A cross-country investigation on the patterns of intergenerational mobility[1]

By CHIARA CAVAGLIÀ

UNIVERSITY OF ESSEX

March 30, 2014

The aim of this paper is to compare the pattern of intergenerational transmissions of labour income between fathers and sons in Germany (West and East), Italy, the United Kingdom (UK) and United States (US). This research uses the Two-Sample-Two-Stage methodology to estimate the intergenerational elasticity (hereafter, the IGE), its trend across cohorts and investigate eventual non-linearities at different levels of earnings. The results suggest that the United States and Germany have the highest level of IGE, followed by Italy and the United Kingdom. For all countries, except for the US, the IGE has increased overtime. The estimates also indicate that the intergenerational elasticity is not constant along the income distribution. This is particularly apparent for the US and Italy, with a higher persistence at both extremes of the income distribution. These findings are robust to a series of controls and across sample specifications.

Keywords: intergenerational earnings mobility, two-sample-two-stage methodology, transition matrices, ordered logit

I INTRODUCTION

The majority of the literature on intergenerational transmissions is empirical[2]. The few theoretical articles were developed to explain the findings of the applied studies. The theoretical foundation of intergenerational studies is the seminal work of Becker and Tomes (1979, 1986) . Their model relies on the relationship among private investment in human capital, future earnings and the financial markets. It assumes that with perfect capital markets parents choose the optimal level of investment in education for their children. Therefore, the intergenerational relationship mainly depends on the inheritance of certain characteristics, which the authors define

[1] I thank Prof. Marco Francesconi, my supervisor, Dr. David Reinstein, Prof. Joao Santos Silva and Dr. Giovanni Mastrobuoni for their valuable comments.

[2] A complete literature review is provided in Solon (1999); Black and Devereux (2011)

as endowments. In the presence of borrowing constraints, investment of less-favoured families in their children's education may be sub-optimal and thus depend on parental income. Applied analyses suggest that higher intergenerational immobility characterises the individuals at the extremes of the income distribution. Therefore, the higher persistence at the bottom of the income distribution[3] is in line with the predictions for low income families with credit constraints.

In 2004, Solon augments Becker and Tomes' model to account for differing IGE across countries and over time. He explicitly models the public investment in human capital. In his model, the IGE is positively correlated to the heritability of human capital endowments, the productivity of human capital investments and the returns to human capital, and negatively correlated to progressive public expenditure in education. Different amounts of IGE are explained by differences in any of these parameters.

Overall, the empirical analyses confirm the predictions of the Becker-Tomes-Solon model (hereafter, BTS). For example, Blanden (2013)[4] who investigates the correlation between the IGE and income inequality, educational investment, and returns to education. The author reviews studies on three indicators of intergenerational mobility (income, occupational status and education) in sixty-five countries. For each country and each indicator he selects the articles that employ similar methods and compare the estimated IGE. Blanden (2013) also compares the IGE estimates from the selected studies to provide an international ranking of countries by IGE. His conclusions are consistent with most cross-country investigations[5]. Globally, the Nordic countries have the lowest level of persistence, whereas Southern Europe and South America are the least mobile societies. The IGE in the United States is higher than in the United Kingdom. West Germany ranks between the Scandinavian and the Anglo-Saxon countries. The empirical analyses (Corak and Heisz, 1999; Jantti et al., 2006; Blanden et al., 2005; Björklund et al., 2012) also confirm another implication of the BTS framework: the elasticity varies along the income distribution. Across a range of the different methods (quantile regression, spline functions or

[3]Bratberg et al. (2007) show that financial constraints may also explain higher persistence at the top of the distribution.

[4]This article is also interesting as it reviews and summarises the findings of a large amount of literature.

[5]Black and Devereux (2011); Björklund and Jäntti (2009); Blanden (2009) review the relevant existing literature.

transition matrices), the literature indicates higher persistence at the top and bottom quantiles of the sons' income distribution[6].

My research positions itself among this literature as it investigates the intergenerational transmission of labour income[7] in four OECD countries: Germany, Italy, the United Kingdom and the United States. Specifically, this study contributes to the intergenerational research in two ways. The first contribution is substantive. I believe that no cross-country or national study performs an exhaustive analysis of IGE at three levels: over time, at different quantiles of the income distribution, and in terms of mobility. Scholars tend to focus on a specific aspect or on a given country and refer to other research in order to complement their analysis. However, the conclusions drawn upon the comparison of different articles might be less trustworthy than custom-made cross-country analyses. The reason is that empirical results are highly dependent on the sample selection criteria and on the selected specification. For example, Grawe (2004) estimates the IGE on two datasets: the National Longitudinal Survey (NLS) and the Panel Study of Income Dynamics (PSID). The IGE for the former is significantly lower than that calculated on the latter. The author motivates this result by underlining the older age of both fathers and sons in PSID. Similarly, the estimates on NLS of this study are around half of those of Zimmerman (1992) who used the same survey but excludes fathers and sons employed for less than thirty weeks per year and thirty hours per week. My article provides comparable cross-country estimates of the patterns of IGE, and fills the gap in the Italian and German literature. For these countries, to the best of my knowledge, no recent examination on the IGE trend is available. I also believe that this is the first study to include East Germany. In addition, it estimates the IGE in the United States based on individual versus paternal earnings, whereas the majority of studies consider individual versus overall family income. The second contribution is methodological. The limitation of the existing empirical studies on mobility and persistence patterns is that the commonly used transition matrices do not allow for any control variables. I construct the transition matrices by combining the outcomes of a generalised ordered logit. The

[6]The exception are Scandinavian countries where lower mobility does not characterise the the bottom quantiles.

[7]Although the reference variable is labour income, the terms earnings and labour income are used as synonyms. Additionally, unless differently specified, the term income refers to labour income.

estimated model not only includes the income indicators but also a set of covariates (individual and paternal characteristics) that aim to control for life-cycle bias.

The rest of the paper is organized as follows. Section II describes the methodology. Section III specifies the sample selection criteria. Section IV describes the datasets and Section V outlines the results of the first stage regressions. Section VI reviews the results of the second stage estimations. Finally, section VII concludes.

II Methodology

This study employs the Two-Sample Two-Stage methodology[8]. Two-Sample Two-Stage Least Squares (TS2SLS) were first used to estimate female labour supply in the UK by Arellano and Meghir (1992), and first applied to intergenerational studies by Björklund and Jäntti (1997). Since then, an increasing number of researchers have relied on this method, especially in order to overcome data limitations. In fact, it combines two independent samples (or datasets) and solves the problem of small sample size that occurs when fathers and sons are matched within a dataset. This approach has two other advantages. Firstly, the larger permitted sample size allows more flexibility to set ex-ante consistent and stringent criteria for sample and variable selection across datasets. Secondly, it could overcome a sample selection bias as highlighted by Francesconi and Nicoletti (2006). In fact, all the relevant surveys[9] require sons and fathers to live in the same household, or to have lived at the time of the first wave, in order to report the earnings of both.

In brief, the first stage consists of dividing the respondents of each survey into two groups, according to a set of criteria (listed in section III)[10]. The first, the auxiliary subsample, constitutes the fictitious (or synthetic) fathers. The characteristics of these individuals are used to predict the paternal earnings of the individuals in the main sample, the sons, through the following

[8] The method is preferred to the Two-Sample Instrumental Variable (TSIV) approach because, ceteris paribus, the former is asymptotically more efficient than the latter, unlike the single-sample case where the two estimators are the same (Inoue and Solon, 2010).

[9] As indicated in section III, the selected surveys are SHIW for Italy, BHPS-USS for the United Kingdom, PSID for the United States and the German Socioeconomic Panel

[10] This technique is extensively explained in Arellano and Meghir (1992); Angrist and Krueger (1992). Examples of its application in the field of intergenerational mobility are Mocetti (2007); Piraino (2007) for Italy, Nicoletti and Ermisch (2007) for the United Kingdom and Grawe (2004) for the United States.

equation:

$$(1) \quad y_f = \lambda Z + \nu$$

where y_f is the vector containing the log of the pseudo-father's labour income, Z is a matrix of regressors about personal characteristics such as age and occupation, λ is the vector of coefficients and ν is a vector of white noise disturbances. The estimated coefficients, $\hat{\lambda}$, are then used in the second stage to predict paternal income in the main sample, $\hat{y}_f = \hat{\lambda} Z$, where Z is a matrix containing the same regressors of eq.1. Specifically, the covariates in Z in the main sample refer to the paternal characteristics when the respondent was a given age, usually fourteen or fifteen [11].

In the second stage, $\hat{y}_f$ is included in the empirical model, best fitted for this research[12]. The baseline framework is TS2SLS:

$$(2) \quad y_s = \beta \hat{y}_f + \delta X + \varepsilon$$

where X is a matrix of control variables for individual and paternal characteristics. In this model, β is the intergenerational elasticity (IGE), and indicates the fraction of income that is on average transmitted across generations. In general, β ranges between zero (complete mobility) and one (complete immobility).

Although very common in intergenerational studies[13], this technique is more problematic than Ordinary Least Square estimations (OLS). In particular, the limited amount of information about the parents provided by the respondents as well as age differences between the synthetic and the real fathers complicate the study. The above shortcomings affect consistency, for which Z need to be identically and independently distributed (i.i.d.) in the two samples and independent of the error term, ε[14]. If eq.2 is correctly specified the estimator β_{TS2SLS} will converge

[11] This applies to the data of Germany, the United Kingdom and the United States, whereas for the Italian case the questions refer to the period when the father is the same age as the individual.

[12] The aim of this research is to uncover and compare non-linearities in the transmission of earnings. Thus, the baseline model will be modified (for example, to study the trend of IGE across cohorts) or substituted (by quantile and logit regressions) according to the different research questions. Additional details on each model are provided in the relevant subsection of section III.

[13] And so far, the only available in countries such as Italy, where the panel of the only suitable survey, SHIW, is still too short for an OLS estimation.

[14] As discussed in Arellano and Meghir (1992)

in probability to β, ($p\lim \beta_{TS2SLS} = \beta$). However, if ε depends on Z and the regressors have not only an indirect effect through the endogenous regressors, but also a direct effect on the dependent variable, then the probability limit will be different from β and the coefficient is not consistent. Nicoletti and Ermisch (2007) generalize the result of Solon (1992) on the instrumental variable estimator:

$$(3) \quad y_s = \theta y_f + \kappa Z + \lambda X + \xi$$

where y_f is the measured paternal income in β_{OLS} and it is equal to $\hat{y}_f$ for the TS2SLS estimator. If for simplicity we ignore λX, then

$$(4) \quad \begin{aligned} p\lim \beta_{OLS} &= \frac{Cov(y_s, y_f)}{Var(y_f)} = \theta + \frac{Cov(y_f, Z)}{Var(y_f)}\kappa \\ p\lim \beta_{TS2SLS} &= \frac{Cov(y_s, P_z y_f)}{Var(P_z y_f)} = \theta + \frac{Cov(P_z y_f, Z)}{Var(P_z y_f)}\kappa \end{aligned}$$

where P_z is the projection matrix created by the variables z. Given that $Cov(y_s, y_f) = Cov(y_s, P_z y_f)$ then the two estimators converge to the same probability if $Var(y_f) = Var(P_z y_f)$, or when $\frac{Var(P_z y_f)}{Var(y_f)} = 1$, which indicates the R^2 of the regression in eq.1. From eq.4 it derives that if y_f and Z are positively correlated and the vector of coefficients κ is also positive, asymptotically β_{TS2SLS} is at least equal to β_{OLS}. Nicoletti and Ermisch draw two conclusions. Firstly, in terms of consistency it is important to select the components of Z so that the covariance between them and the error term of eq.2 is as close to 0 as possible. Secondly, and if, ε and Z are not independent a way to reduce the difference between β_{OLS} and β_{TS2SLS} is to select instruments that are highly correlated with y_f in the first stage.

