equation 4 (our baseline) and $\hat{\beta}_2$ can be interpret as the fraction of the elasticity associated to a single mediating factor. However, this is true only if this mediating variable included in the vector $X_{i,t}^{s}$ is not correlated with the error term. Conversely, if the mediating variable is positively (negatively) correlated with other unobservable factors that influence sons' earnings, the coefficient $\hat{\beta}_2$ is upward (downward) biased and the channel of influence is overestimated (underestimated).

Figure 1.2: T2TSLS estimated coefficient of the association between son's and father's earnings, including sons' outcomes among the covariates[a]. Fathers and sons observed for 5 years[b].

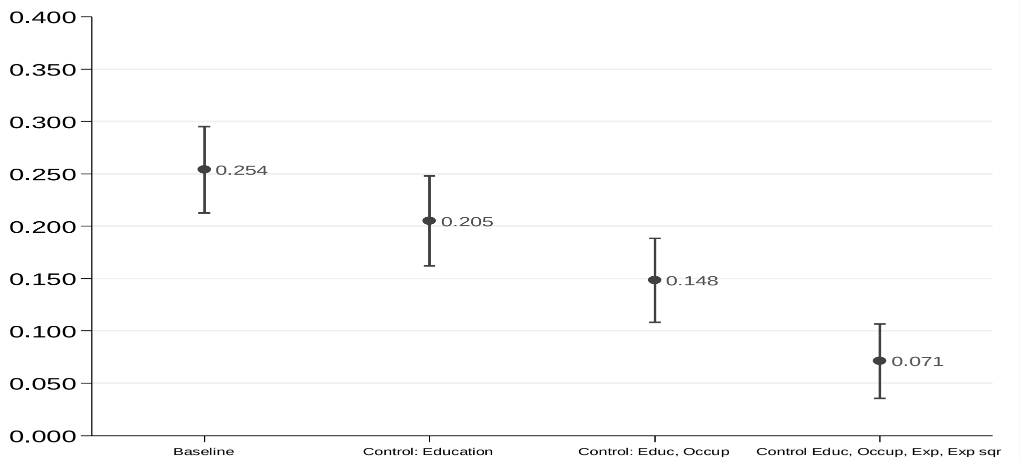


[a] Dummies on sons' year of birth are included among covariates in all estimated models. b Fathers and sons are observed, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. 90% confidence intervals. Source: elaborations on AD-SILC dataset

Estimated elasticities obtained by means of equation 10 are presented in figure 1.2 together with their 90% confidence intervals. Estimates suggest that including all three control variables together is the only result statistically different from the baseline estimate (when no sons' characteristics are controlled for).

On the contrary, the estimated IGE obtained by including either sons' educational levels or educational levels and occupation and work status are not statistically different from the baseline. These results provide some evidence that higher income of fathers' may influence their sons' economic outcomes in ways other than through the mere investment human capital. More specifically, sons from lower income families may obtain less stable occupations which negatively affects their experience and may reduce their gross annual earnings, as shown in the "full model"

Figure 1.3: Rank to rank estimated coefficient of the association between son's and father's earnings percentiles, including sons' outcomes among the covariates[a]. Fathers and sons observed for 5 years[b].



[a] Dummies on sons' year of birth are included among covariates in all estimated models. [b] Fathers and sons are observed, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. 90% confidence intervals. Source: elaborations on AD-SILC dataset.

Results obtained by estimating equation 10 are confirmed when we examine the relative importance of the three mechanisms by means of rank to rank estimates carried out starting from predicted incomes obtained by the two-stage procedure (Figure 1.3). As in the case of elasticity, we compare our baseline rank to rank slopes estimate with those obtained using different set of controls: first we add to our baseline model dummies on education, then both dummies on education and work status and, in the last model, dummies for education, work status, and effective experience. Education, again, seems to capture only a small part of the intergenerational earnings persistence. In the next section, we will deepen this results by means of a decomposition analysis.

1.7. Decomposition approach

A further detailed way to examine the role of mechanisms driving the intergenerational correlation of earnings is to exploit the sequential decomposition approach suggested by Blanden, Gregg and Macmillan (2007) and further developeded in Blanden et al. (2014).

