Intergenerational earnings elasticity in Italy along sons’ lifecycle*

Francesco Bloise†, Michele Raitano‡

February 1, 2019

Abstract

This study provides first comparable estimates of the intergenerational earnings elasticity in Italy being able to observe actual fathers-sons pairs. Using high-quality administrative data managed by the Italian Social Security Institute merged with former Italian waves of the European Union Statistics on Income and Living Conditions, we observe a representative sample of sons from 1 to 6 years after they stopped studying (i.e. potential work experience) and each father over a 15 years period starting in the birth year of his son. Results show an high intergenerational earnings elasticity of 0.401 computed 6 years after sons left education. According to subsequent empirical tests, the choice of observing sons by potential work experience rather than by age prevent our estimated elasticity to be strongly affected by lifecycle bias. Hence, after re-estimating the elasticity by observing two subsamples of sons 8 or 11 years after they stopped to study and correcting for the residual estimated lifecycle bias, we report high background-related earnings advantages in Italy finding the intergenerational elasticity to be around 0.45.

JEL Classification: J62, D31, D63. Keywords: intergenerational earnings elasticity; earnings inequality; Italy.

*This is a preliminary draft. Please do not cite or distribute without permission of the authors.
†University of Urbino Carlo Bo. Mail to: francesco.bloise@uniurb.it
‡Sapienza University of Rome. Mail to: michele.raitano@uniroma1.it
1 Introduction

Previous studies that estimate the intergenerational earnings elasticity (IGE) in Italy showed a high degree of intergenerational economic persistence compared to most of other developed countries (Mocetti, 2007; Piraino, 2007; Barbieri et al., 2018). However, due to the lack of longitudinal data which follow two subsequent generations during their entire working careers, previous empirical works were forced to estimate the IGE by imputing fathers’ lifetime earnings through the famous two-sample two-stage least squares (TSTLS) method. This empirical approach is very similar to the two-sample instrumental variable (TSIV) method firstly proposed by Angrist and Krueger (1992) and Arellano and Meghir (1992) and later introduced in the empirical literature on intergenerational mobility by Björklund and Jäntti (1997) to compare the IGE between Sweden and the US. Though the TSTLS method is essential to estimate the IGE in countries for which it is not possible to directly link earnings of sons to those of their fathers, it may produce coefficients which are not perfectly comparable to the ones obtained from an OLS regression of the logarithm of sons’ earnings on their actual fathers’. This happens because, when parental income is imputed, the estimated IGE is assumed to be upward biased (Nicoletti and Ermisch, 2008; Jerrim et al., 2016).

To overcome this potential problem, we provide comparable estimates of the IGE in Italy by directly linking earnings of sons to those of their actual fathers. Using administrative archives managed by the Italian Social Security Institute (INPS) merged with the 2004 to 2008 Italian waves of the European Survey on Income and Living Conditions (IT-SILC), we take sons interviewed no later than 2 years after they achieve their highest educational attainment. Since Italy is characterized by a high share of individuals co-residing with their parents before and right after leaving education, this selection into sample allow us to identify a parent for more than 90 percent of the full sample of respondents taken at that specific point of the life cycle.

This study has several advantages with respect to most of previous evidence provided for Italy. Firstly, as previously stated, we can measure the IGE by directly linking sons to their actual fathers. Secondly, we measure intergenerational mobility by taking earnings of the two generations from high-quality administrative data which are minimally affected by measurement errors and attrition. Secondly, we minimize right-hand side measurement errors in the estimation of the IGE, by exploiting the longitudinal dimension of the INPS archives to measure averaged earnings of fathers followed for 15 years since the birth year of their sons (12 years on average). Thirdly, we observe sons
over the early stages of their career, from 1 to 6 years after their highest educational attainments. This selection into sample allows us to estimate the IGE by observing all sons approximately at the same point of their working career, despite their age, to reduce the amount of lifecycle bias. Eventually, we adapt the error-in variables model firstly proposed by (Haider and Solon, 2006) to our selection of sons into sample to analyze the career-earnings profile of a large sample of Italian males which left education between 1978 and 1984, followed for the subsequent 30 years.

According to the selection rules adopted to obtain father-son pairs, we find important results showing that the IGE in Italy is high and increases along the sons’ career, with an estimated coefficient of 0.401 obtained observing sons 6 years after they stopped studying. Though this initial result may suggest that the degree of intergenerational mobility in Italy is very low, it can still be a lower bound given that sons are observed at early stages of their career. However, when we exploit the error-in-variables model proposed by Haider and Solon (2006) to estimate the so called “forward regression” of yearly earnings on a proxy of lifetime earnings, we show that the estimated IGE 6 years after sons left education is only marginally affected by lifecycle bias. This results derives from the fact that our selection rules allow us to observe sons approximately at the same point of their career. Hence, we are basically reducing that fraction of lifecycle bias related to the fact that, at each specific age, more educated individuals have less work experience since they left education later than other workers the same birth cohort.

In any case, to obtain a comparable measure of intergenerational persistence, we re-estimate the IGE observing two subsamples of sons with 8 or 11 years of potential work experience. In this case, we obtain an estimated IGE of 0.373 and 0.429 respectively. Given that the estimated differences between the latter two coefficients and the one obtained 6 years after the full sample of sons left education are highly comparable with the career-earnings profile estimated using the “forward regression”, we correct our estimated measures for the residual lifecycle bias reporting a IGE of 0.44/0.45.

According to these final results, we show that the estimated IGE is high and similar in magnitude to the ones obtained by Piraino (2007), Mocetti (2007) and Barbieri et al. (2018) exploiting the TSTSL method. On the contrary, our IGE is consistently higher than that provided by (Acciari et al., 2016) that proxy fathers’ permanent income using only two yearly observations. Accordingly, our result suggests that most of previous estimates obtained by using the TSTSL method are likely not to be significantly upward biased, despite the use of imputed fathers’ earnings. Italy is a very low mobility
country likewise the US where previous evidence show the IGE to be around 0.4/0.5 (Solon, 1992; Zimmerman, 1992).