Considering the above differences between β_{OLS} and β_{TS2SLS}, and in order to ensure the comparability of the results across different surveys, the research applies TS2SLS to all four countries, including Germany and the United States, where the OLS could be used. TS2SLS are necessary with the selected Italian and British datasets.

III Empirical strategy and sample selection

It is important to be aware of the consequences of sample selection, even more so when comparing findings across time or countries. The intergenerational literature reviewed in Solon (1999);

Black and Devereux (2011) shows that the characteristics of the dependent and independent variables, as well as the sample features affect the results. For example, Solon (1992), Zimmerman (1992) and Mazumder (2005a) demonstrate that the use of point estimates of income leads to biased results. To reduce the impact of transitory shocks on annual earnings they propose to substitute yearly earnings with five-year averages. Another important methodological contribution is Haider and Solon (2006). They confirm the patterns of the so-called attenuation bias and, more importantly, they draw the attention to the fact that a bias does not only arise from measurement errors in paternal earnings. So far, the literature had not devoted great importance to measurement error in the dependent variable as it assumed the validitiy of the textbook errors-in-variables model[15]. However, the articles that Haider and Solon review and the results of their empirical analysis lead them to conclude that the textbook errors-in-variables model is not suitable[16]. Although the authors suggest that the textbook framework can be applied if the income is measured between the early thirties and mid-forties, they also notice that the majority of studies report the earnings of the sons at earlier ages. As the attenuation factor changes with age, and the outcome of sons is usually observed at an earlier age than the age at which paternal earnings are considered, β might suffer from life-cycle bias[17], of which the direction depends on the age of the son[18].

Given the crucial role played by the sample selection and the choice of the variables on the estimates, I develop identical guidelines for the four countries[19]. A set of three criteria is common to both the main and auxiliary samples.

The first criterion concerns the unit of analysis. In order to simplify the interpretation of the results and the cross-country comparison, women are excluded from the sample. For example,

[15] According to the textbook model, a measurement error in the dependent variable would only affect the standard errors and not the consistency of the estimator.

[16] In particular,the authors show that the deviation is not random.

[17] A life-cycle bias appears when the slope of a regression of log earnings at a given age on the log of the present value of lifetime income is not the same for fathers and sons.

[18] Other applied literature also confirms the existence of biases. For example, Grawe (2006) who reviews twenty studies on the United States and Canada and Nilsen et al. (2008, in Black and Devereux 2011) on Nordic countries.

[19] The summary statistics for each dataset are indicated in Appendix A.

the career choices of women, and consequently their labour income, might be influenced by a higher number of factors[20].

The second criterion relates to the variable on which the elasticity is calculated: positive annual[21] earnings. Studies using the parent's overall income and parent's earnings suggest that the former yields higher values of intergenerational elasticity (Grawe, 2004). In that case, non-linearities might be explained by the inheritance of non-labour income. Therefore, considering total income might obfuscate the role of labour market and education. Moreover, considering family earnings rather than one parent's income might require additional assumptions. For example, Finnie and Irvine (2006) analyse the composition of couples at the top quantiles of the distribution, using tax information in Canada and they find that whereas high-earning women are usually coupled with high-earning men, the reverse is not true.

The third principle is the exclusion of individuals with missing information about their job position. Considering that the research focuses on intergenerational transmissions of earnings of active individuals, it aims to reduce the number of men who are not working.

Following Nicoletti and Ermisch (2007), the year of birth is used to allocate the respondents to the two samples. For each dataset, the males born after 1949 belong to the main sample. Additionally the information provided about their real fathers must refer to the period when the latter are between thirty and fifty-nine of age. The real fathers also need to be born between 1900 and 1949. The males born before 1950 compose the subsample of the fictitious fathers. A requirement for the individuals to be classified in this sample is that they have children. For the case of Italy, considering only individuals with non-missing information about their own children would reduce drastically the number of observations. Thus, the presence of children in the household is considered instead.

The main and auxiliary samples consist of males between thirty and fifty-nine years of age (hereafter, *ALL*), with repeated information over the years. The analysis is based on *ALL*. In

[20] Additional details on the difficulties for the intergenerational transmission of earnings between fathers and daughters are mentioned in Black and Devereux (2011).

[21] The only exception is the last wave of UK data, that is the second wave of Understanding Society survey, where I derived this variable by multiplying the monthly income by twelve. Furthermore, observations for which annual labour income is imputed are excluded from the research.

order to check the robustness of the results, however, three alternative subsamples are also considered [22]. The first includes respondents with at least three-year averages of own earnings ($AV3$)[23]. The second retains one observation per individual: the values of the variables referring to the year when the age is closer to forty ($MIN40$). The third sample adds age restrictions to *ALL* and considers only the respondents between thirty-five and fifty-five years of age ($A55$)[24].

Two additional robustness checks are reported in Appendix I. In this appendix, the first exercise excludes the outlying observations, whereas in the second case self-employment is explicitly accounted for.

A final task is performed in Appendix H. I compare and contrast the change in the IGE estimates in *ALL* when four dummy variables for educational attainment are added (from no secondary education to university degree). Despite the limitations of this analysis, such as the endogeneity of education, this exercise can provide interesting insight.

IV Datasets

For Germany, the selected survey is the German Socioeconomic Panel (GSOEP), a longitudinal household based study which started in 1984 in the Federal Republic of Germany, and extended to eastern German households from 1990. Whereas all the studies I am aware of consider West Germany only, this research examines the whole of Germany. The time ranges from 1991 to 2010 (2008 for the auxiliary sample).

The Italian Survey on Household Income and Wealth (SHIW) started in the 1960s. Over the years it has been extended to include several aspects of the economic and financial behaviour of around 8,000 households (24,000 individuals). Given its construction, it is not straightforward to follow individuals over time, as there is no unique individual identifier. Nonetheless, as long as they belong to the same household, individuals can be linked across waves by exploiting information about the identifier of the previous wave. Since 1987 the survey has been carried out

[22]The relevant Appendices report the estimations for these three additional samples.

[23]As Appendix A highlights, however, the average number of years is greater than five in all the datasets.

[24]Originally, $A55$ was set to be the main sample. However, the estimation results suggest that given the empirical specification, there is no substantial difference between $A55$ and *ALL*. The latter, however, has the advantage of a larger sample size. Moreover, an analysis on the intergenerational trend can capture also later cohorts (up to 1979 for *ALL* versus 1975 for $A55$).

every other year, which reduces the number of observations to a maximum of ten per individual, from 1991 to 2010 (2008 for the auxiliary sample).

For the United Kingdom, the research combines British Household Panel Survey (BHPS) and Understanding Society Survey (USS) to obtain observations about the original BHPS sample from 1991 to 2010 (2008 for the auxiliary sample). BHPS began in 1991 with a representative sample of about 5,500 units and 10,300 individuals, to which households of Scotland, Wales and Northern Ireland were added subsequently. In 2009, the survey converged into Understanding Society Survey (USS), where since the second wave of this survey, the original BHPS sample has been inserted. The reference unit is the household, although all individuals of the original sample are followed throughout the years, even after joining a new household[25].

For the United States, the selected dataset is the Panel Study of Income Dynamics (PSID). The study began in 1968 and interviewed a sample of over 18,000 individuals from 5,000 families. It was done annually until 1997 and then biannually. The survey collects details on individual's income, occupational position, education and other topics. Each individual is asked details about his or her parents, including occupation of the father when the respondent was fifteen of age[26]. For consistency with the other datasets, the considered years are from 1991 (1990 labour income) to 2009 (2008 earnings).

V FIRST STAGE REGRESSIONS

The selection of instruments is a crucial stage for consistency. Despite disagreement over the relative contribution of each of the factors affecting earnings (Mincer, 1974; Spence and Stiglitz, 1975; Heckman, 1985), a consensus seems to exist on the importance of variables such as age, education, experience in the labour market and gender. The main obstacle to predict paternal earnings are data constraints. Thus, the instruments used vary from one study to another according to data availability. For example, Grawe (2004) uses only educational levels; Aaronson

[25] Whereas only two studies have used BHPS to assess the extent of IGE in the United Kingdom ((Ermisch and Francesconi, 2004; Nicoletti and Ermisch, 2007), this survey is more appropriate than the yet more popular National Child Development Study (NCDS)and British Cohort Study (BCS). The main reason lies in their construction that does not allow to examine the variation of IGE overtime. In fact, the only follow a given cohort: 1958 for the NCDS and 1971 for BCS.

[26] Other surveys have been used in intergenerational studies such as the National Longitudinal Surveys (NLS). These surveys, however, focus on specific cohorts. Consequently, they are less comparable to other datasets and face the same limitations mentioned for the British NCDS.

and Mazumder (2008) use the year and the state of birth. Björklund and Jäntti (1997) who perform the study on Sweden include eight occupational categories and two dummy variables: one for the fathers living in Stockholm and one if education is higher than the compulsory level. Mocetti (2007) and Piraino (2007) include geographical area, occupation, education and work status to study intergenerational transmission of earnings in Italy. Nicoletti and Ermisch (2007) add age and age squared, and the Hope-Goldthorpe score as well as occupation and education.

The current study faces an additional constraint with respect to the literature mentioned above. Indeed, whereas the above research only focuses on one country (except for Grawe 2004), the need for comparabilitymeans that only those instruments that are available in all four datasets can be considered[27]. For this reason, the Hope-Goldthorpe score is not used as a regressor for the first stage model of the United Kingdom; the same applies to the Treiman score for Germany.

Overall, the selected instruments provide information about occupation, education, geographical origin, date of birth and age, as indicated in eq. 5:

$$\begin{aligned} y_i^f &= \lambda_0 + \sum_g \lambda_g educ_{ig} + \sum_j \lambda_j job_{ij} + \sum_k \lambda_k race_{ik} \\ &\quad + \sum_l \lambda_l location_{il} + \sum_h dadanasc_{ih} + \sum_m \lambda_m dadcat_{im} + \nu_i, \\ where\ i &= 1...n \end{aligned} \tag{5}$$

where *educ* is a set of dummies for completed education; *job* is a series of indicators providing information about the type of occupation, the sector and if the individual is employed or self-employed; *race* indicates ethnicity[28]; *location* is a set of indicator variables providing information about the geographical position of the individual; *dadanasc* indicates the year of birth. Finally, $dadcat_m$ includes m categorical variables about age. Specifically, the variable includes four age categories: up to 43 years of age; from 44 to 48; from 49 to 53; and from 54 to 59[29]. $dadcat_m$ substitutes the age and age squared used in Nicoletti and Ermisch (2008). The

[27] The variables are not exactly the same for the four surveys. Obviously, this is also related to the fact that institutions, such as educational systems, differ from country to country.

[28] For Italy it is not possible to obtain information about racial origin. However, according to the latest report of the national institute for statistics (ISTAT), the percentage of non-Caucasian immigrants accounts for about 1.5 percent of the population, and 20 percent of these are children.

[29] The excluded category is the third for both analyses in the first and second stage.

coefficients of this equation are used to predict the paternal income in the second stage. The prediction is based on information from the son about the father when the son was fourteen or fifteen of age[30]. However, it may be extremely difficult for him to recall the exact occupation of his father in a specific year. Moreover, the quadratic term could further amplified the error. The recollection error may be attenuated if a broader age category is used instead of the precise age. The combination of the paternal year of birth and age categories may take into account cultural and socioeconomic changes, such as the changing status of a given occupation, or the evolving returns to education. It is relevant as the considered time span is twenty years and labour income is registered in different waves for different individuals.