Following Blanden et al. (2014) we decompose the IGE into two parts: the first one is how much of the father-son earnings relationship is accounted for by transmission factors – that is, some sons' outcomes, e.g. education or occupation, that are affected by parental characteristics and then influence sons' earnings –, whereas the second one is the unexplained persistence in earnings that is not transmitted through the considered mediating variables. The part of the intergenerational persistence explained by the pathway factors is the product of two measures: the relationship between fathers' earnings and the pathway factor and its monetary return in the labour market.

Among all possible transmission mechanism, this section focuses on two mediating variables: educational attainments and occupational qualification. The first step of the decomposition method consists in estimating the univariate relationship between sons' educational attainments and the prediction of logarithm of fathers' earnings:

$$Educ_i^s = \alpha_{ed} + \lambda_{ed}\hat{y}_f + e_{1i} \tag{11}$$

Then, to combine the estimated association with the return of educational attainments in the labour market, the logarithm of sons' earnings is regressed on sons' educational attainments. We control for the prediction of the logarithm of fathers' earnings, thus estimating the effect of education on sons' earnings independent of that estimated in equation 11:

$$lnY_i^{son} = \omega_1 + \rho_{ed}Educ_i^{son} + \gamma_{inc}\hat{y}_f + \nu_{1i} \tag{12}$$

It follows that the IGE estimated in equation 4 (i.e. our baseline) can be decomposed into two parts:

$$\beta = \lambda_{ed}\rho_{ed} + \gamma_{inc} \tag{13}$$

where $\lambda_{ed}\rho_{ed}$ is the indirect effect of fathers' earnings on sons' through the educational channel and γ_{inc} is the unexplained persistence in earnings that is not transmitted through education.

Then, we account for occupational attainments only by estimating in equation 14 the association between occupational status and father's earnings and in equation 15 its monetary pay-off in the labour market:

$$Occ_i^s = \alpha_{occ} + \lambda_{occ}\hat{y}_f + e_{1i} \tag{14}$$

$$y_i^{son} = \omega_2 + \rho_{occ}Occ_i^{son} + \gamma_{inc}\hat{y}_f + \nu_{2i} \tag{15}$$

In this case, the decomposition becomes:

$$\beta = \lambda_{occ}\rho_{occ} + \gamma_{inc} \tag{16}$$

Moreover, we want to consider the interaction between educational attainments and occupational choices. Therefore, once we have estimated the relationship of each variables with fathers' earnings, we estimate an equation where we consider together the return to education and occupation in the labour market. In the next equation 17, we obtain the monetary pay-off of each variable, conditional on the others.

$$y_i^s = \omega_1 + \rho_{ed}Ed_i^s + \rho_{occ}Occ_i^s + \gamma_{inc}\hat{y}_f + \nu_{3i} \tag{17}$$

Now β can be decomposed as follows:

$$\beta = \lambda_{ed}\gamma_{ed} + \lambda_{occ}\gamma_{occ} + \gamma_{inc} \tag{18}$$

where we can distinguish the component of beta accounted for by educational attainments and the part of beta accounted for by occupational outcomes. Thus,

$\lambda_{ed}\gamma_{ed} - \lambda_{ed}\rho_{ed}$ gives the extent to which the influence of education is transmitted through occupation.

According to Hirvonen (2010), in order to obtain consistent estimates for the coefficients of the two mediating variables, error terms of equation 11 and equation 14 must be uncorrelated with the error term in the return equation 17. However, this assumption is likely to be violated since both educational attainments and occupational status could be related to other variables, such as cognitive and non-cognitive skills, education quality and other hardly observables factors as, for example, social networks and family ties. Unfortunately, our dataset does not provide information on education quality (e.g. marks or field of study), on cognitive and non-cognitive skills and social network (e.g. channels used to find job) to control for other sons' characteristics.