The structure of the work is the following. Section 2 describes the empirical framework associated to the intergenerational transmission of inequality and previous evidence for Italy in cross–country comparison. Section 3 presents the data and the selection of sons and fathers into the final samples. Section 4 discusses the results obtained in terms of IGEs following sons from 1 to 6 years after they left education. Section 4 presents the results of the “forward regression” of yearly earnings on lifetime earnings for a representative sample of Italian workers observed either for each age or year of potential experience. Section 6 shows final results obtained after observing two subsamples of sons 8 or 11 years after their highest educational attainments and correcting for the residual estimated lifecycle bias. Section 7 concludes.

2 Empirical framework and previous literature

In the last few decades, many empirical studies carried out by economists and social scientists analyze to what extent economic advantages are transmitted from one generation to the next\textsuperscript{1}. In this literature, many indicators have been used to summarize the degree of intergenerational mobility or persistence. The most common index used by economists to measure degree of intergenerational earnings persistence is the IGE which can be estimated by regressing the logarithm of children’s lifetime earnings on the logarithm of parents’ as in the following equation:

$$\ln y_c = \alpha + \beta \ln y_p + \epsilon_i$$

(1)

According to this specific measure of persistence, a country is completely mobile when the estimated $\beta$ equals 0 and the higher the estimated $\beta$, the higher the degree of intergenerational economic persistence.

Since background-related earnings advantages last over the whole working career of individuals, empirical studies analyzing the persistence of earnings aim to estimate the IGE by considering lifetime rather than yearly earnings of the two generations. However, lifetime earnings of children and their parents are usually not available because of the lack of panel data following two subsequent generations during their entire working

\textsuperscript{1}For more details about the approaches used by empirical researchers to estimate mobility across generations, see Black and Devereux (2011)
career. This is the reason why researchers have to face several measurement issues when estimating the IGE. For instance, earlier studies estimating the intergenerational earnings persistence in the US, report IGEs around 0.2 and describe the US as a very mobile society according to the ideal of “American Dream” (Becker and Tomes, 1986; Behrman and Taubman, 1986). Subsequent works demonstrate that previous estimates were substantially downward biased due to right-hand side measurement errors. For instance, Solon (1992) and Zimmerman (1992), using better data to average fathers’ earnings over four or five years, estimate the IGE to be around 0.4/0.5, consistently higher than the coefficient estimated in the 80s. Subsequently, Mazumder (2005) and Chen et al. (2017) in two studies on the US and Canada respectively show that even using 5-years averaged earnings may lead to an underestimation of the IGE since transitory shocks which cause right-hand side measurement errors are likely to be extremely persistent.

Turning to left-hand side measurement errors, despite they are usually not assumed to cause any bias in an OLS regression, they may cause the so-called lifecycle bias lowering the estimated $\beta$ if earnings of children are observed when they are too young. This is because the age-earnings profile is steeper for high-skilled workers as their earnings growth rate is often higher than that of other individuals at the same age. Moreover, at any given age, individuals have different years of work experience according to their education, since tertiary graduated individuals usually enter the labor market several years after low-skilled workers. For the two reasons just described, the earnings dispersion of young workers generically selected by age is likely to be consistently lower than the dispersion of their lifetime earnings. Therefore, the estimated $\beta$ from equation 1 is equal to the following expression:

$$p\lim \hat{\beta} = \frac{Cov(y_{ci}, y_{pi})}{Var(y_{pi})} = \frac{\rho_{cp} \sigma_{c}}{\sigma_{p}}$$

where $\sigma_{cp}$ is the correlation between children’s lifetime earnings and parents’, $\sigma_{c}$ is the standard deviation of children’s lifetime earnings and $\sigma_{p}$ is the standard deviation of parents’. The estimated $\beta$ is downward biased if we include $y_{t}^{yc}$ rather than $y_{ci}$ in the right-hand side of equation 1, where $y_{t}^{yc}$ are yearly earnings of the young child, with $\sigma_{yc} < \sigma_{c}$.

Nevertheless, two empirical estimations of the lifecycle earnings variation made by Haider and Solon (2006) and Böhlmark and Lindquist (2006) for the US and Sweden respectively, suggest that this specific source of downward bias can be greatly reduced.
by selecting the second generation around median ages. This happens because the difference between earnings of individuals at mid-careers should be the closest to the difference between their lifetime earnings. Subsequently Nybom and Stuhler (2016), while confirming that the best way to minimize the lifecycle bias is to take individuals at median ages, warn that it can be very difficult to identify the specific age at which children’s lifetime earnings are perfectly approximated for all individuals as the age-earnings profile may be worker specific.

2.1 Previous evidence for Italy in cross-country comparison

Previous empirical studies for Italy carried out by Mocetti (2007), Piraino (2007), and Barbieri et al. (2018) estimate the IGE by exploiting the TSTSLS methodology due to the impossibility to directly link children observed at their mid-careers to their actual parents. Following the TSTSLS method and according to microdata at their disposal, these scholars exploit retrospective time-invariant socioeconomic information of Italian fathers (e.g. educational achievements, occupational status, region of residence, sector of activity) recalled by their sons and an auxiliary sample of pseudo-fathers to get a prediction of lifetime fathers’ earnings and estimate the IGE\(^2\). Using this estimation methodology they obtain an IGE between 0.44 and 0.50 depending on the number of predictors used to impute fathers’ lifetime earnings, the number of years the two generations are observed, the data used and the income definition adopted (i.e. net or gross of taxes and deductions). According to these results, Italy is reported to have a higher IGE compared to most of other developed countries such as Norway, Sweden, Canada, France and Germany and close to the ones estimated for Spain, the US and the UK.