For each dataset, the same characteristic is measured by several variables. Similarly, the number of categories of a given variable can be altered. For this reason, alternative combinations of instruments are used to estimate eq.5. At least four sets of first stage equations are estimated, except for Italy where the information about parental characteristics is limited. The details about the model selection and the regression outputs for the four countries are provided in Appendix B. Globally, the criteria to idenfity the preferred model are the higher value of the Adjusted R^2 and the comparability of the regressors across surveys. Moreover, the Ramsey Regression Equation Specification Error Test (RESET) and the link test are taken into account[31]. Ceteris paribus, the model with the higher number of instruments is preferred, as it ensures higher variability in the prediction in the second stage. Finally, in line with Nicoletti and Ermisch, the endogeneity of the instruments is tested by assessing the correlation of the residual from the intergenerational equation (eq.2) with the auxiliary variables (of eq.5). Non-statistically significant covariates provide evidence in favour of exogeneity. For statistically significant coefficients, a lower R^2 is preferred as it indicates a lower correlation. Appendix C reports the regressions and discusses the outputs.

[30] Only in SHIW the information refers to the period when the father had the same age as the current age of the respondent.

[31] Both tests assess the general specification of the model. They achieve this goal by estimating a model with the linear predicted value of a regression and its squared. The reasoning behind the test is that if the model is correctly specified (the null hypothesis) any additional regressor (in this case the squared fitted value) should have no predictive power.

VI Second stage regressions: model specifications and results

As mentioned in section III, the choice of the sample and of the variable to proxy lifetime earnings might produce biased estimates. In order to reduce bias, some authors favour average earnings (Solon 1992), others explore the possibility to use educational attainment as an instrument for life-time earnings. Both methods, however, have some limitations. The drawback of the Instrumental Variables (IV) approach is the difficulty in finding a reliable instrument, whereas the use of averages depends on data availability.

An alternative way to reduce the bias is suggested by Hertz (2007) and Lee and Solon (2009). The authors augment the basic intergenerational regression with an interaction between the individual's age and the paternal income, and apply this model to study the pattern of IGE over time. The current research uses (and adapts to TS2SLS) the specification proposed by Lee and Solon (2009) to obtain the following baseline model:

$$
\begin{aligned}
y_{st} = {} & \delta_0 + \beta \hat{y}_f + \delta_1 (age_s - 40) + \delta_2 (age_s - 40)^2 + \delta_3 (age_s - 40)^3 \\
& + \delta_4 (age_s - 40)\, \hat{y}_f + \delta_5 \,(age_s - 40)^2\, \hat{y}_f + \delta_6 \,(age_s - 40)^3\, \hat{y}_f \qquad (6) \\
& + \delta_7 birthy_s + \delta_8 birthy_f + \sum_{j=1}^{4} \mu_j dadcat_{fj} + \varepsilon_t
\end{aligned}
$$

where $(age_s - 40)$ is the son's age normalised at 40, $(age_s - 40)^2$ and $(age_s - 40)^3$ are its square and cube respectively; $(age_s - 40)\, \hat{y}_f$ $((age_s - 40)^2 \hat{y}_f$, $(age_s - 40)^3 \hat{y}_f)$ is the interaction between normalised age (squared, and cubed) and paternal income[32]. The statistical significance of the coefficients and the results of the RESET test suggest that the introduction of a polynomial of order three better suits the data than a polynomial of order one and two [33].

[32] Following Lee and Solon (2009) I included $(age_s - 40)^4$. However, for the four datasets it was not significantly different from zero at standard levels. So, it was dropped.

[33] In order to select the suitable order of the polynomial a Wald test on the joint significance of the coefficients, the RESET test and the t-test for the statistical significance of each coefficient were considered. With GSOEP, a polynomial of order three represents the best compromise between a quadratic term, where all the coefficients are significant and with a lower adjusted R^2 and a model with a quartic in age and in its interaction, where neither the two regressors with the highest power nor the cube terms are significant. The same exercise on Italian data provides different results: data suggest that a model with only the quadratic term is worse than a linear model or than a model with the cubic term. The two latter perform better in terms of adjusted R^2 and the individual significance of the coefficients. Between these two, the model specified in eq.6 is selected for the following reasons: firstly, it has a slightly higher adjusted R^2; secondly, although δ_2,δ_3,$\delta_5$and δ_6 are not statistically significant individually, their joint significance cannot be rejected; finally, this specification is consistent with the model selected for the other datasets. Anyway, there is no significant difference among IGE calculated with these two models: 0.450 with a linear interaction versus 0.459 for the model specified

Differently from Lee and Solon (2009), but consistently with Hertz (2007) along with Nicoletti and Ermisch (2007), the model includes the year of birth of the son ($birthy_s$) rather than the year labour income is observed, which explicitly accounts for the fact that individuals are not observed at the same age[34].

Whereas Lee and Solon (2009) include paternal age and its square, the current article employs dummies for age, for the reasons specified with reference to eq.5. I also introduce the father's year of birth ($birthy_f$). This variable is especially relevant with TS2SLS, as it is a way to control for cultural changes across the different generations of fathers, such as modifications in the occupational prestige. In fact, the impact on income of the characteristics used to predict paternal earnings could change over time (e.g. different returns to education). Considering that on average the real fathers are born earlier than the fictitious ones, the predictive power of the coefficients of the first stage regressions might be affected if this divergence was not accounted for.

The covariates specified in eq.6 remain the same for all the equations that will be estimated in this research.

Table 1 reports the baseline results[35]. For each dataset, two models are estimated. Model (2) is specified in eq.6, whereas (1) is eq.6 without the interaction terms between paternal income and the age of the respondent. Table 1 shows that the impact of the interaction term differs: in Italy it slightly decreases the value of β (from 0.48 to 0.46), whereas the IGE increases in the United Kingdom, United States and in Germany. Overall, Table 1 suggests that the IGE is higher in the United States, followed by Germany, Italy and United Kingdom. This

in eq.6. With BHPS-USS, the quadratic and of the cubic terms have no individual statistical significance. Nonetheless, they are jointly statistically different from zero. Additionally, the adjusted R^2 is higher than with the linear interaction only. None of the specifications performs significantly better than the others with the RESET test. Moreover, the results of the linear, quadratic and cubic models are similar (0.337, 0.342, 0.343, respectively). So, for consistency with the other countries the cubic term is introduced. With PSID the cubic improves the statistical significance of the model. Whereas with the quadratic term the relevant coefficients are not statistically significant, introducing the cubic term adds statistical significance to two of them (δ_2, δ_5). Furthermore, it is not possible to reject the null hypothesis that model is correctly specified at 1 percent significance level.

[34] If individuals were observed at the same age, selecting the year at which the outcome is observed or the cohort would be the same.

[35] In order to check whether the magnitude of the coefficient is affected by estimating the first and the second stage separately, Table 29 in Appendix D reports the estimated coefficients obtained by simultaneously bootstrapping the first and the second stage. The results do not change in a relevant way from those of model (2) in Table 1.

TABLE 1
TS2SLS, BASELINE MODEL WITHOUT AND WITH THE INTERACTION WITH PATERNAL INCOME

	(1) Germany	(2) Germany	(3) Italy	(4) Italy	(5) UK	(6) UK	(7) US	(8) US
Father's ln(LI)	0.521***	0.606***	0.483***	0.459***	0.337***	0.343***	0.569***	0.614***
	[0.012]	[0.016]	[0.022]	[0.029]	[0.018]	[0.022]	[0.022]	[0.030]
Age-40	0.099***	-0.000	0.075***	-0.061**	0.046***	0.019	0.025***	0.021
	[0.002]	[0.024]	[0.004]	[0.031]	[0.002]	[0.034]	[0.003]	[0.073]
$(\text{Age-40})^2$	-0.001***	0.024***	0.000	0.002	-0.001***	0.001	-0.001**	0.012**
	[0.000]	[0.003]	[0.000]	[0.007]	[0.000]	[0.003]	[0.000]	[0.006]
$(\text{Age-40})^3$	-0.000	-0.001***	-0.000	0.000	0.000	-0.000	-0.000*	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]
LI(Age-40)		0.011***		0.016***		0.003		0.000
		[0.003]		[0.004]		[0.004]		[0.007]
$\text{LI}(\text{Age-40})^2$		-0.003***		-0.000		-0.000		-0.001**
		[0.000]		[0.001]		[0.000]		[0.001]
$\text{LI}(\text{Age-40})^3$		0.000***		-0.000		0.000		0.000
		[0.000]		[0.000]		[0.000]		[0.000]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27442	27442	6860	6860	14363	14363	7530	7530
Adjusted R^2	0.270	0.275	0.366	0.368	0.125	0.125	0.102	0.102

Bootstrapped errors in brackets
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

is partially consistent with a common pattern that seems to emerge from the cross-country literature: Nordic countries, Canada and Germany are characterised by lower IGE than the Anglo-Saxon countries. Between the two, the United Kingdom is more mobile than the United States ((Blanden et al., 2005; Bratsberg et al., 2007; Jantti et al., 2006; Grawe, 2004)). Despite the limitations of comparing different studies, these results are also consistent with the majority of research on a single country. Italy is not usually included in cross-country literature, perhaps because it is not possible to use OLS to estimate the elasticity with SHIW. The IGE estimates of Mocetti (2007) and Piraino (2007) are in line with the ranking suggested in this research. In particular, β is 0.459, which lies within 0.499 and 0.435, the elasticities in Mocetti (2007) and Piraino (2007) who employ the same technique and dataset.

In the United Kingdom, the IGE differs widely across studies. The majority uses either the NCDS (Dearden et al., 1997), or the BCS, or both (Blanden et al., 2004). It is not straightforward to compare their results with Table 1. Both surveys are different from BHPS by construction, since they follow one cohort of individuals over time. Ermisch and Francesconi (2004) are the first researchers who use BHPS to estimate the IGE, 0.247, although they consider an occupational prestige score and not earnings[36]. A more comparable study is Nicoletti and Ermisch (2007). They estimate the IGE on a restricted sample of sons born between 1962 and 1972, aged between 31 and 45, who have co-resided with their father for at least one wave of the BHPS. Their estimate, 0.365, is similar to β in Table 1 (0.34).

Whereas TS2SLS is necessary to investigate IGE with SHIW and BHPS-USS, PSID and GSOEP are larger datasets. This allows individuals to be matched with their real fathers. Consequently, it is difficult to find a recent study that employs TS2SLS. An example for the US are Björklund and Jäntti (1997). The authors consider five-year averages of parental earnings and control for the age of fathers and sons. On PSID they obtain a β of 0.516. The value of IGE when earnings are measured as averages ($AV3$) is higher in this research: 0.63 (0.61 without the interaction term). The differing results might be attributed to the sampling technique: the authors use income averages for the fathers, whereas for the sons they consider the earnings in 1987, for individuals between 28 and 36 years of age. Moreover, Björklund and Jäntti (1997) like other cross-country articles using TS2SLS for the United States, such as Grawe (2004), use a reduced number of instruments in the first stage regressions, which may reduce the precision of the predicted income. A more relevant exercise is Lee and Solon (2009). To compare the results of Table 1 with Lee and Solon (2009), I convert my estimates following the calculation of Solon (1992), cited in Blanden (2009). According to him, the IGE estimates computed through IV for PSID are larger than those based on OLS by a factor of 0.75. If the β_{PSID} in Table 1 was scaled by that factor, the result, 0.46, would be very similar to the average value of 0.44 in Lee and Solon (2009) (with a sample of individuals from 25 to 48).

The estimated β for Germany is higher than the findings of the literature so far. Eisenhauer and Pfeiffer (2008) assess intergenerational mobility for full-time employed West Germans from 30 to 50 years old. Their estimates range from 0.205 (OLS, single-year earnings) to 0.282 (OLS, five-years averages) to 0.374 (IV where they instrument earnings with years of education). In the same article, they review previous research on Germany (Couch and Dunn

[36]Precisely, the Hope-Goldthorpe occupational prestige score, based on a survey in England and Wales where respondents had to provide information about social desirability of male occupations (Goldthorpe and Hope, 1974).