Observe however, that our decomposition approach cannot be directly compared to that proposed by Blanden et al. (2007) and Blanden et al. (2014) as we are using imputed instead of actual fathers' earnings. In fact, estimated λ_{ed} and γ_{inc} may be biased due to unobservables included in the vector ν_{1i} of equation 12 that are not captured by the set of auxiliary variables used to predict fathers' earnings. More specifically, there could be some variables (e. g. soft skills, social networks, cultural factors, cognitive and non-cognitive abilities) that are positively correlated to ̀earnings of the two generations (i.e. they are in the error term of our baseline equation 12) and to educational attainments of offspring. Therefore, we are likely to be underestimating the mediating role of education if imputed earnings are less correlated to the mediating variable than actual earnings (i.e. imputed earnings are less correlated to unobservables in equation 12 which are correlated to educational attainments)

With all this in mind, we proceed to decompose intergenerational mobility into different channels. Educational attainments of sons included in the decomposition analysis are provided by the 2005 wave of the IT-SILC survey. More specifically, educational levels are coded according to the five main International Standard Classification of Education levels (ISCED). Here we rely on a four-modal distribution of education: "tertiary graduates", "high-school graduates", "middle school graduates" and "elementary". However, when we estimate the univariate relationship between fathers' earnings and sons' education, exclusive dummies would lead to ambiguity in the interpretation of the coefficient for the middle category. Thus, following Blanden et al. (2014) we redefine our dummy on education as equal to one for all those who are at the relevant education level or above: "tertiary graduates", "at least high-school" and "at least middle school". In this case the coefficient must be interpreted as the incremental effect of that education level compared to the next lower level of education.

Regarding occupational status, it was originally classified according to ISCO codes: the lowest ISCO code indicates the highest occupational quality. We convert ISCO categories in a four-modal distribution of occupation: "higher managerial and professional" (corporate managers, professionals, legislators), "lower managerial and professional" (associate professionals, managers of small enterprises), "intermediate" (clerks and service workers), "bottom occupation" (assemblers, agricultural, crafts, elementary occupations). As with education, we then redefine our variables equal to one for all those who are at the relevant occupational level or above[11].

The decomposition analysis is carried out on IGE estimates run on 5-year earnings averages for both generations (see section 1.4). Results are summarized in Table 1.9. The overall IGE can be decomposed into the relationship between father's earnings and the mediating variables (λ) multiplied by the return to those variables in the labor market (γ), plus the unexplained persistence in earnings that is not transmitted by those factors. Column (i) considers only education as mediating variable, column (ii) only occupation and column (iii) consider the interaction between the two variables.

[11] We use less categories for both mediating variables than we did in the previous sections for computational reasons

Table 1.9: Decomposition: share of β explained by mediating variables.

factor	(i)	(ii)	(iii)
college degree	0.022		0.011
at least highschool	0.047		0.035
at least middle school	0.018		0.018
Total educational outcomes	0.087		0.064
higher manangerial or professionals		0.004	0.001
at least lower managerial or professional		0.021	0.017
at least intermediate		0.026	0.017
total occupational outcomes		0.051	0.035
total accounted for (λ*γ)	0.087	0.051	0.099
not accounted for	0.409	0.445	0.397
total	0.496	0.496	0.496
% through ed.outcomes	17.61%		12.91%
% through occupational outcomes		10.34%	7.08%
% of total	17.61%	10.34%	19.99%

Following Blanden (2014) we add to the full decomposition, model (iii), the two variables in the order in which they occur in the aging process. The sum of the explained and unexplained component of β is 0.496, that is the total association between fathers' earnings and sons'. In model (i) we first include only education that explains 17.6% of the intergenerational persistence, whereas around 82% can be accounted as the direct effect of father's earnings. In model (ii) we include only occupations that accounts for almost 10.3% of the intergenerational elasticity coefficient. When in model (iii), our complete decomposition, we include both education and occupation, the magnitude of the indirect effect implies that around 20% of total effect is mediate by education and occupation. The share of persistence accounted for by education decreases from 17.6% to 12.9%. Thus, the two mechanisms are clearly correlated and occupation takes over some of the explanatory power of education. Parental background exerts its effect on education, education effects sorting into occupation and occupation influences earnings. When we move from model (i) to model (iii) we notice that proportion of the intergenerational elasticity coefficients explained increase only by 2.4 percentage points thus occupation contributes directly only marginally in explaining intergenerational persistence, but it is education that works through occupational sorting.