As showed in Table 1, the TSTSLS has been often used to estimate IGEs in countries such as France, Spain and Italy where it is not possible to directly link earnings of sons to those of their fathers. However this estimation methodology may produce coefficients which are not perfectly comparable to the ones obtained from an OLS regression of the logarithm of sons’ earnings on their actual fathers’. More specifically, the TSTSLS estimator produces estimated IGEs which are usually considered upward biased, as the standard deviation of imputed fathers earnings is by construction lower than the standard deviation of actual fathers earnings and may not predict all components of

\(^2\)As in most of empirical studies estimating the IGE, they focus on father-son pairs to avoid the selection bias related to the low labor force participation of women.
earnings which are correlated across generations (Olivetti and Paserman, 2015). For instance Blanden (2013), suggests to re-scale all estimated IGEs obtained by using the TSTLS method by a factor of 0.75 according to the results reported by Björklund and Jäntti (1997) that estimate the IGE in the US by using both the OLS and the TSIV estimator.\footnote{However, it is well acknowledged by Blanden (2013) that it is not possible to generalize the bias found by Björklund and Jäntti (1997) for the US to studies focused on other countries.}

Table 1: Intergenerational earnings elasticity: cross-country comparison

<table>
<thead>
<tr>
<th>Country</th>
<th>Source</th>
<th>Empirical approach</th>
<th>IGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>Zimmerman (1992)</td>
<td>OLS</td>
<td>0.54</td>
</tr>
<tr>
<td>Italy</td>
<td>Barbieri et al. (2018)</td>
<td>TSTLS</td>
<td>0.44-0.50</td>
</tr>
<tr>
<td>Italy</td>
<td>Mocetti (2007)</td>
<td>TSTLS</td>
<td>0.50</td>
</tr>
<tr>
<td>Italy</td>
<td>Piraino (2007)</td>
<td>TSTLS</td>
<td>0.44</td>
</tr>
<tr>
<td>UK</td>
<td>Gregg et al. (2017)</td>
<td>OLS</td>
<td>0.43</td>
</tr>
<tr>
<td>US</td>
<td>Solon (1992)</td>
<td>OLS</td>
<td>0.42</td>
</tr>
<tr>
<td>Spain</td>
<td>Cervini-Plá (2015)</td>
<td>TSTLS</td>
<td>0.42</td>
</tr>
<tr>
<td>France</td>
<td>Lefranc and Trannoy (2005)</td>
<td>TSTLS</td>
<td>0.40</td>
</tr>
<tr>
<td>Germany</td>
<td>Schnitzlein (2016)</td>
<td>OLS</td>
<td>0.39</td>
</tr>
<tr>
<td>Canada</td>
<td>Chen et al. (2017)</td>
<td>OLS</td>
<td>0.32</td>
</tr>
<tr>
<td>Sweden</td>
<td>Björklund and Chadwick (2003)</td>
<td>OLS</td>
<td>0.24</td>
</tr>
<tr>
<td>Norway</td>
<td>Bratberg et al. (2005)</td>
<td>OLS</td>
<td>0.13</td>
</tr>
</tbody>
</table>

\textbf{Notes:} To maximize the degree of comparability, almost all of reported IGEs which use the OLS estimator have been estimated after averaging fathers’ earning using 4/5 yearly observations. Three exceptions are the studies by Chen et al. (2017), that take fathers with positive earnings in at least 10 years, Schnitzlein (2016) who observes fathers for ten years excluding those with less than 5 positive earnings observations and Gregg et al. (2017) that take two parental income observations. The two estimates by Barbieri et al. (2018) differ from the number of years used to measure sons’ earnings (i.e. single year or 5-years average).

3 Data and sample selection

Following Barbieri et al. (2018), we estimate the intergenerational earnings persistence in Italy by exploiting the so-called AD-SILC dataset, built merging the 2004 to 2008
waves of the IT-SILC – a specific version of the Italian EU-SILC dataset\textsuperscript{4} – with high-quality administrative data managed by the Italian Social Security Institute (INPS). The latter records employment characteristics and gross earnings of Italian workers from the moment they entered the labor market together with other demographic characteristics (e.g. gender, year of birth, region of residence) and detailed information on every job relationship that individuals experience in a specific year (e.g. duration, fund where workers pay contributions).

The IT-SILC waves include other important information at both the household and personal level which is absent in administrative archives and is fundamental for our empirical purposes. More in detail, it provides information about the specific relation between the respondent and the person of reference of the household and the year when respondents obtain their highest educational attainment. Additionally, all respondents are asked if they are still in education at the time of the interview. Therefore, by combining the latter two information reported in the IT-SILC, it is possible to identify the specific year when all individuals left education.

Similarly to the selection approach proposed by Raitano and Vona (2018) to analyze the association between fathers’ years of education and sons’ earnings over their entire working career, we select samples of sons followed by \textit{potential experience} defined as years of distance from when they attained the highest educational degree and finished to study. Consistently with most studies estimating the IGE, we focus on father-son pairs to avoid the potential selection bias arising from the low labor force participation of women. To link sons to their actual fathers, we consider sons aged 16 to 30 in the 2004 to 2008 IT-SILC waves conditionally to have left education (i.e. having achieved their highest educational level and not being anymore in education) no later than 2 years before the year of the interview. This means that selected sons interviewed in the IT-SILC wave of 2004 left education between 2002 and 2004; selected respondents interviewed in the wave of 2005 between 2003 to 2005 and so on for the subsequent IT-SILC waves. Therefore, since the INPS administrative archives record individuals earnings until the end of 2014, we are able to observe all selected sons’ with positive earnings from 1 to 6 years of \textit{potential experience}, despite their age. Obviously, at any given \textit{potential experience}, age differences will simply depend on educational differences as more educated sons leave education at a higher age.