1997, Wiegand 1997 and Vogel 2007): the reported β go from 0.124 to 0.266 for OLS, to 0.402 for IV (Wiegand). All the above estimates are significantly lower than 0.607 (or 0.522 without the interaction) in Table 1. The exclusion of Est Germany is likely to explain the difference between my article and the above mentioned literature. In fact, despite diverse sample sizes, selected variables, methodology and functional form, the studies of intergenerational mobility in Germany all share a common characteristic: they only consider West Germany. Table 30 in Appendix D reports the estimates for West Germany: β decreases to 0.460 with the interaction terms (or 0.320 without). These results are much closer to the above mentioned IV estimates of β. This suggests differing IGE patterns between East and West Germany: particularly, the mobility decreases in a considerable way when East Germany is included. The literature may have not provided a complete picture of the IGE in a united Germany so far. Nonetheless, further investigation is required to rule out the hypothesis that these results are driven by sample characteristics ignored at this stage [37].

Finally, Appendix D reports the IGE on $AV3$, $MIN40$ and $A55$. The interaction between paternal income and age of the respondent introduced by Lee and Solon aims to reduce the life-cycle bias. Consequently, the estimations are expected to be more robust across samples. Indeed, the introduction of the interaction renders the estimates of β more similar. This is particularly evident with PSID.

To sum up, this section establishes the consistency of my empirical specification with other IGE studies. In order to obtain further clarifications on cross-country differences, however, it is useful to disaggregate the average impact of paternal earnings.

(a) IGE Trend across cohorts

The investigation of the pattern of IGE over time might provide additional insight on the differences between IGE across countries. The goal of this exercise is to contribute to clarify the

[37] Although beyond the scope of the current research, it might be interesting to further explore regional differences in IGE. The variation between East and West Germany has motivated a similar exercise for Italy, a country with large regional disparities. Table 33 in Appendix D reports the elasticities for the North, the Centre and the South. In terms of IGE, however, regional differences in Italy are not as strong as in Germany.

contrasting results in the literature about the IGE trend in the United Kingdom[38], and to fill the gap in the literature on Italy and Germany. Such an analysis might also provide preliminary information on the effectiveness of public policies for inequality reduction. Björklund and Jäntti (2009) combine the Gini coefficient of annual disposable income for the Luxemburg Income Study with β from different studies on IGE. Although they cannot establish the direction of causality, they confirm that high intergenerational elasticities are associated with high levels of inequality in the population. Therefore, higher mobility over time should be positively correlated with lower inequality.

Lee and Solon (2009) and Hertz (2007) analyse the evolution of IGE in the United States. They regress family income on individual income of the respondents born between 1952 and 1975[39]. The current research follows closely both articles and estimates the following model:

$$(7) \qquad y_{st} = \sum \beta_c D_c \hat{y}_f + \delta X + \varepsilon_t$$

where $D_c \hat{y}_f$ is an interaction term between paternal income and cohort indicators. Differently from Solon and Lee (2009), however, my model displays a constant intercept, whereas the slope is allowed to change with cohorts. This facilitates the interpretation of the coefficients and their graphical representation. In addition, a matrix of covariates, X, controls for the year effects[40]. I add cohort interactions to eq.6:

$$\begin{aligned}(8) \qquad y_{st} \;=\; & \gamma_0 + \sum_{c=1950}^{1980} \beta_c D_c \hat{y}_f + \gamma_1 (age_s - 40) + \gamma_2 (age_s - 40)^2 \\ & + \gamma_3 (age_s - 40)^3 + \gamma_4 (age_s - 40)\, \hat{y}_f + \gamma_5 (age_s - 40)^2\, \hat{y}_f \\ & + \gamma_6 (age_s - 40)^3\, \hat{y}_f + \gamma_7 birthy_s + \gamma_8 birthy_f + \sum_{j=1}^{4} \lambda_j dadcat_{jf} + \varepsilon_t\end{aligned}$$

[38]These studies are reviewed in Nicoletti and Ermisch (2007). To my knowledge, this article also constitutes the most recent attempt to estimate changes in IGE across cohorts in Britain.

[39]Hertz augments his model to account for heterogeneity. Particularly, he introduces dummies for each individual, and additional interactions of paternal income with, for example, dummies for the presence of Afro-American children, and the mean education of the head of household and his/her spouse Furthermore, the sensitivity tests performed by Hertz highlight the positive bias that arises when only one-year observation per individual is considered. This is consistent with our results for PSID and SHIW: the estimates of MIN40 are higher than those of the other samples (Appendix D).

[40]The IGE trend across the years is calculated in Appendix E.

Figure 1[41] represents graphically the coefficients β_c for *ALL*[42].

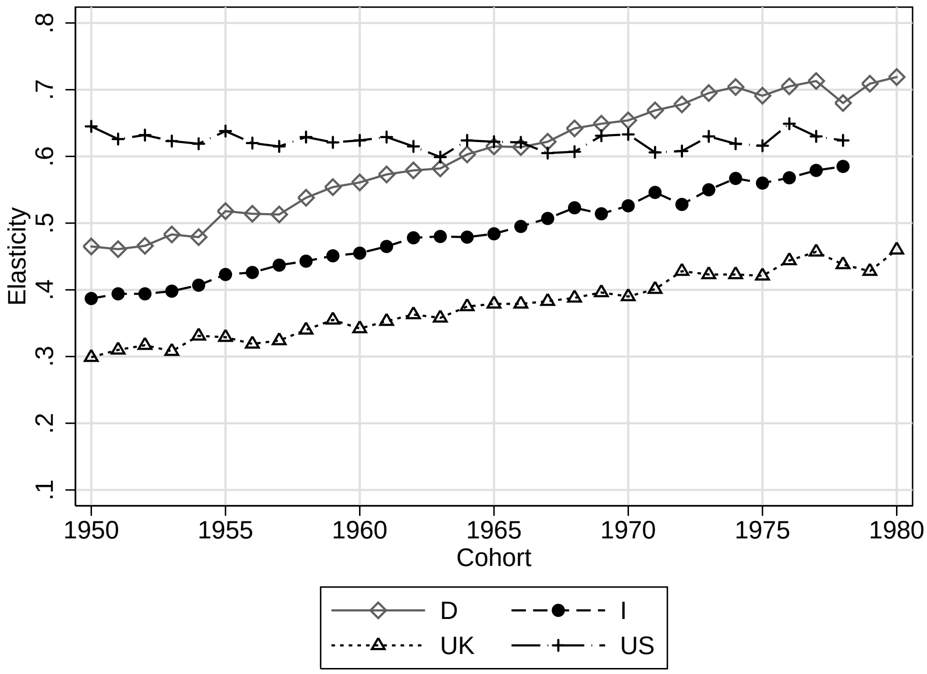

FIGURE 1
INTERGENERATIONAL ELASTICITY BY COHORT

My findings are robust across all sample specifications (in Appendix E) and in line with the literature. For the United States the conclusions are consistent with Lee and Solon (2009) and Hertz (2007): the former assess the trend over time, whereas the latter investigates the IGE across cohorts. Both articles provide evidence against the existence of a trend in IGE over time. Figure 1 indicates that β_c are very similar to each other and the Wald tests suggest that the differences among the coefficients are not significantly different from zero at all standard levels[43]. Further evidence in favour of a constant elasticity across cohorts in the US is provided in Appendix E: for the subsamples *AV*3 and *MIN*40,there are no statistically significant differences

[41] The detailed results and standard errors are reported in the Appendix E, (Table 40)

[42] The main advantage of this sample is especially clear with this exercise, where it is possible to examine a greater number of cohorts: specifically, 1950 to 1978 for the United States, 1950 to 1980 for the United Kingdom and Germany from 1950 to 1978 for Italy. As per Italy, the sample consists of only 1 individual born in 1979; so, this observation is included in 1978 cohort.

[43] There is little evidence about some differences when distant cohorts are considered (for example, 1950 and 1965, or 1950 and 1964, 1950 and 1975) at five or ten percent significance level. Differences among the statistically significant β_c would suggest a slight decrease in IGE between the cohort of 1950 and the respondents born after 1960. However, the gap is so narrow (0.645 in 1950 versus 0.616 in 1975) that some caution is required before coming to conclusions about a possible trend.

among β_c. All the above is only partially in contrast with Aaronson and Mazumder (2008) who apply TS2SLS on US Census data. Their findings indicate an increased mobility from 1950 to 1980, but a decrease after 1980. The authors consider earlier cohorts (individuals born between 1921 and 1975, from 25 to 54 years of age) and they only use the state of birth and the cohort in the first stage to predict parental income, which may partly explain the different conclusions. An alternative explanation might posit the existence of a positive trend for mobility that requires a huge number of observations to be detected as the extent is very limited. PSID (and especially when real fathers are matched with real sons) might be unsuitable for this purpose.

In the other three countries Figure 1 and Table 40 indicate a reduction of mobility over time. In United Kingdom, the IGE ranges from 0.299 (1950 cohort) to 0.421 (1975 cohort) to 0.460 (1980 cohort). The coefficients are statistically different from each other at 1 percent level of significance, except for later cohorts where the strength of the significance decreases[44]. This is consistent with Blanden et al. (2004) and Blanden et al. (2007) who estimate the impact of family income on individuals born in 1958 (NCDS) and 1970 (BCS) and observe an increase in the IGE for the later cohort [45]. The literature based on BHPS provides contrasting results. Figure 1 partially supports Nicoletti and Ermisch (2009) who conclude in favour of a statistically significant positive IGE trend for the individuals born after 1960, whereas a decrease in IGE emerges from Ermisch and Francesconi (2004). However, the exercise of Ermisch and Francesconi (2004) is not directly comparable. The sons in their sample are younger and they use the Hope-Goldthorpe score rather than labour income [46]. Finally, both articles assess the trend by cohort groups. Given that the cohort groups differ, the results may be sensitive to the categorisation. In Italy, β increases from 0.387 for the 1950 cohort to 0.585 for the individuals

[44] For example, β_{1975} is statistically different from β_{1978} or from β_{1980} at 10 percent level. However, we can reject the null that β_{1978} and β_{1980} are the same at 5 percent confidence level.

[45] According Blanden et al.(2004), this difference might be explained by the lower percentage of working mother for the older cohort.

[46] Goldthorpe and Jackson (2007) suggest that either the measurement error in income is greater, or social class and income do not capture the same socioeconomic dimension. The authors, indeed, estimate and compare the results of changes in IGE of income and class mobility across cohorts, using NCDS and BCS, for men and women. They uncover a negative trend for income mobility. However, little difference emerges on class mobility for the two cohorts.

It is worth mentioning, however, that the study of Ermish and Francesconi is not necessarily comparable to Goldthorpe (2007): the latter calculates IGE on a occupational class that is a categorical variable with seven outcomes, based on the individual's Socioeconomic Group; the former instead use the Hope-Goldthorpe score, where values range from 17.52 to 82.25.

born in 1978. The β_c are different from each other at 1 percent level, although the significance level decreases when two close cohorts are examined, and particularly for later cohorts, similar to BHPS-USS. Figure 1 suggests a positive IGE trend in Germany as well. The IGE ranges from 0.466 (1950 cohort) to 0.711 (1980 cohort). This exercise illustrates that disaggregating the indicator of intergenerational elasticity can provide further insight. According to Table 1, β_{GSOEP} is 0.607 (similar to β_{PSID}, 0.614). However, whereas the IGE is constant in the United States, in Germany the elasticity increases in a statistically significant way across cohorts. Thus, despite the same average IGE, the intergenerational inequality is increasing in Germany. Summary statistics by cohort show that the proportion of West Germans does not change significantly over time. Thus, the higher value of IGE for later cohorts in Germany does not appear to be determined by a higher percentage of East Germans in the sample. This suggests that the negative mobility trend involves both East and West Germans.