Table 1.10: Detailed decomposition results

	association with father's earnings (λ)	return in the labour market (γ)		
		(i)	(ii)	(iii)
<u>Education:</u>				
college	0.217	0.1024		0.0501
	[0.0331]	[0.0520]		[0.0586]
at least high school	0.2752	0.1704		0.1262
	[0.0376]	[0.0382]		[0.0383]
at least middle school	0.065	0.2803		0.2836
	[0.0161]	[0.1119]		[0.1071]
<u>Occupation:</u>				
high managerial and professional	0.0964		0.0421	0.0155
	[0.0277]		[0.0685]	[0.0750]
at least low managerial and professional	0.1955		0.1094	0.0869
	[0.0414]		[0.0541]	[0.0536]
at least intermediate	0.2392		0.108	0.0695
	[0.0420]		[0.0476]	[0.0476]

Author's elaboration based on the AD_SILC dataset.

In table 1.10 we report the estimates that are behind the decomposition presented in table 1.9. The first column reports the λ coefficient estimated in the set of regression of the relationship between the mediating variables and father's earnings. The second pair of columns presents the γ coefficients from the single regression of log sons' earnings on the set of included pathway variables. Column from two to fourth display the γ coefficient from the regression of sons' earnings on the mediating factors: equation (i) regress son's earnings on education, equation (ii) on occupation and equation (iii) on both educational and occupational levels.

1.8. Concluding remarks

This chapter provides new evidence on the degree of earnings correlation across generations in Italy, which may be considered as a low mobility country (Piraino, 2007; Mocetti, 2007). New results are provided by relying on the AD-SILC, a very rich panel dataset built merging the 2005 wave of the Italian sample of the Survey on Income and Living Condition (IT-SILC) conducted by ISTAT (the National Italian Statistical Institute) with information collected from administrative archives managed by the Italian Social Security Institute (INPS) that cover individual earnings histories from the moment they enter the labour market to the end of 2013. The advantages of exploiting this dataset are twofold. Firstly, unlike previous estimates of the intergenerational earnings elasticity (IGE) for Italy obtained by using net earnings, we are able to examine the extent of intergenerational mobility in the labour market which is not mediated by the redistributive effect of taxes and, once reconstructed net earnings, we can compare the two measures. Secondly, we can rely on a large panel dimension which permits to obtain a measure of earnings which is less affected by lifecycle and attenuation biases.

As in previous studies on intergenerational economic mobility in Italy, we exploited the two-sample two-stage least squares (TSTSLS) method to obtain different measures of intergenerational earnings associations. Nonetheless, unlike previous studies for Italy, we exploited the large panel dimension of the dataset to measure earnings over a 5–year period both in the first and second stage of the

TSTSLS approach. Moreover, our auxiliary variables have a higher number of categories, thus allowing us to obtain a higher predictive power in the first stage.

Results showed an IGE of 0.496 for gross earnings and an IGE of 0.428 for net earnings. The two measures both become lower if estimates are obtained by using point in time measures of earnings of the two generations or young sons.

We also provided estimates of rank to rank slopes for Italy that proved to be more robust across different specifications, samples and measures of earnings. Since this measure remove the “within generation” inequality component, it is particularly suited for cross-country comparisons. Moreover, rank to rank slope measures are available only for few countries, and thus a cross country ranking based on this measure is not easy. However, according to our results, conversely on what we find for the IGE, Italy is not so distant from Nordic European countries and very closed to the level of mobility of Germany. Therefore, it is highly desirable for future research to provide, besides the IGE, also measures based on the rank to rank slope.

Education is usually recognized as the most prominent mechanism affecting the intergenerational transmission of income from parents to children. We presented additional estimates including as regressors children characteristics affected by parental circumstances – e.g. education, occupation – to assess whether a residual influence of parental background still emerges when these characteristics are controlled for.

Furthermore, we also applied the decomposition method proposed by Blanden et al. (2014) in order to compare the role played by education and occupation as transmission mechanisms of intergenerational inequality. We find that, when considered together, education and occupation account for around 20% of father-son earnings relationship: education contributes to 12.9% of the earnings persistence, while occupation account for around 7.1%.