For each selected son, we can exploit information provided by the IT-SILC on his

\textsuperscript{4}The Italian version of the EU-SILC waves differs from the baseline dataset only for the inclusion of additional country-specific variables.
parental relationships with members of the household to identify his father. Since Italy is a country where almost all young individuals live with their parents at least until they leave education, this selection rule allows us to link the 90.2% of all respondents that left education no later than 2 years before the year of the interview. It is important to note that in Italy about 4% of all children who were born in the 80s and 90s did not grow up with their parents\textsuperscript{5}. This means that the percentage of sons that we are able to link to their actual fathers is extremely high.

Tough we can directly link to their actual fathers the vast majority of sons that left education no later than 2 years before the interview, it may be important to compare some summary statistics regarding the characteristics of sons in the selected sample to the same characteristics observed in the full sample of individuals at the same point of the lifecycle. For instance, male workers in the two samples are highly comparable in terms of gross earnings as it is shown in Figure A.1 in appendix A. More in detail, on average sons’ earnings increases quickly in the first 4 years after they left education to stabilize starting from the fifth year. Similarly, the earnings dispersion measured by the standard deviation seems to increase until the fifth year in both the selected and the full sample. Table A.1 summarizes many other socio-economic characteristics of sons comparing again our selected sample to the full sample of male workers interviewed no later than 2 years after they left education whose earnings profile is observed until potential experience \textsuperscript{6}. As it is clear from the table, the observed distribution of characteristics is similar in the two samples since selected sons are highly comparable to the full sample of sons in terms of age when leaving education (20.31 vs. 20.49 years old), weeks of work experience when leaving education (26.16 vs. 29.39), years of education (13.98 vs. 14.11) and fund where workers pay contributions.

Regarding the first generation, fathers are observed for 15 years since their sons were born. In order to decrease the incidence of right-hand measurement errors causing a downward bias in the estimated \( \beta \), we include in our baseline model only fathers with at least 4 positive earnings observations. We thus compute an average according to the number of positive observations available in the 15 years period. Given that many fathers have positive earnings observations in most of the 15 years considered (54.20% of all selected fathers have positive earnings observations in all the period) with an average of 12.96 observations, our measure of fathers’ earnings should be a very good proxy of fathers lifetime earnings. In any case, in Section 4 additional estimates of

\textsuperscript{5}We obtain this percentage by observing a representative sample of Italian children aged 1 to 14 taken from the Bank of Italy’s Survey on Household Income and Wealth (SHIW).
the IGE are provided by considering a different number of observations to assess the robustness of our baseline estimates to right-hand measurement errors.

Table 2 reports summary statistics of our final selected sample for both sons and fathers. Our main measure of economic outcome of the two generations is computed as the sum of all CPI adjusted earnings gross of personal taxes and pension contributions received by employees and self-employed workers. In the first years after the sons obtain their highest educational attainment, their earnings are obviously lower and less dispersed than fathers’ since they are measured when the latter were around 38 years old on average. In any case, as we could expect, differences in terms of earnings level and dispersion become lower as we move along the sons’ working career.

Since our selected sons are only 21.48 years old on average 1 year after they left their fathers. Our main measure of economic outcome of the two generations is computed as the sum of all CPI adjusted earnings gross of personal taxes and pension contributions received by employees and self-employed workers. In the first years after the sons obtain their highest educational attainment, their earnings are obviously lower and less dispersed than fathers’ since they are measured when the latter were around 38 years old on average. In any case, as we could expect, differences in terms of earnings level and dispersion become lower as we move along the sons’ working career.

Table 2: Summary statistics of sons observed by potential experience and their actual fathers.

<table>
<thead>
<tr>
<th>Potential experience</th>
<th>Statistic</th>
<th>Earnings Sons</th>
<th>Earnings Fathers</th>
<th>Age Sons</th>
<th>Age Fathers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean</td>
<td>12382.88</td>
<td>24155.01</td>
<td>21.48</td>
<td>38.00</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>8294.09</td>
<td>9997.78</td>
<td>3.26</td>
<td>5.17</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>566</td>
<td>566</td>
<td>566</td>
<td>566</td>
</tr>
<tr>
<td>2</td>
<td>Mean</td>
<td>15431.82</td>
<td>23654.71</td>
<td>22.33</td>
<td>37.86</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>8764.11</td>
<td>9999.63</td>
<td>3.26</td>
<td>5.19</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>648</td>
<td>648</td>
<td>648</td>
<td>648</td>
</tr>
<tr>
<td>3</td>
<td>Mean</td>
<td>17002.94</td>
<td>23700.75</td>
<td>23.29</td>
<td>37.85</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>9416.93</td>
<td>10194.70</td>
<td>3.30</td>
<td>5.10</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>707</td>
<td>707</td>
<td>707</td>
<td>707</td>
</tr>
<tr>
<td>4</td>
<td>Mean</td>
<td>18445.34</td>
<td>23498.71</td>
<td>24.16</td>
<td>37.93</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>9892.46</td>
<td>10035.78</td>
<td>3.32</td>
<td>5.28</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>739</td>
<td>739</td>
<td>739</td>
<td>739</td>
</tr>
<tr>
<td>5</td>
<td>Mean</td>
<td>19867.97</td>
<td>23599.96</td>
<td>25.22</td>
<td>37.99</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>10744.28</td>
<td>10177.38</td>
<td>3.32</td>
<td>5.25</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>739</td>
<td>739</td>
<td>739</td>
<td>739</td>
</tr>
<tr>
<td>6</td>
<td>Mean</td>
<td>20334.37</td>
<td>23500.93</td>
<td>26.11</td>
<td>37.97</td>
</tr>
<tr>
<td></td>
<td>Sd</td>
<td>10860.47</td>
<td>9954.44</td>
<td>3.29</td>
<td>5.22</td>
</tr>
<tr>
<td></td>
<td>Obs.</td>
<td>750</td>
<td>750</td>
<td>750</td>
<td>750</td>
</tr>
</tbody>
</table>

Notes: All earnings are CPI adjusted in 2012 Euro. At any given potential experience, only workers with positive earnings and their fathers with at least 4 positive earnings observations recorded in a 15 years period starting when sons were born are considered. Source: Authors’ elaborations on the AD-SILC dataset.
education and 26.11 years old 6 years after, it is clear from Table 2 that we are not able to follow the selection rules suggested by Haider and Solon (2006) to minimize the lifecycle bias arising when lifetime earnings of the second generation are not available. Nevertheless, the specific selection rules we adopt in this work allow us to reduce one of the two aspects described in Section 2 related to the lifecycle bias. More in detail, despite of sons’ age, our estimated IGEs are obtained by taking all workers of the second generation approximately at the same point of their career since estimates are obtained observing sons by potential experience. This means that we are not estimating the intergenerational earnings persistence by generically selecting young sons within a given age rage and thus at different points of their working career. Hence, the only potential source of lifecycle bias in our estimated coefficients is related to existing differences in the earnings growth rate over the working career which are likely to persist after potential experience.  