Finally, in Appendix E the relationship between the IGE and the business cycles is investigated. The explorative analysis on macrodata suggests that the IGE may be correlated to business cycles, although it is not possible to identify the mechanisms and to what extent this relationship is robust. One channel might be the public expenditure in education, one of the key parameters in Solon's model. [47].

To sum up, the analysis of the trend of intergenerational elasticity indicates the importance of disaggregating the overall measure in order to better understand the implications of a given β in each country. Moreover, the section uncovered a possible link between the economic performance and the intergenerational transmission of earnings. The following step investigates whether the IGE is constant or changes according the economic status of the individual.

(b) Two-stage quantile regression

The previous section suggested that the IGE may not only change over time but also across different types of individuals. Specifically, it may not be constant along the individual's distribution of earnings. This would be consistent with the theoretical framework described in section I,

[47] Further details are provided in Appendix E

which maintains that less-favoured parents cannot afford an optimal investment in the education of their children. In order to assess whether IGE is constant along the sons' income distribution, equation 2 needs to be estimated at the relevant quantile, α, through the 2SQR (two-stage quantile regression)[48]. As explained by Cameron and Trivedi (2005, p. 88) and Koenker (2005, p. 5-25), the quantile regression measures the effect of the explanatory variables at a given quantile in the conditional distribution of the dependent variable. Koenker and Hallock (2001) provide a detailed explanation of quantile regression. The main difference between mean and quantile regression is the methodology used to estimate the coefficients. Whereas OLS are obtained by minimizing the sum of squared residuals, the estimators of quantile regression minimize the sum of absolute deviations. A main advantage of this methodology is its robustness to outliers. An outlying observation can still be highly influential, but not as much as with mean regression because the impact is bounded. That is, if the observations change on the y-direction outside the given quantile fit, the latter is not affected. In addition, this approach does not make assumptions about the parametric distribution of errors. These render quantile regression particularly suitable for skewed and heteroscedastic data. In particular, the estimated model follows eq.9:

$$(9) \quad Q_\alpha(y_s \mid y_f, X) = \hat{y}_f \rho(\alpha) + X\delta(\alpha) + \varepsilon(\alpha)$$

where $\rho(\alpha)$ indicates the IGE at quantile α, and X is the matrix of regressors of eq.2. If applied to intergenerational studies, differing slopes indicate different sensitivity of different portions of the son's distribution to small changes in father's income.

Table 45 in Appendix F suggests that the relationship between the earnings of fathers and sons is not constant. The impact of paternal income varies across the quantiles and it is not the same in the four countries. It is higher at both extremes of the sons' distribution in Italy, whereas in the United States it rises with the quantiles. In Germany, the transmission of earnings from fathers to sons appears rather constant. Finally, in the United Kingdom the relationship is increasing until the 85^{th} percentile and then decreasing. The pattern indicated by the coefficients is consistent with the results of the Wald test for the statistical difference between pairs of

[48] The first stage remains the same as above: least squares regression is used to predict paternal income. In the second stage, the predicted income is used as a regressor in the quantile regression.

slopes[49]. The number of intergenerational studies relying on quantile regression to investigate the IGE in Germany, US, Italy and UK is limited[50]. Among the first researchers to employ this methodology are Eide and Showalter (1999): they use quantile regression on 612 matched pairs of fathers and sons from PSID. They consider three-years averages of paternal earnings and averages of son's earnings from 1984 to 1992. The overall time period ranges from 1968 to 1992. The authors uncover non-linearities, although the pattern is somewhat contrasting with Table 45. Indeed, they find that IGE is higher at the bottom of the son's distribution. The authors obtain similar results by doing the same estimation using the High School and Beyond dataset, with over 5,000 observations. The difference might be explained by the two variables for the main analysis: they use family total income and not paternal earnings. Additionally, the analysis is performed for an earlier period. Thus, my research might be considered as an update of Eide and Showalter (1999), with a different methodology and the introduction of an additional tool to control for life-cycle bias[51]. To my knowledge, there are no comparable studies on the United Kingdom and Germany. For Italy, my results are consistent with Mocetti (2006). He uncovers higher sensitiveness at both extremes of the offspring's income distribution, although the statistical significance of their estimates is only confirmed for the median and the quantiles above it. Their estimates are higher than those in Table 45, but the difference could be due to the additional controls for life-cycle bias [52].

Further support to these results also emerges from studies analysing the correlation between the son's and paternal incomes at different levels of the fathers' income distribution. For example, Bratberg et al. (2007) compare Norway, Finland, Denmark (with register data), the Unites States (data from NLSY79) and the United Kingdom (NCDS). They explicitly account for non-linearities in the transmission of earnings through polynomials of different order according to

[49]More details about the Wald tests are provided in Appendix F

[50]Examples of studies for countries not included in this research are Corak and Heisz (1999), who uses non-parametric estimation techniques on Canadian men; Bratberg et al. (2007) for Norway who apply quantile regression to register data for Norwegians born in 1950, 1955, and 1960.

[51]Grawe (2004) also applies 2SQR. However, his study is not easily comparable with other literature: he only uses 354 matched pairs of sons and fathers and only two variables as instruments for the first stage, as mentioned in section VI. Grawe uses the same methodology, with the same shortcomings, for the United Kingdom and Germany.

[52]In fact, Table 1 indicate that, for Italy, β decreases when the interaction between paternal income and son‘s age is included.

the country. The authors estimate the elasticity at different percentiles of paternal earnings (10^{th}, 50^{th}, and 90^{th}). For all countries, the estimates of the elasticity at the 10th percentile are much lower than at the 50th, in turn lower than at the 90th, although the difference is less striking in the Anglo-Saxons countries.[53]. Finally, Couch and Lillard (2004) test for non-linearities in Germany and the United States by estimating the IGE of earnings and the log of earnings with higher order terms in paternal earnings. Whereas for the United States, the higher order polynomials are statistically significant in the majority of cases, in Germany it is less so. Only the parameter for the squared term in the level specifications is statistically different from zero. This is consistent with the results of the quantile regression on German data, where IGE does not change significantly along the son's income distribution, except at the 95^{th} percentile.

(c) Ordered logit transition matrices

The quantile regression does not provide explicit information about the father's income quantile. Therefore, further indications about mobility patterns require an alternative method. The transition matrices are square matrices that estimate the relative frequencies (or probabilities) of moving from one income quantile to another, conditional on the paternal quantile. The income classes can vary and are usually defined according to the quantile of the marginal distribution. More frequently, the width of the classes is based on quartiles, quintiles or deciles. Some studies, especially those concerned with top income literature, employ percentiles. Most articles using transition matrices suggests that mobility tends to be lower at the tail end of the distribution, and even more so at the top. For the United Kingdom, Dearden et al. (1997), for example, analyse earnings and school years of father-son and father-daughter pairs from the UK National Child Development Survey (NCDS). The analysis of the transition matrix suggests that downward mobility from the top quartile is rarer than upward mobility from the bottom quartile. Similar conclusions are drawn by Blanden et al. (2005) who use quartile transition matrices on

[53]For Nordic countries, this shows that the elasticity in lower percentiles might be overestimated; it highlights as well that overall differences between Anglo-Saxon and Nordic countries might be overestimated by the linear model. Indeed, differences are smaller in the middle and at the top of the distribution than at the bottom. These conclusions are consistent with Björklund et al. (2012), who uses non-linear regressions by means of a spline function with pre-defined knots corresponding to paternal income percentiles. They use a large sample of Swedish fathers and sons from register data.

data from NCDS and BCS. Corak and Heisz (1999) and Mazumder (2005b) use decile transition matrices to assess mobility, respectively in Canada (with register data) and in the United States (PSID). Corak and Piraino (2010) review the two studies after arguing that the selected sample is similar. In particular, they compare the matrices for the sons whose fathers are in the bottom or in the top deciles. In both countries the probability of the sons of being in the same decile as their father is higher than that of being in other deciles. Moreover, it gradually decreases the further the decile is from the father's. Additionally, downward mobility from the top and upward mobility from the bottom are less likely in the United States than in Canada. The above results find confirmation in Jantti et al. (2006), who compare Anglo-Saxons and Nordic countries. The authors find that whereas mobility in the middle quintiles is similar, the UK and the US show stronger persistence in both ends of the distribution, which according to their estimates accounts for the majority of the difference between the two groups of countries. A similar picture emerges from studies on Italian data. Checchi et al.(1999) and Piraino (2007) use transition matrices on a different dataset and with a different methodology. Specifically, the former use 1,666 matched pairs of fathers and sons on a survey conducted in 1985, the Indagine Nazionale sulla Mobilità Sociale whereas the latter is based on SHIW and the TS2SLS procedure. Additionally, Checchi et al. do not use earnings but an indicator of occupational mobility[54]. Despite the differences, both studies uncover higher persistence at the extremes of the distribution. Finally, Checchi et al. use the same technique and the same variable to investigate mobility patterns in the United States (1050 father-son couples from PSID) and compare the two countries. The United States are also characterised by lower mobility at the extremes of the occupational ranking: specifically, the persistence at the top is higher in Italy, whereas the persistence at the bottom is higher in the United States.

Despite the large number of literature on transition matrices, two main shortcomings can be highlighted. One criticism is that the extent of mobility might depend on the selected discrete categorisation. For example, fewer categories (for example, quartiles) might underestimate mobility with respect to other discretisations (for example, percentiles). Or else, lower mobility

[54]The authors explain that occupations are ranked according to the occupational median income in the generation of the children.

at the extremes of the distribution might only be a mechanical consequence of the matrix design, particularly of the existence of floors and ceilings which impede the individuals at the bottom to go further down and those at top to move further up. Overcoming this problem is not straightforward if the goal is to maintain all the information provided by the transition matrices. For example, Black and Devereux (2011) mention Bhattacharya and Mazumder (2008) who propose an alternative measure. They use the probability that a son's percentile in the earnings distribution of sons exceeds the father's percentile in the earnings distribution of fathers. This measure, however, does not capture the extent of mobility, as it does not estimate of how much the son has exceeded the father. An alternative solution is to focus on mobility indices (Jantti et al., 2006), which take into account the different elements and diagonals of the matrix. However, this does not provide information about the direction of mobility either. A way to check whether the results are driven from matrix construction is suggested by Corak and Heisz (1998). If the higher persistence towards the extremes of the distribution was exclusively the result of a floor-ceiling effect, then only the top and bottom income quantiles would present significant spikes in the transition probabilities. Instead if that characteristic is common to the neighbouring classes, then something else is at work. Similarly, if a finer disaggregation (in their case decile versus quartiles) does not reduce the spikes significantly, then the higher persistence at the extremes can not be explained solely by the matrix design.

Another criticism is that transition matrices describe probability transitions from an income quantile to another only conditioning on the corresponding paternal income quantile. Indeed, the IGE literature so far has used Markovian matrices to indicate the proportion of individuals in each earning class, given the paternal income quantile, usually with no further controls. As indicated by Zimmerman (1992) the interpretation of transition matrices is complicated by the fact that this type of specification cannot account for life-cycle bias. The only exception, to my knowledge, is Blanden et al. (2005), who control for education[55]. As mentioned in section VI, averaging earnings over several years can reduce the bias [56].

[55]In a way, Jantti et al. (2006) correct for the ages of the fathers and the sons, although to my understanding this is done in the calculation of the paternal income.