Thanks to the sequential decomposition we assessed that the proportion of β accounted for by education becomes smaller when occupational attainments, that occur later in life with respect to the educational ones, are included. Therefore, part of the effect of education is absorbed by occupational choices. Considering education alone, as the only transmission mechanism, may overstate its explanatory power of the persistence of socio-economic outcomes (Hirvonen, 2010). Even when we considered both education and occupation, it is highly likely that the part of the intergenerational elasticity coefficient ascribed to education, still conceals some of the explanatory power that should be ascribed to other mechanisms (e.g. soft and hard skills). Therefore, it is likely that we are overestimating the effect of the education and these results should be interpreted as an upper bound. The role of education in Italy, compared to that found by Blanden at al. (2014) for the UK and the US is very limited and it is even more limited if our results should be interpreted as an upper bound.

D'Addio (2007) suggests that in many country – among which we find both USA and UK - high skill premia are associated with low level of intergenerational mobility. However, this picture does not fit for Italy, where we can find the coexistence of low labor market rewards for education and a high level of intergenerational persistence. Recent literature posits that in Italy is possible to detect a decrease in the earnings differential between educated and less-educated workers (Lovaglio & Verzillo 2016; Naticchioni et al. 2010) and, in particular, most recent cohorts of high-skilled workers are suffering much heavier earnings penalty with respect to unskilled workers (Naticchioni et al. 2016). Thus, low wage premia for highly qualified workers may disincentive family in investing in their children education and this may explain why education accounts for a limited part of the intergenerational resemblance of earnings, that is more likely to be driven by other mechanisms, for example the importance of family connections and social ties in finding highly rewarded jobs (Raitano & Vona, 2015)

References

Angrist, J. D., & Krueger, A. B. (1992). The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. *Journal of the American statistical Association*, *87*(418), 328-336.

Arellano, M., & Meghir, C. (1992). Female labour supply and on-the-job search: an empirical model estimated using complementary data sets. *The Review of Economic Studies*, *59*(3), 537-559.

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *The Journal of Political Economy*, 1153-1189.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3),

Bidisha, S. H., Das, A., & McFarlane, A. A. (2013). Microeconometric analysis of earnings mobility of immigrants and ethnic minorities: evidence from the UK. *Journal for Labour Market Research*, *46*(1), 35-42.

Björklund, A., & Chadwick, L. (2003). Intergenerational income mobility in permanent and separated families. *Economics Letters*, *80*(2), 239-246.

Björklund, A., & Jäntti, M. (1997). Intergenerational income mobility in Sweden compared to the United States. *The American Economic Review*, *87*(5), 1009-1018.

Björklund, A., & Jäntti, M. (2009). Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality, Oxford University Press, Oxford*.

Björklund, A., Roine, J., & Waldenström, D. (2012). Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity?. *Journal of Public Economics*, *96*(5), 474-484.

Black, S. E., & Devereux, P. J. (2010). Recent developments in intergenerational mobility (No. w15889). National Bureau of Economic Research.

Blanden, J. (2013). Cross‐country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys*, *27*(1), 38-73.

Blanden, J., Gregg, P., & MacMillan, L. (2007). Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education. *The Economic Journal, 117*.

Breen, R., & Goldthorpe, J. R. (2001). Class, Mobility and Merit. The Experience of Two British Cohorts. *European Sociological Review, 17*(2), 81-101.

Büchner, C., Cörvers, F., Traag, T., & Velden, R. V. D. (2012). How do education, cognitive skills, cultural and social capital account for intergenerational earnings persistence? Evidence from the Netherlands. *Netspar Discussion Paper.*

Cavaglia, C. (2015). A new cross-country investigation on the patterns of intergenerational mobility. *Available at SSRN 2624799.*

Cervini-Plá, M. (2015). Intergenerational earnings and income mobility in Spain. *Review of Income and Wealth, 61*(4), 812-828.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. The Quarterly Journal of Economics, 129(4), 1553-1623

Corak, M., & Heisz, A. (1999). The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. *Journal of Human Resources*, 504-533.

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

D'Addio, A. C. (2007). Intergenerational Transmission Of Disadvantage: Mobility Or Immobility Across Generations? A Review Of The Evidence For Oecd Countries. *Oecd Social, Employment And Migration Working Papers No. 52*.

Dahl, M., & DeLeire, T. (2008). The Association between Children's Earnings and Fathers' Lifetime Earnings: Estimates Using Administrative Data. *Institute for Research on Poverty, University of Wisconsin-Madison*.