4 Estimates of the intergenerational elasticity in Italy

In this section, we present our estimated IGEs for Italy, according to the selection rules previously described. To obtain our main measure of intergenerational earnings persistence in Italy, we estimate equation 1 presented in Section 2 by considering only sons and their fathers. As a proxy of fathers’ lifetime economic outcomes we use averaged earnings calculated by considering only fathers with at least 4 positive observations over a 15 years period. On the contrary, sons’ earnings are selected by potential experience to obtain different estimated IGEs at different points of their careers. As a general control variable, we include in our model the year dummies to exclude variability related to the year in which sons’ earnings are observed (model 1 henceforth). An additional model is estimated by controlling also for the number of sons’ weeks of work experience gained before they left education (model 2 henceforth). However, we do not expect very large differences in the results obtained from the 2 models because, as we know from Table A.1 in the Appendix A that sons have on average only about only about 26 weeks of work experience when leaving education.  

The estimated IGEs are plotted in Figure 1 for both the two models. The estimated

---

6Observe that we do not control for a polynomial of sons’ age as it is usually done in the literature that selects individual by age since, at any given potential experience, sons’ age is related to their human capital, an important channel through which fathers’ economic status is transmitted to their sons (Becker and Tomes, 1979, 1986).
Figure 1: Estimated elasticity of sons’ yearly earnings with respect to fathers’ by sons’ potential experience

Notes: At any given potential experience, those with positive earnings and their fathers with at least 4 positive earnings observations are considered. In model 1 we estimate the IGE by controlling for the year dummies. In model 2 we control for year dummies and the sons’ work experience gained before leaving education. Source: Authors’ elaborations on the AD-SILC dataset.

IGE using model 1, is around 0.166 taking sons 1 year after they left education and, as we could expect, it increases as we move over the sons’ career reaching the value of 0.401 at the latest potential experience considered (Figure 1). Therefore, according to these results the intergenerational earnings persistence appears to be considerably high in Italy though we are considering sons’ at early stages of their career. However, if we compare our maximum estimated IGE at potential experience 6 to previous evidence for Italy obtained using the TSTSLS method, we can see that the degree of intergenerational mobility in Italy could appear to be slightly higher than previously suggested. However, it is not possible to say in this section whether these differences derive from the fact that our estimated IGE is downward biased because of left-hand side measurement errors or if previous studies report upward biased estimates of the IGE due to the TSTSLS methodology adopted. In any case, we will try to answer to
this important question in subsection 6.

When we control also for the sons’ weeks of work experience gained before leaving education, the estimated IGEs are slightly higher than those obtained from the baseline model 1 (dashed line in Figure 1). This estimated difference in more relevant considering the first years after sons’ educational achievements and becomes lower the closer we get to potential experience 6. This result suggests that especially at potential experience 1 there are some differences in the degree of actual work experience deriving from the fact that some selected individuals start working before leaving education. In any case, as this estimated difference between model 1 and model 2 becomes smaller as we move over the sons’ career, we choose model 1 as our baseline.

To test the sensitivity of our estimated IGEs to right-hand side measurement errors, we compare different estimated elasticities by varying the number of years used to average fathers’ earnings in order to verify whether and to what extend the IGE is influenced by this kind of methodological choices. More in detail, our baseline estimates obtained by excluding fathers with less than 4 positive earnings observations in the 15 years period considered is compared with alternative estimates produced: 1. by observing fathers in an unique year when sons were 14 years old; 2. by excluding all those fathers with less than 5 positive earnings observations in the period when sons were 1 to 14 years old.

Figure 2 shows that our baseline estimates appear to be highly comparable to the ones obtained by considering fathers with at least 5 positive observations. This result suggests that our measure of fathers’ earnings is basically robust to right-hand side measurement errors probably because, as we already stated in Section 3, the selected fathers have positive earnings in most of the 15 years period. On the contrary, estimates produced by measuring fathers’ earnings in an unique year when sons were 14 years old are considerably lower than our baseline. The attenuation bias is particularly relevant at potential experience 6 when the estimated IGE is about 50% lower than our baseline. Therefore, confirming previous evidence for the US (Solon, 1992; Zimmerman, 1992; Mazumder, 2005), it is extremely important to take into account right-hand measurement errors in order to avoid an overestimation of the degree of intergenerational mobility. This is probably the reason why (Acciari et al., 2016) find a very low elasticity of 0.22 by averaging fathers’ income using only two yearly observations. Their result is very similar to the one we obtain at potential experience 6 when only one positive

\[ \text{We also try to estimate the IGE by excluding those fathers with less than 6 positive earnings observations with no significant differences from our baseline estimate.} \]
Figure 2: Estimated elasticity of sons’ yearly earnings with respect to fathers’: sensitivity to attenuation bias

Notes: Alternative estimated IGEs are provided by averaging fathers’ earnings using a different number of positive earnings observations: 1 positive earnings observation when sons were 14 years old (dashed line); 4 or more positive earnings observations in the 15 years period (solid line); 5 or more positive earnings observations in the 15 years period (dotted line). In all the estimates we control for the year dummies. Source: Authors’ elaborations on the AD-SILC dataset.

earnings observations of fathers is considered.
5 Empirical estimates of the lifecycle bias

In this section we adapt the textbook error-in-variables model to assess the career-earnings profile in Italy, following the approach firstly proposed by Haider and Solon (2006) in their study on the US and subsequently applied by Böhlmark and Lindquist (2006) and Chen et al. (2017) to Swedish and Canadian data respectively. Given that we are only interested in the bias arising from left-hand side measurement error, we will from now use the simplifying assumption of no right-hand side measurement errors, though we are aware that even if fathers’ earnings are averaged using many yearly observations, we are not able to observe “true” lifetime earnings of fathers.