[56]For example, Jantti et al. (2006) does it as a sensitivity analysis with transition matrices

This research suggests an alternative. I estimate an ordered logistic regression, where the dependent variable is a categorical variable with as many categories as the number of income quantiles. The other covariates are a categorical variable for the income quantiles of the father, and the covariates of eq.6. I then compute the predicted probabilities for each outcome given the paternal income class and I recompose the transition matrix. Two similar models are available for this purpose: the ordered logit and the generalised ordered logit[57]. The generalized ordered logit is more efficient than models that do not account for data ordering, such as the multinomial logit. It is however less parsimonious than the ordered logit. The main difference is that the former estimates a series of coefficients (including one for the constant) for all the j points at which the dependent variable can be dichotomised, whereas the latter assumes that the coefficients are the same (Fu, 1999). The ordered logit relies on the proportional odds assumption, according to which the coefficients of the covariates across a series of cumulative logits on the response variable outcomes do not change. If this assumption is violated, then the coefficients of the ordered logit might not correctly capture the relationship between the dependent and independent variables (Williams, 2013).

In order to select the type of ordered model suitable to the data, this assumption was tested on the four datasets with the Brant test, on a five-category dependent variable[58]. For all datasets the overall test rejects the proportional odds assumption with a high χ^2 statistic and with a p-value equal to 0[59]. Additional test statistics also suggest that the generalised ordered logit performs

[57]The two models are described in Cameron and Trivedi (2005, p. 519-520), Williams (2006); Fu (1999).

[58]Under the null, the Brant test assumes that the slope coefficients of the four cumulative binary logits implied by the ordered regression model are the same.

[59]As per test on the individual regressors, it needs to be highlighted that with SHIW the failure of the Brant test might rather be attributed to other regressors, such as the year of birth, and not to the dummies indicating the paternal income quantiles. Instead, with GSOEP, PSID and BHPS-USS also these variables contribute to the violation of the assumption. In general, the higher quintiles are more likely to violate the parallel odds assumption. Instead, the first and the second paternal quintiles tend to have a statistically similar impact across the cut-off points of the response variable, at least at 5 percent significance level. The exception is the bottom quintile with GSOEP that violates the assumption at any standard level of significance.

better than the ordered logit for all datasets[60]. The generalised ordered model is:

$$
\begin{aligned}
\Pr(y &= j|c) = \Pr(\varepsilon \leq \alpha_j - \sum \lambda_{ij} q_{fi} - \delta_j X) \\
&= \frac{\exp\ (\alpha_j + \sum \lambda_{ij} q_{fi} + \delta_j X)}{1 + [\exp\ (\alpha_j + \sum \lambda_{ij} q_{fi} + \delta_j X)]}\ , \qquad (10) \\
\text{with } j &= 1, 2, ..m-1, \\
\text{and } i &= 1, 2, ..q \qquad (11)
\end{aligned}
$$

where α_j are the $m-1$ cut-off points, q_{fi} are the q income quantiles of the father, X is the matrix of of eq.2[61] and ε are the errors from the following latent variable framework:

$$
\begin{aligned}
y^* &= \sum \lambda_{ij} q_{fi} + \delta_j X + \varepsilon \\
\text{where } y &= \begin{cases} 1\ if\ \ y^* \leq \alpha_1 \\ 2\ if\ \ \alpha_1 < y^* \leq \alpha_2 \\ ... \\ j\ if\ \ y^* > \alpha_{m-1} \end{cases} \qquad (12)
\end{aligned}
$$

For the main analysis $j = i = 5$. The choice to divide fathers and sons according to the quintiles of their income distributions is suggested by data characteristics. Specifically, the smallest dataset, SHIW, has $6,860$ observations. Considering five categories leaves a large number of observations for each outcome.

Figure 2 illustrates the probabilities on the leading diagonal (left panel) and on the antidiagonal (right panel). The complete matrix is available in Table 46 in Appendix G[62]. The elements of the main diagonal have higher probabilities than the one associated with perfect mobility

[60]Other χ^2 statistics (the Wald test, the Likelihood Ratio test, the Wolfe-Gould test and the score test) were also checked. The results are consistent with the Brant test for the overall model. Finally, the generalized ordered logit and the ordered logit were compared with the Aikake and Bayesian Information Criteria (AIC and BIC respectively), although the results are contrasting. Indeed, whereas the AIC suggests that the former minimizes the information loss for all datasets; for Germany, the United Kingdom and Italy the BIC is in favour of the ordered logit.

[61]Matrix X includes the son's age normalised at 40, its square and cube; the interaction between normalised age (squared, and cubed) and paternal income; father's and son's years of birth. Finally,a categorical variable for paternal age. Specifically, the variable includes four age categories: up to 43 years of age; from 44 to 48; from 49 to 53; and from 54 to 59. These variables are the controls specified in eq.6.

[62]The same table also reports the probabilities calculated from an ordered logit to check how and whether the results change. It appears that when the probabilities are predicted from an ordered logit the variation is not elevated and the patterns that characterise the probabilities computed on this specification are similar to those predicted on the basis of the generalised ordered logit. As expected, this is particularly evident for the Italian case.

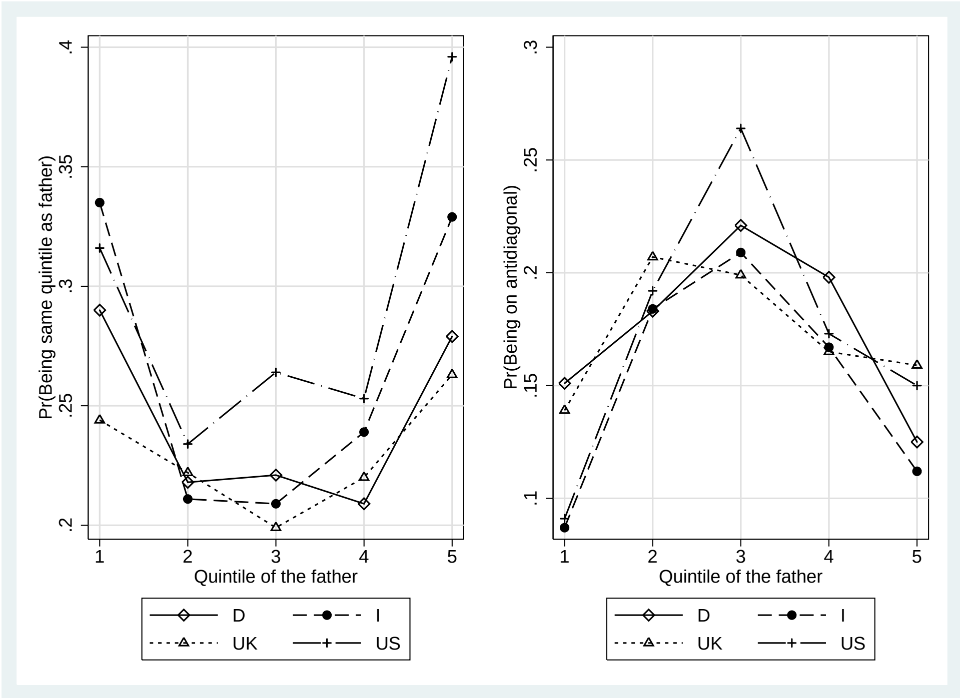

FIGURE 2
ORDERED LOGIT TRANSITION MATRIX: PROBABILITIES ON THE MAIN AND MINOR DIAGONALS

(0.20), except for $\Pr(y = 3|q_{f,3}) = 0.199$ in the United Kingdom. Overall, the results are in line with the literature based on the Markov transition matrices that is outlined above. All datasets are characterised by higher probabilities at the top and bottom quintiles of the main diagonal. Specifically, the United States has the highest persistence at the top, that is to say a lower extent of downward mobility: the sons whose fathers are in the top quintile have 39 percent probability of being in the top quintile as well, whereas the probability ranges from 10 to 15 percent if their father is in the first three quintiles. In Italy and Germany the values at the bottom and top quintiles on the main diagonal are similar, although their magnitude is lower in Germany (0.329 versus 0.335, and 0.29 versus 0.279, respectively). Finally, in the United Kingdom, the probability of upward mobility from the bottom quintile is higher than that of downward mobility from the top; particularly, the $\Pr(y = 1|q_{f,1}) = 0.244$ is the lowest among the four countries. It is also worth highlighting that the likelihood of being in a given quintile is lower the further that quintile is from the same paternal quintile. For example, in the United States the probability of being in the fifth quintile decreases from 0.396 if $i = 5$ to 0.091 if $i = 1$. Additionally, for all countries the individuals with the richest families have more chance to be in lowest quintile than the sons of the poorest fathers have to be in the highest quintile (that is, $\Pr(y=5|q_{f,1}) < \Pr(y=1|q_{f,5})$).

These results are reinforced, statistically, by the failure of the parallel line assumption. The implication is that the effect of the paternal quantile on the probability of the son of being in a certain quantile differs according to the income class of the son. This violation can be interpreted as a confirmation and a reinforcement of the results of the quantile regression, according to which the impact of paternal income changes along the offspring's income distribution; and it provides some additional information. Indeed, the statistical tests on the coefficients suggest that within a given income quantile j, λ_{ij} are statistically different from each other in several cases, and especially for higher paternal quintiles. Thus, the findings suggest that not only the effect of paternal income changes with the son's income quantile, but they also imply that the extent of the effect within a given j may be affected by the specific income quantile of the father. This is particularly evident for high earning fathers.

The above conclusions provide some evidence against the claim that higher persistence at the extremes is mainly determined by a floor-ceiling effect. Particularly, higher probabilities characterise the extremes as well as the elements close to them (the first and the second of the first two rows, the fourth and fifth of the last two in Table 46). As a further check the generalised ordered logit is re-estimated on three different subsamples. The first subsample excludes the 1^{st} and the 20^{th} vintile of the earnings distributions of both fathers and sons; the second excludes the 1^{st} and 10^{th} deciles; the third, the 1^{st} and 5^{th} quintiles. The expectation is that the higher the reduction, the lower the persistence at the extremes. In fact, if the higher persistence is only caused by the fact that the individuals at the extremes of the leading diagonal cannot move further (up or down) then excluding the lowest and the highest percentiles of the distribution would not affect them. However, we do not expect higher persistence to disappear completely, as Table 46 illustrates that it is not only a feature of the extremes, it also involves the neighbouring quintiles. The results are reported in Appendix G and confirm our claim. As expected, the extreme values gradually decrease with the sample size.

Another possible objection is that the size of the quantile might affect the results. In particular, if higher persistence was caused by the fact that a quintile is a broad category and that as such it underestimates mobility, then a reduction in the size of the quantile should reduce the probabilities at the extremes. To check the impact of a finer disaggregation on the probabilities at the extremes of the distribution, the generalised ordered logit has been estimated on deciles. Table 51 in Appendix G reports the results of this task. The increase in the number of categories does not reduce the persistence at the tails: instead, in certain cases it appears to enhance it. What is interesting is that even with a finer categorisation the patterns are the same of Table 46: for the majority of datasets, at least the first three elements of the first two rows and the last three of the last two rows are characterised by higher persistence[63].

VII CONCLUSIONS

This article analysed the patterns of intergenerational transmissions of earnings in Germany, Italy, the United Kingdom and the United States. By doing so, it filled the gap in the intergenerational literature at two levels. Firstly, it provided comparable cross-country estimates of IGE over time and in terms of mobility patterns. Secondly, it calculated a matrix of transition probabilities from a generalised ordered logit with control variables to account for life-cycle and attenuation biases.

The main findings indicated that the United States and Germany have the highest value of the IGE, followed by Italy and the United Kingdom. This was the first research on IGE that includes East Germany. The results suggested that doing so increases the estimated IGE, which explains the significantly lower values for Germany in the existing literature. Moreover, this was the first study to analyse the IGE trend in Italy and Germany: the elasticity appeared to increase across cohorts, as in the UK.

In terms of mobility patterns, the transition matrices suggested higher levels of immobility for very rich and very poor individuals, especially for the United States and Italy. With PSID, the individuals whose fathers are in the top income quintile have 40 percent chance of being

[63]These results present some variation across countries: for example, for the United States this would also apply to the last three elements of the eighth row and the first three of the third row; instead in the United Kingdom the bottom decile is not characterised by a high level of persistence, whereas the second row is.

there as well. A similar pattern emerged with GSOEP but the chance of downward mobility if at the top, or of upward mobility if at the bottom are higher than in Italy and the United States. Finally, the United Kingdom is the country with the lowest level of persistence at the extreme quintiles of the income distribution.