Dunn, C. E. (2007). The intergenerational transmission of lifetime earnings: Evidence from Brazil. *The BE Journal of Economic Analysis & Policy*, *7*(2).

Ferreira, S. G., & Veloso, F. A. (2006). Intergenerational mobility of wages in Brazil. *Brazilian Review of Econometrics*, *26*(2), 181-211.

Franzini, M., & Raitano, M. (2009). Persistence of inequality in Europe: the role of family economic conditions. *International Review of Applied Economics*, *23*(3).

Franzini, M., Raitano, M., & Vona, F. (2013). The channels of intergenerational transmission of inequality: a cross-country comparison. *Rivista Italiana degli Economisti*, *XVIII*(2).

Goldthorpe, J., & Jackson, M. (2008). Education-based meritocracy: the barriers to its realization. In A. Lareau, & D. Conley, *Social Class. How does it work?*

Grawe, N. D. (2001). Intergenerational mobility in the US and abroad: Quantile and mean regression measures (No. 30-19920 UMI).

Grawe, N. D. (2004). Intergenerational mobility for whom? The experience of high- and low-earning sons in international perspective. *Generational income mobility in North America and Europe*, 58-89.

Gregg, P., Macmillan, L., & Vittori, C. (2014). *Moving towards estimating lifetime intergenerational economic mobility in the UK* (No. 13/332). Department of Economics, University of Bristol, UK.

Haider, S., & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *The American Economic Review*, *96*(4), 1308-1320.

Hirvonen, L. (2010). Accounting for Intergenerational Earnings Persistence: Can We Distinguish Between Education, Skills, and Health?. *SOFI Working Paper*

Hudson, J., & Session, J. (2011). Parental Education, Labor Market Experience and Earnings: New Wine in an Old Bottle. *Economic Letters, 113,* 111-115.

Hussain, M. A., Munk, M. D., & Bonke, J. (2008). How Sensitive is Intergenerational Earnings Mobility to Different Measures. *Danish National Centre for Social Research, Working Paper*, *7,* 2008.

Inoue, A., & Solon, G. (2010). Two-sample instrumental variables estimators. *The Review of Economics and Statistics*, *92*(3), 557-561

Jerrim, J., Choi, A., & Simancas Rodriguez, R. (2016). *Two-Sample Two-Stage Least Squares (TSTSLS) estimates of earnings mobility: how consistent are they?.* In *Survey Research Methods* (Vol. 10, No. 2, pp. 85-102).

Kan, K., Li, I., & Wang, R. H. (2015). Intergenerational income mobility in Taiwan: Evidence from TS2SLS and structural quantile regression. *The BE Journal of Economic Analysis & Policy*, *15*(1), 257-284.

Lefranc, A., & Trannoy, A. (2005). Intergenerational earnings mobility in France: Is France more mobile than the US?. *Annales d'Economie et de Statistique*, 57-77.

Lefranc, A. (2011). Educational expansion, earnings compression and changes in intergenerational economic mobility: Evidence from French cohorts, 1931-1976. *THELMA working paper 2011-11*

Leigh, A. (2007). Intergenerational mobility in Australia. The BE Journal of Economic Analysis & Policy, 7(2).

Lovaglio, P. G., & Verzillo, S. (2016). Heterogeneous economic returns to higher education: evidence from Italy. *Quality & Quantity*, *50*(2), 791-822.

Macmillan, L. (2013). The role of non-cognitive and cognitive skills, behavioural and educational outcomes in accounting for the intergenerational transmission of worklessness *(No. 13-01). Department of Quantitative Social Science-UCL Institute of Education, University College London*.

Mazumder, B. (2005a). Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data. *Review of Economics and Statistics, 87*(2), 235-255.

Mazumder, B. (2005b). The apple falls even closer to the tree than we thought. *Unequal chances: Family background and economic success*, 80-99.

Mocetti, S. (2007). Intergenerational Earnings Mobility in Italy. *The B.E. Journal of Economic Analysis & Policy*, 7(2).

Naticchioni, P., Ricci, A., & Rustichelli, E. (2010). Far away from a skill-biased change: falling educational wage premia in Italy. *Applied Economics*, *42*(26), 3383-3400.