Beside giving information on the specific amount of bias related at any given age or potential experience to the use of yearly earnings instead of lifetime earnings, this method allows us to evaluate the potential downward bias affecting our highest estimated elasticity obtained at potential experience 6. The method introduced by Haider and Solon (2006) consists in evaluating the bias associated to the use of yearly instead of lifetime earnings by regressing the former on the latter having at disposal a proper dataset which follows individuals approximately over their entire working career. Ideally, we would like to estimate the IGE by means of OLS:

\[ y_s^i = \alpha + \beta y_f^i + \epsilon_i \] (3)

where \( y_s^i \) and \( y_f^i \) are the logarithm of the son’s and father’s lifetime earnings respectively, \( \epsilon_i \) is the classical disturbance and the estimated coefficient \( \beta \) is the IGE. Since we can not directly observe lifetime earnings of sons, we are likely to obtain biased estimates of the IGE due to left-hand side measurement errors. However, we can directly measure a proxy of lifetime earnings for a representative sample of Italian workers thanks to the large longitudinal dimension of the AD-SILC dataset. We can thus regress their yearly earnings on their lifetime earnings according to the so-called “forward regression” of \( y_{it} \) on \( y_i \):

\[ y_{it} = \theta_i y_i + \omega_{it} \] (4)

where \( y_{it} \) are yearly earnings which can be observed either at a given age or potential experience. Therefore, by assuming that the estimated relation between yearly and lifetime earnings is the same among our selected sons and substituting the 4 into the 3
we obtain:

\[ y_{it} = \theta_t \alpha + \theta_t \beta y_{it}^f + (\omega_{it} + \theta_t \epsilon_i) \]  

Equation 5 is very useful to give us properly information on the amount of bias affecting the estimated IGE in Italy when lifetime earnings of sons are not available. In particular, assuming that \( \text{Cov}(\omega_{it}, y_{it}^f) = 0 \), the lifecycle bias equals \( \theta_t \) and it disappears only when this estimated coefficient equals 1. On the contrary, the estimated \( \beta \) is likely to be downward biased when sons are observed when they are too young (\( \theta_t < 1 \)) and upward biased after a given median age (\( \theta_t > 1 \)). Therefore, according to this empirical approach, it could be theoretically possible to correct a biased estimated \( \beta \) by simply exploit the estimated \( \theta_t \) corresponding to the specific age at which sons are observed.

In any case, Haider and Solon (2006) suggest that it is not correct to use an estimated \( \theta_t \) for a specific country, to correct the estimated \( \beta \) in a different country as the age-income profile is likely to differ from a specific labor market to another. Likewise, the age-income profile might change within the same country from one cohort to another. Eventually, Nybom and Stuhler (2016) and Chen et al. (2017) point out that another source of bias affecting the estimated \( \theta_t \) may arise if \( \text{Cov}(\omega_{it}, y_{it}^f) \neq 0 \). Nevertheless, though all previous studies are perfectly aware of all possible biases related to the estimated \( \theta_t \), they still acknowledged that this is the best way available so far to choose the age at which sons should be taken when it is not possible to measure their lifetime earnings.

In this work, we estimate the “forward regression” by selecting a representative sample of male Italian workers that left education from 1978 to 1984, followed from the subsequent 30 years (i.e. from 1979 to 2008 those that left education in 1978; from 1980 to 2009 those that left education in 1979 and so on) for a final overall sample of 4606 male workers for which we are able to get a proxy of their lifetime earnings. Following Haider and Solon (2006) and Chen et al. (2017) we compute a proxy of lifetime earnings by averaging earnings of workers with at least 10 years positive observations taken from the 6th to 25th year of distance form their highest educational achievement. Though, we are not able to measure their “true” lifetime earnings since we are not considering some potential earnings obtained at the beginning of the career or in the last years before retirement, we can consider our measure of lifetime earnings as a good proxy of

\(^8\) Alternative results are presented in Figure A.2 in appendix A by estimating the “forward regression” considering individuals with at least either 8 or 12 positive earnings observations with no significant differences with respect to the baseline estimate.
“true” lifetime earnings as our selected Italian workers are observed for 17.74 years on average.

The results from the estimated “forward regression” are presented in the left-end graph in Figure 3 for each age between 20 and 55\(^9\). As in previous studies, we can see

Figure 3: Forward regression of log annual earnings observed by age or potential experience on log lifetime earnings.

Notes: Workers that left education between 1978 and 1984 are followed for the 30 subsequent years. Their proxy of lifetime earnings has been calculated by averaging at least 10 positive observations in a 20 years period which begins at the 6th year after they left education and ends with the 25th. In all the estimates we control for the year dummies. Source: Authors’ elaborations on the AD-SILC dataset.

that the estimated \(\theta_a\) is lower than 1 when individuals are very young; it becomes greater than 1 later during the lifecycle; it decreases again in the later stages of the individuals working career. According to our estimated \(\theta_a\) the age at which the estimated parameter is the closest to one is between 33 and 34 years old. This optimal age range is only slightly higher than the one suggested in the study by Haider and Solon (2006) for the US (i.e. around 32 years old).