The above results also provided further support the claim stated at the beginning of this article about the possible challenges of cross-country review articles. For example, the four countries might be ranked differently if the outcome of each exercise was considered separately. In terms of equality of opportunities, the United States and Italy appear as the countries where the paternal income has the highest influence. It might be wise to rank Germany on a lower level because despite the elevated and increasing IGE, the transitional probabilities suggest that the individuals are more likely to reach earnings quintiles further from their fathers'. Finally, in the United Kingdom the career opportunities are the least related to the father's own status.

As mentioned in the outset of this paper, the theoretical literature has mainly explored the effect of parental income through the investment in their offspring's education. The explorative exercise in Appendix H suggested that the role of education in promoting intergenerational mobility seems to decrease across the cohorts. Additionally, conditioning on education does not eliminate the higher dependence on paternal income at the extreme quintiles of the transition matrices, at least in Italy and US. This calls for further research, empirical and theoretical, on the drivers of intergenerational immobility.

REFERENCES

AARONSON, D. AND B. MAZUMDER (2008): "Intergenerational economic mobility in the United States, 1940 to 2000," *Journal of Human Resources*.

ANGRIST, J. D. AND A. B. KRUEGER (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, 328–336.

ARELLANO, M. AND C. MEGHIR (1992): "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets," *The Review of Economic Studies*, 59, 537–559.

BECKER, G. S. AND N. TOMES (1986): "Human capital and the rise and fall of families," *Journal of Labor Economics*, Part 2 4(3.

BHATTACHARYA, D. AND B. MAZUMDER (2008): "Nonparametric analysis of intergenerational income mobility with application to the United States," *Forthcoming Journal of Econometrics*.

BJÖRKLUND, A. AND M. JÄNTTI (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *The American Economic Review*, 87, 1009–1018.

——— (2009): "Intergenerational income mobility and the role of family background," in *Oxford Handbook of Economic Inequality, Oxford ...*, 1–66.

BJÖRKLUND, A., J. ROINE, AND D. WALDENSTRÖM (2012): "Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity?" *Journal of Public Economics*, 96, 474–484.

BLACK, S. AND P. DEVEREUX (2011): "Recent developments in intergenerational mobility," in *Handbook of labor economics*, vol. 4, chap. Volume 4,, 1487–1541.

BLANDEN, J. (2009): "How much can we learn from international comparisons of intergenerational mobility?" Tech. rep., Sutton Trust.

——— (2013): "Cross-Country Rankings In Intergenerational Mobility: A Comparison Of Approaches From Economics And Sociology," *Journal of Economic Surveys*, 27, 38–73.

BLANDEN, J., A. GOODMAN, P. GREGG, AND S. MACHIN (2004): "Changes in intergenerational mobility in Britain," in *Generational income mobility in North America and Europe*, 122–46.

BLANDEN, J., P. GREGG, S. MACHIN, AND W. TRUST (2005): "Intergenerational mobility in Europe and North America: a report supported by the Sutton Trust," .

BLANDEN, J., P. GREGG, AND L. MACMILLAN (2007): "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education," *The Economic Journal*, 117, C43–C60.

BRATBERG, E., O. A. NILSEN, AND K. VAAGE (2007): "Trends in intergenerational mobility across offspring's earnings distribution in Norway," *Industrial Relations: A Journal of Economy and Society*, 46, 112–129.

BRATSBERG, B., K. RØ ED, O. RAAUM, R. NAYLOR, T. ERIKSSON, AND OTHERS (2007): "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons*," *The Economic Journal*, 117, C72–C92.

CAMERON, A. C. AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*, Cambridge University Press.

CORAK, M. AND A. HEISZ (1999): "The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data," *Journal of Human Resources*, 34, 504–533.

CORAK, M. AND P. PIRAINO (2010): "Intergenerational earnings mobility and the inheritance of employers," *IZA Discussion Paper No. 4876*.

COUCH, K. AND D. LILLARD (2004): "Non-linear patterns of intergenerational mobility in Germany and the United States," in *Generational Income Mobility in North America and Europe*, 147–89.

DEARDEN, L., S. MACHIN, AND H. REED (1997): "Intergenerational mobility in Britain," *The Economic Journal*, 47–66.

EIDE, E. R. AND M. H. SHOWALTER (1999): "Factors Affecting the Transmission of Earnings across Generations: A Quantile Regression Approach," *The Journal of Human Resources*, 34, 253–267.

ERMISCH, J. AND M. FRANCESCONI (2004): "Intergenerational mobility in Britain: new evidence from the British Household Panel Survey," in *Generational Income Mobility in North America and Europe*, 147–89.

FINNIE, R. AND I. IRVINE (2006): "Mobility and gender at the top tail of the earnings distribution," *Economic and Social Review*, 37, 149.

FRANCESCONI, M. AND C. NICOLETTI (2006): "Intergenerational mobility and sample selection in short panels," *Journal of Applied Econometrics*, 21, 1265–1293.

FU, V. K. (1999): "Estimating generalized ordered logit models," *Stata Technical Bulletin*, 8.

GOLDTHORPE, J. H. AND K. HOPE (1974): *The social grading of occupations: A new approach and scale*, Clarendon Press Oxford.

GOLDTHORPE, J. H. AND M. JACKSON (2007): "Intergenerational class mobility in contemporary Britain: political concerns and empirical findings1," *The British Journal of Sociology*, 58, 525–546.

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

——— (2006): "Lifecycle bias in estimates of intergenerational earnings persistence," *Labour Economics*, 13, 551–570.

HAIDER, S. AND G. SOLON (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *The American Economic Review*, 96, 1308–1320.

HERTZ, T. (2007): "Trends in the Intergenerational Elasticity of Family Income in the United States," *Industrial Relations: A Journal of Economy and Society*, 46, 22–50.

INOUE, A. AND G. SOLON (2010): "Two-Sample Instrumental Variables Estimators," *The Review of Economics and Statistics*, 92, 557–561.

JANTTI, M., B. BRATSBERG, AND K. RØ ED (2006): "American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States," *Available at SSRN . . .* .

KOENKER, R. (2005): *Quantile Regression*, Cambridge University Press.

KOENKER, R. AND K. HALLOCK (2001): "Quantile regression: An introduction," *Journal of Economic Perspectives*, 15, 43–56.

LEE, C.-I. AND G. SOLON (2009): "Trends in Intergenerational Income Mobility," *Review of Economics and Statistics*, 91, 766–772.

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, 235–255.

——— (2005b): "The apple falls even closer to the tree than we thought," in *Unequal chances: Family background and economic success*, Princeton University Press.

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

NICOLETTI, C. AND J. F. ERMISCH (2007): “Intergenerational Earnings Mobility: Changes across Cohorts in Britain,” *The B.E. Journal of Economic Analysis & Policy*, 7.

PIRAINO, P. (2007): “Comparable Estimates of Intergenerational Income Mobility in Italy,” *The B.E. Journal of Economic Analysis & Policy*, 7.

SOLON, G. (1992): “Intergenerational Income Mobility in the United States,” *The American Economic Review*, 82, 393–408.

——— (1999): “Intergenerational mobility in the labor market,” Elsevier, vol. Volume 3,, 1761–1800.

——— (2004): “A model of intergenerational mobility variation over time and place,” in *Generational income mobility in North America and Europe*, 38–47.

WILLIAMS, R. (2006): “Generalized ordered logit/partial proportional odds models for ordinal dependent variables,” *Stata Journal*, 6, 58–82.

——— (2013): “Generalized Ordered Logit/ Partial Proportional Odds Models,” *forthcoming*.

ZIMMERMAN, D. (1992): “Regression toward mediocrity in economic stature,” *The American Economic Review*, 409–429.

APPENDIX A: SUMMARY STATISTICS

TABLE 2
GSOEP: SUMMARY STATISTICS FOR SONS

	AV3	MIN40	A55	ALL
ln (Labour Income)	10.475	10.259	10.484	10.399
	(0.69)	(0.93)	(0.82)	(0.84)
Age	41.315	40.055	43.501	41.611
	(6.86)	(5.44)	(5.60)	(7.30)
Year of birth	1962.981	1964.025	1961.261	1962.588
	(7.06)	(7.69)	(5.89)	(6.79)
Less than secondary education	0.058	0.064	0.050	0.056
	(0.23)	(0.24)	(0.22)	(0.23)
Secondary education	0.015	0.014	0.012	0.016
	(0.12)	(0.12)	(0.11)	(0.12)
Vocational training	0.659	0.644	0.678	0.671
	(0.47)	(0.48)	(0.47)	(0.47)
University education	0.267	0.278	0.261	0.258
	(0.44)	(0.45)	(0.44)	(0.44)
self	0.126	0.134	0.132	0.123
	(0.33)	(0.34)	(0.34)	(0.33)
N. of averaged LI obs.	7.809	1.000	8.052	9.033
	(3.49)	(0.00)	(3.45)	(3.83)
N	3232	4221	19563	27442

TABLE 3
SHIW: SUMMARY STATISTICS FOR SONS

	AV3	MIN40	A55	ALL
ln (Labour Income)	9.906	9.679	9.768	9.749
	(0.49)	(0.68)	(0.65)	(0.65)
Age	43.456	41.653	43.439	42.501
	(6.06)	(5.68)	(5.39)	(6.69)
Year of birth	1959.155	1959.773	1958.611	1959.263
	(6.09)	(6.28)	(5.72)	(6.21)
Up to primary educ.	0.059	0.070	0.065	0.063
	(0.24)	(0.25)	(0.25)	(0.24)
Lower sec educ.	0.376	0.420	0.409	0.412
	(0.48)	(0.49)	(0.49)	(0.49)
Upper sec educ.	0.447	0.398	0.419	0.415
	(0.50)	(0.49)	(0.49)	(0.49)
Higher education	0.102	0.113	0.107	0.110
	(0.30)	(0.32)	(0.31)	(0.31)
self	0.244	0.262	0.243	0.245
	(0.43)	(0.44)	(0.43)	(0.43)
N. of averaged LI obs.	4.537	2.443	3.615	3.886
	(1.77)	(1.95)	(2.24)	(2.47)
N	1027	2740	5860	6860

TABLE 4
BHPS: SUMMARY STATISTICS FOR SONS

	AV3	MIN40	A55	ALL
ln (Labour Income)	10.508	10.402	10.519	10.467
	(0.49)	(0.73)	(0.66)	(0.66)
Age	40.055	39.509	42.613	40.622
	(5.01)	(2.10)	(5.25)	(6.56)
Year of birth	1961.078	1960.370	1958.508	1959.892
	(6.62)	(6.15)	(5.41)	(6.07)
No completed ed.	0.056	0.067	0.055	0.053
	(0.23)	(0.25)	(0.23)	(0.22)
Some ed.	0.340	0.319	0.302	0.319
	(0.47)	(0.47)	(0.46)	(0.47)
Further ed.	0.392	0.378	0.390	0.383
	(0.49)	(0.49)	(0.49)	(0.49)
University ed.	0.212	0.236	0.252	0.245
	(0.41)	(0.42)	(0.43)	(0.43)
self	0.109	0.142	0.112	0.110
	(0.31)	(0.35)	(0.32)	(0.31)
N. of averaged LI obs.	12.789	1.035	12.584	14.210
	(4.47)	(0.18)	(4.12)	(4.00)
N	1144	834	8363	11976