Naticchioni, P., Raitano, M., & Vittori, C. (2016). La Meglio Gioventù: earnings gaps across generations and skills in Italy. *Economia Politica*, *33*(2), 233-264.

Nicoletti, C., & Ermisch, J. F. (2007). Intergenerational earnings mobility: changes across cohorts in Britain. The BE Journal of Economic Analysis & Policy, 7(2).

Núñez, J., & Miranda, L. (2011). Intergenerational income and educational mobility in urban Chile. *Estudios de economía*, *38*(1), 195-221.

Olivetti, C., & Paserman, M. D. (2015). In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850–1940. *The American Economic Review*, *105*(8), 2695-2724

Olivetti, C., Paserman, M. D., & Salisbury, L. (2016). Three-generation Mobility in the United States, 1850-1940: The Role of Maternal and Paternal Grandparents (No. w22094). National Bureau of Economic Research.

Pekkarinen, T., Uusitalo, R., & Kerr, S. (2009). School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, *93*(7), 965-973.

Piraino, P. (2007). Comparable estimates of intergenerational income mobility in Italy. *The BE Journal of Economic Analysis & Policy*, *7*(2).

Piraino, P. (2015). Intergenerational earnings mobility and equality of opportunity in South Africa. *World Development*, *67*, 396-405.

Raitano, M., & Vona, F. (2015). Direct and indirect influences of parental background on children’s earnings: a comparison across countries and genders. *The Manchester School*, *83*(4), 423.-450.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, 393-408.

Solon, G. (1999). Intergenerational mobility in the labor market. *Handbook of labor economics*, *3*, 1761-1800.

Solon, G. (2002). Cross-Country Differences in Intergenerational Earnings Mobility. *The Journal of Economic Perspectives, 16*(3), 59-66.

Ueda, A. (2013). Education and Intergenerational Earnings Transmission: The Case of Japan and South Korea.

Ueda, A. (2013). Intergenerational mobility of earnings in South Korea. *Journal of Asian Economics*, *27*, 33-41

Vogel, T. (2006). Reassessing intergenerational mobility in Germany and the United States: the impact of differences in lifecycle earnings patterns. *SFB*, *649*, 2006-055.

Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *The American Economic Review*, 409-429.

Appendix

Tab. A1: Association between son's and father's earnings. Prime age sons. Fathers and sons observed for 5 years[a]. OLS estimates in the second stage dropping one coefficient at a time in the first stage[b]

	Regressors dropped in the first stage			
	Education dummies	Occupation dummies	Dummy on self-employment	Dummies on region of work
Father's earnings	0.489***	0.466***	0.640***	0.370***
s.e.	[0.055]	[0.057]	[0.064]	[0.057]
Obs	1445	1445	1445	1445
R^2	0.055	0.052	0.069	0.031
R^2 first stage	0.380	0.351	0.282	0.374

[a] When observed in a 5-year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 2009. [b] TSTSLS are carried out: in the first stage father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage sons' log annual gross earnings are regressed on predicted fathers' log earnings, also controlling for sons' year of birth. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset.

Tab. A2: Association between son's and father's earnings. Prime age sons. Fathers and sons observed for 5 years[a]. Rank to rank estimates dropping one coefficient at a time in the first stage[b]

	Regressors dropped in the first stage			
	Education dummies	Occupation dummies	Dummy on self-employment	Dummies on region of work
Father's earnings	0.244***	0.237***	0.262***	0.187***
s.e.	[0.025]	[0.026]	[0.026]	[0.026]
Obs	1445	1445	1445	1445
R^2	0.066	0.063	0.075	0.042
R^2 first stage	0.380	0.351	0.282	0.374

[a] When observed in a 5-year period, fathers and sons are considered, respectively, when aged 40-44 in the period 1980-1988 and 35-39 in the period 2005-2013. When observed in a single year, fathers and sons are considered, respectively, in 1985 and 2009. [b] TSTSLS are carried out: in the first stage father's earnings are imputed regressing log annual gross earnings on dummies on education, occupation, self-employment, region of work; in the second stage sons' log annual gross earnings are regressed on predicted fathers' log earnings, also controlling for sons' year of birth. *** p<0.01, ** p<0.05, * p<0.10. Source: elaborations on AD-SILC dataset