When we re-estimate the “forward regression” by observing yearly earnings of male individuals by potential experience rather than by age, we obtain for each potential experience between 1 and 30, many estimated \(\theta_{pe}\) which are plotted in the right-end graph in Figure 3. This second graph shows that left-hand side measurement errors are minimized between potential experience 12 and 14. Moreover, at potential experience 6 the estimated downward bias is only around 9%. According to this result, we can

\(^9\)Even if results are not presented in the right-end graph in Figure 3, we observe individuals also after 60 years old if they have no more than 30 years of potential experience
argue that our choice of selecting sons by potential experience prevent our estimated elasticity at potential experience 6 to be strongly affected by downward bias.

Another interesting result derives from the fact that our selection rule seems to reduce, at least partially, the downward bias related to the young age at which sons are taken. For instance, given than our selected sons are on average 26 years old at potential experience 6, the downward bias would have been around 17% if the second generation were generically selected by age.

6 Corrected measures of intergenerational earnings persistence

The results from the estimation of the “forward regression” suggest that our estimated IGE is unlikely to be strongly downward biased taking sons at potential experience 6. In this section, using the estimated coefficients in equation 4, we correct our highest estimated IGE to obtain measures which are comparable with those previously presented for other countries.

Firstly, if we use the estimated parameter from the “forward regression” of the logarithm of yearly earnings on lifetime earnings obtained at potential experience 6, we obtain a corrected IGE of 0.440 (Table 4, panel A). However, according to what suggested by Haider and Solon (2006), Nybom and Stuhler (2016) and Nybom and Stuhler (2017), we are perfectly aware that the estimated coefficients in the “forward regression” can be exploited to correct estimates of persistence affected by the lifecycle bias only if the potential experience-earnings profile of individuals observed during their entire career is comparable to that of selected sons. In other terms, one should not use the estimated θ, for a given country to correct downward biased IGE of other countries. Moreover, even estimating the “forward regression” for the same country for which we want to analyze the degree of the intergenerational persistence, other sources of bias may arise if the age or potential experience-earnings profiles vary from one cohort to another.

In our case, we are not worried by the first source of bias as we properly estimate the “forward regressions” considering male Italian workers to correct the estimated IGEs of Italian sons. On the contrary, there could be a source of bias deriving from the fact that we estimate the “forward regression” for male Italian workers that left education from 1978 to 1984, thus about 20/25 years before the selected sons. To
verify whether this potential source of bias is affecting our corrected IGE, we test the
goodness of our estimated $\theta_{pe}$ by correcting the estimated IGE at potential experience
1 and additional estimates at potential experience 8 and 11. The latter new estimates
have been obtained on two subsample of sons that left education between 2002 and 2006
and between 2002 and 2003 respectively. Descriptive statistics for these newly built
subsamples are presented in table 3. Obviously, as potential experience increases sons

Table 3: Summary statistics of the subsamples of sons observed at potential experience
8 and 11

<table>
<thead>
<tr>
<th>Potential experience: 8</th>
<th>Potential experience: 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings</td>
<td>Age</td>
</tr>
<tr>
<td>Mean</td>
<td>21137.69</td>
</tr>
<tr>
<td>Sd</td>
<td>12142.35</td>
</tr>
<tr>
<td>Obs.</td>
<td>608</td>
</tr>
</tbody>
</table>

Notes: The two subsamples are obtained by considering individuals that left education from 2002
and no later than 2006 in the potential experience 8 case and between 2002 and 2003 in the potential
experience 11 case. Monetary values are CPI adjusted at 2012 prices. Source: Authors’ elaborations
on the AD-SILC dataset.

get older, both mean earnings and standard deviation increases, and, as we expected,
sample dimensions become smaller $^{10}$ The latter result derives from the fact that, as
we are selecting sons interviewed no later than 2 years after their highest educational
achievements, we are forced to exploit the earlier waves of the AD-SILC dataset only.

These newly estimated measures of persistence are presented in Table 4. The estimated
IGE is 0.373 at potential experience 8 (Panel C) and 0.429 at potential experience
11 (Panel D). Reassuringly, even if we are using smaller samples to estimate IGEs at
potential experience 8 and potential experience 11, our resulted are highly compatible
with the potential experience-earnings profile showed in the left-hand graph in Figure
3. Accordingly, when we correct the estimated IGEs obtained at potential experience
1 (Panel A), potential experience 6 (Panel B), potential experience 8 (Panel C)
and potential experience 11 (Panel D) using the corresponding estimated $\theta_{pe}$, we ob-
tain very similar corrected IGEs in all 4 cases considered. This means that estimated
$\theta_{pe}$ from the “forward regressions” are applicable to our selected sons. In more detail,

$^{10}$As in our case, many empirical studies that estimate the IGE are forced to use very small samples
due to data limitations. See for instance the works by Björklund and Jäntti (1997), Mazumder (2005)
and Schnitzlein (2016).
Table 4: Comparable estimated IGEs

<table>
<thead>
<tr>
<th>Panel A: Full sample of sons at potential experience 1</th>
<th>Panel B: Full sample of sons at potential experience 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated IGE</td>
<td>Estimated (1-θ_{pe})</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>0.166</td>
<td>0.627</td>
</tr>
<tr>
<td>(0.0101)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>Obs. 566</td>
<td>1649</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Subsample of sons at potential experience 8</th>
<th>Panel D: Subsample of sons at potential experience 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated IGE</td>
<td>Estimated (1-θ_{pe})</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>0.373</td>
<td>0.171</td>
</tr>
<tr>
<td>(0.080)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Obs. 608</td>
<td>3600</td>
</tr>
</tbody>
</table>

Notes: Estimates in panel A are obtained by considering the full sample of sons that left education starting from 2002 and no later than 2008. Their positive earnings are observed 1 year after their highest educational achievements. Estimates in panel B are obtained by considering the full sample of sons that left education starting from 2002 and no later than 2008. Their positive earnings are observed 6 years after their highest educational achievements. Estimates in panel C are obtained by considering the subsample of sons that left education starting from 2002 and no later than 2006. Their earnings are observed 8 years after their highest educational achievements. Estimates in panel D are obtained by considering the subsample of sons that left education starting between 2002 and 2003. Their earnings are observed 11 years after their highest educational achievements. In all cases we control for the year dummies. Standard errors in parantheses. Authors’ elaborations on the AD-SILC dataset.

Using the results obtained after the correction, we can finally compare our estimated IGE to answer an important question: “Is the degree of intergenerational earnings persistence overestimated in Italy due to the TSTLS method used in previous studies?” Given that our corrected IGEs are between 0.440 and 0.45 and extremely close to the ones obtained by Piraino (2007) and Barbieri et al. (2018) (see Table 1) we can answer to the important question with a clear “no”. The degree of intergenerational earnings persistence in Italy is very high and comparable to those estimated for the UK and the US which are broadly considered as the countries with the lowest level of intergenerational mobility among developed countries.
7 Concluding remarks

In this work we presented comparable estimates of intergenerational earnings elasticity in Italy being able to observe actual fathers-sons pairs rather than imputed fathers’-sons’ earnings. Using the AD-SILC dataset built by merging INPS administrative archives to the 2004 to 2008 Italian samples of the EU-SILC, we are able to link sons to their actual fathers by observing that more than 90% of all sons are still coresiding with their parents when they have just left education. We then obtain averaged fathers’ earnings going back in time to take fathers with at least 4 positive observations in a 15 years period which starts when their sons were born. Instead, sons’ earnings have been observed by potential experience defined as years of distance from sons’ highest education achievement.

According to these selection rules we found an estimated IGE of 0.401 at the 6th year after sons’ highest educational achievements. This estimated coefficient is not that far from those obtained in previous works for Italy which exploit the TSTLS method and from the ones estimated for the US. Moreover, to evaluate their career-earnings profile of Italian workers, we adapted the methodology firstly proposed by Haider and Solon (2006) to a consistent sample of males which left education between 1978 and 1984 whose earnings are observed for the subsequent 30 years.

Results showed that, the estimated IGE obtained at potential experience 6 is unlikely to be consistently underestimated due to left-hand side measurement errors. This result derives from the fact that our selection of sons into sample minimizes that component of lifecycle bias related to sons’ differences in work experience. However, we re-estimate the two measures on intergenerational earnings persistence for a subsample of sons that left education between 2002 and 2006, so that they are followed at potential experience 8, and an additional one built taking sons observed 11 year after they left education. Hence, using these newly estimated coefficients and exploiting the predicted lifecycle bias obtained from the “forward regressions” of yearly earnings on lifetime earnings for each point in career, we showed that the best estimated IGE in Italy is between 0.44/0.45. According to these final results, it is possible to summarize that the degree of intergenerational earnings persistence in Italy is very high if compared to other developed countries either considering. Comparable levels of background-related earnings advantages are obtained only in the US and the UK, the less mobile developed countries. Therefore, contrary to what is often predicted in the empirical literature on intergenerational mobility (Blanden, 2013), the estimated IGEs obtained in previous
studies (Barbieri et al., 2018; Mocetti, 2007; Piraino, 2007) which exploited the TST-SLS method do not seem to be strongly upward biased even if these studies were not able to directly link earnings of sons to those of their actual fathers. Therefore, the level of intergenerational earnings mobility in Italy is likely to be very low regardless of the estimation method used. Moreover, given that it is still not possible to directly follow the two generations over their entire careers, our final IGE might still be a lower bound of the true values because of potential residual left/right-hand side measurement errors.

References


A Appendix

Figure A.1: Sons’ earnings by potential experience: selected sample v.s. full sample

Notes: Only individuals with positive earnings are considered. Source: Authors’ elaborations on the AD-SILC dataset.
Figure A.2: Empirical test of lifecycle bias: sensitivity test to the number of positive earnings observations used to proxy lifetime earnings

Notes: Workers that left education between 1978 and 1984 are followed for the 30 subsequent years. Their proxy for lifetime earnings has been calculated by averaging at least 10 positive observations in a 20 years period which begins at the 6th year after they left education and ends with the 25th (black line). Alternative measures of lifetime earnings have been calculated by averaging at least 8 (dotted line) or 12 (dashed line) positive observations in the 20 years period. In all estimates we control for the year dummies. Source: Authors’ elaborations on the AD-SILC dataset.
Table A.1: Summary statistics of sons: selected sample v.s. full sample

<table>
<thead>
<tr>
<th></th>
<th>Selected sample</th>
<th>Full sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age when leaving educ.</td>
<td>20.31 (3.47)</td>
<td>20.49 (3.57)</td>
</tr>
<tr>
<td>Weeks of work experience when leaving educ.</td>
<td>26.16 (65.36)</td>
<td>29.39 (72.57)</td>
</tr>
<tr>
<td>Years of educ.</td>
<td>13.98 (2.72)</td>
<td>14.11 (2.76)</td>
</tr>
<tr>
<td>Fund where worker pay contributions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private employee</td>
<td>76.30%</td>
<td>75.52%</td>
</tr>
<tr>
<td>Public employee</td>
<td>2.56%</td>
<td>2.39%</td>
</tr>
<tr>
<td>Cococo</td>
<td>3.32%</td>
<td>3.70%</td>
</tr>
<tr>
<td>P. IVA</td>
<td>8.95%</td>
<td>9.16%</td>
</tr>
<tr>
<td>Craft</td>
<td>1.53%</td>
<td>1.39%</td>
</tr>
<tr>
<td>Salesman</td>
<td>1.79%</td>
<td>2.23%</td>
</tr>
<tr>
<td>Agricultural worker</td>
<td>2.47%</td>
<td>2.31%</td>
</tr>
<tr>
<td>Other funds for self-employment</td>
<td>2.47%</td>
<td>2.77%</td>
</tr>
</tbody>
</table>

**Notes:** Only sons with positive earnings are considered. Standard deviations in parentheses. **Source:** Authors’ elaborations on the AD-SILC dataset.