TABLE 5
PSID: SUMMARY STATISTICS FOR SONS

	AV3	MIN40	A55	ALL
ln (Labour Income)	10.824	10.759	10.855	10.836
	(0.66)	(0.99)	(0.95)	(0.92)
Age	37.390	38.577	43.640	41.871
	(4.19)	(3.20)	(5.49)	(7.02)
Year of birth	1963.889	1966.238	1960.121	1961.701
	(6.65)	(5.51)	(5.77)	(6.78)
Up to 11th grade	0.068	0.067	0.059	0.061
	(0.25)	(0.25)	(0.24)	(0.24)
High school	0.321	0.336	0.336	0.327
	(0.47)	(0.47)	(0.47)	(0.47)
Further education	0.252	0.256	0.229	0.236
	(0.43)	(0.44)	(0.42)	(0.42)
University	0.357	0.341	0.375	0.375
	(0.48)	(0.47)	(0.48)	(0.48)
self	0.113	0.130	0.139	0.128
	(0.32)	(0.34)	(0.35)	(0.33)
N. of averaged LI obs	7.603	1.000	7.466	8.937
	(3.45)	(0.00)	(3.50)	(3.57)
N	849	945	5148	7488

TABLE 6
GSOEP: FIRST STAGE REGRESSIONS, INSTRUMENTS

	AV3f	AV3s	MIN40f	MIN40s	A55f	A55s	ALLf	ALLs
F Age 30-43	0.024	0.460	0.134	0.437	0.065	0.498	0.056	0.465
	(0.15)	(0.50)	(0.34)	(0.50)	(0.25)	(0.50)	(0.23)	(0.50)
F Age 44-48	0.356	0.313	0.265	0.329	0.357	0.315	0.304	0.312
	(0.48)	(0.46)	(0.44)	(0.47)	(0.48)	(0.46)	(0.46)	(0.46)
F Age 49-53	0.405	0.151	0.295	0.158	0.436	0.155	0.372	0.149
	(0.49)	(0.36)	(0.46)	(0.36)	(0.50)	(0.36)	(0.48)	(0.36)
F Age 54-59	0.215	0.076	0.306	0.077	0.141	0.032	0.269	0.074
	(0.41)	(0.27)	(0.46)	(0.27)	(0.35)	(0.18)	(0.44)	(0.26)
F Turkish national	0.057	0.081	0.041	0.091	0.051	0.066	0.049	0.075
	(0.23)	(0.27)	(0.20)	(0.29)	(0.22)	(0.25)	(0.22)	(0.26)
F Secondary school	0.412	0.641	0.463	0.628	0.439	0.671	0.434	0.650
	(0.49)	(0.48)	(0.50)	(0.48)	(0.50)	(0.47)	(0.50)	(0.48)
F Intermediate school	0.194	0.142	0.204	0.152	0.219	0.136	0.211	0.143
	(0.40)	(0.35)	(0.40)	(0.36)	(0.41)	(0.34)	(0.41)	(0.35)
F Technical school	0.022	0.003	0.037	0.002	0.028	0.001	0.030	0.003
	(0.15)	(0.05)	(0.19)	(0.05)	(0.16)	(0.03)	(0.17)	(0.05)
F Upper secondary school	0.209	0.132	0.181	0.137	0.184	0.124	0.191	0.126
	(0.41)	(0.34)	(0.39)	(0.34)	(0.39)	(0.33)	(0.39)	(0.33)
F Other educ.	0.117	0.041	0.084	0.041	0.092	0.034	0.096	0.042
	(0.32)	(0.20)	(0.28)	(0.20)	(0.29)	(0.18)	(0.29)	(0.20)
F No degree	0.045	0.042	0.032	0.039	0.038	0.035	0.039	0.036
	(0.21)	(0.20)	(0.18)	(0.19)	(0.19)	(0.18)	(0.19)	(0.19)
F Worker	0.357	0.420	0.309	0.426	0.316	0.429	0.314	0.418
	(0.48)	(0.49)	(0.46)	(0.49)	(0.46)	(0.49)	(0.46)	(0.49)
F Foreman, supervisor	0.061	0.076	0.066	0.072	0.073	0.079	0.070	0.079
	(0.24)	(0.27)	(0.25)	(0.26)	(0.26)	(0.27)	(0.26)	(0.27)
F Self-employed	0.099	0.136	0.136	0.132	0.116	0.133	0.120	0.136
	(0.30)	(0.34)	(0.34)	(0.34)	(0.32)	(0.34)	(0.33)	(0.34)
F White collar	0.355	0.268	0.370	0.268	0.383	0.262	0.379	0.268
	(0.48)	(0.44)	(0.48)	(0.44)	(0.49)	(0.44)	(0.49)	(0.44)
F Civil servant	0.128	0.099	0.119	0.101	0.113	0.097	0.117	0.099
	(0.33)	(0.30)	(0.32)	(0.30)	(0.32)	(0.30)	(0.32)	(0.30)
dadloc1989	0.881	0.845	0.851	0.853	0.863	0.844	0.867	0.847
	(0.32)	(0.36)	(0.36)	(0.35)	(0.34)	(0.36)	(0.34)	(0.36)
Father's ln(LI)	10.088		9.880		10.029		10.038	
	(0.55)		(0.70)		(0.63)		(0.66)	
N. of averaged LI obs.	6.173	7.809	1.000	1.000	6.138	8.052	6.480	9.033
	(2.96)	(3.49)	(0.00)	(0.00)	(3.34)	(3.45)	(3.92)	(3.83)
Predicted F's ln(LI)		9.260		9.370		9.422		9.269
		(0.53)		(0.53)		(0.44)		(0.51)
N	536	3232	1330	4221	3695	19563	4534	27442

f: from the synthetic fathers; s: reported information of sons about real fathers
F stands for father
LI stands for Labour Income

TABLE 7
SHIW: FIRST STAGE REGRESSIONS, INSTRUMENTS

	AV3f	AV3s	MIN40f	MIN40s	A55f	A55s	ALLf	ALLs
F No education	0.022	0.138	0.025	0.166	0.019	0.153	0.053	0.145
	(0.15)	(0.35)	(0.16)	(0.37)	(0.14)	(0.36)	(0.22)	(0.35)
F Primary education	0.234	0.576	0.357	0.546	0.281	0.574	0.331	0.568
	(0.42)	(0.49)	(0.48)	(0.50)	(0.45)	(0.49)	(0.47)	(0.50)
F Lower sec. education	0.292	0.176	0.317	0.171	0.348	0.167	0.316	0.169
	(0.46)	(0.38)	(0.47)	(0.38)	(0.48)	(0.37)	(0.47)	(0.37)
F Upper sec. ed.	0.336	0.079	0.234	0.082	0.273	0.076	0.221	0.084
	(0.47)	(0.27)	(0.42)	(0.27)	(0.45)	(0.26)	(0.42)	(0.28)
F University ed.	0.078	0.031	0.067	0.035	0.078	0.031	0.079	0.033
	(0.27)	(0.17)	(0.25)	(0.18)	(0.27)	(0.17)	(0.27)	(0.18)
F Production worker	0.268	0.519	0.346	0.524	0.346	0.523	0.320	0.517
	(0.44)	(0.50)	(0.48)	(0.50)	(0.48)	(0.50)	(0.47)	(0.50)
F Teacher, clerk	0.288	0.163	0.237	0.151	0.290	0.163	0.343	0.165
	(0.45)	(0.37)	(0.43)	(0.36)	(0.45)	(0.37)	(0.48)	(0.37)
F Junior manager, officer	0.027	0.036	0.061	0.037	0.052	0.038	0.049	0.036
	(0.16)	(0.19)	(0.24)	(0.19)	(0.22)	(0.19)	(0.22)	(0.19)
F Manager	0.023	0.013	0.021	0.016	0.020	0.013	0.022	0.015
	(0.15)	(0.11)	(0.14)	(0.13)	(0.14)	(0.11)	(0.15)	(0.12)
F Self-employed	0.354		0.335		0.293		0.266	
	(0.48)		(0.47)		(0.46)		(0.44)	
F Agriculture	0.089	0.189	0.086	0.233	0.067	0.209	0.134	0.204
	(0.29)	(0.39)	(0.28)	(0.42)	(0.25)	(0.41)	(0.34)	(0.40)
F Industry	0.346	0.289	0.375	0.266	0.395	0.262	0.257	0.269
	(0.48)	(0.45)	(0.48)	(0.44)	(0.49)	(0.44)	(0.44)	(0.44)
F Gov, commerce, craft	0.557	0.523	0.539	0.501	0.538	0.528	0.610	0.527
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)
F Northern Italy	0.410	0.443	0.424	0.396	0.418	0.415	0.000	0.419
	(0.49)	(0.50)	(0.49)	(0.49)	(0.49)	(0.49)	(0.00)	(0.49)
F Central Italy	0.176	0.149	0.206	0.181	0.198	0.167	0.000	0.165
	(0.38)	(0.36)	(0.40)	(0.39)	(0.40)	(0.37)	(0.00)	(0.37)
F Southern Italy	0.414	0.408	0.370	0.423	0.384	0.418	1.000	0.416
	(0.49)	(0.49)	(0.48)	(0.49)	(0.49)	(0.49)	(0.00)	(0.49)
F Age 30-43	0.000	0.504	0.021	0.695	0.016	0.546	0.010	0.585
	(0.00)	(0.50)	(0.14)	(0.46)	(0.13)	(0.50)	(0.10)	(0.49)
F Age 44-48	0.100	0.253	0.187	0.167	0.223	0.243	0.143	0.205
	(0.30)	(0.43)	(0.39)	(0.37)	(0.42)	(0.43)	(0.35)	(0.40)
F Age 49-53	0.538	0.190	0.316	0.100	0.484	0.171	0.297	0.144
	(0.50)	(0.39)	(0.47)	(0.30)	(0.50)	(0.38)	(0.46)	(0.35)
F Age 54-59	0.362	0.053	0.476	0.037	0.277	0.040	0.550	0.066
	(0.48)	(0.22)	(0.50)	(0.19)	(0.45)	(0.20)	(0.50)	(0.25)
F ln(LI)	9.516		9.345		9.382		9.219	
	(0.52)		(0.62)		(0.56)		(0.66)	
N. of averaged LI obs.	4.027		1.000		2.809		3.074	
	(1.21)		(0.00)		(1.57)		(1.80)	
F Self-employed		0.027		0.041		0.034		0.035
		(0.16)		(0.20)		(0.18)		(0.18)
F year of birth		1926.501		1927.385		1925.806		1926.609
		(9.63)		(9.52)		(9.12)		(9.51)
Predicted F ln(LI)		8.217		8.533		8.799		8.417
		(0.67)		(0.47)		(0.36)		(0.50)
N	226	1027	722	2740	941	5860	643	6860

f: from the synthetic fathers; s: reported information of sons about real fathers
F stands for father
LI stands for Labour Income

TABLE 8
BHPS: FIRST STAGE REGRESSIONS, INSTRUMENTS

	AV3f	AV3s	MIN40f	MIN40s	A55f	A55s	ALLf	ALLs
F No completed ed	0.246	0.393	0.265	0.395	0.226	0.424	0.232	0.402
	(0.43)	(0.49)	(0.44)	(0.49)	(0.42)	(0.49)	(0.42)	(0.49)
F Some ed.	0.318	0.217	0.323	0.232	0.312	0.200	0.317	0.208
	(0.47)	(0.41)	(0.47)	(0.42)	(0.46)	(0.40)	(0.47)	(0.41)
F Further ed.	0.325		0.305		0.328		0.330	
	(0.47)		(0.46)		(0.47)		(0.47)	
F University ed.	0.111	0.085	0.107	0.087	0.134	0.086	0.120	0.091
	(0.31)	(0.28)	(0.31)	(0.28)	(0.34)	(0.28)	(0.33)	(0.29)
F Legislator, manager	0.249	0.220	0.240	0.218	0.263	0.217	0.246	0.219
	(0.43)	(0.41)	(0.43)	(0.41)	(0.44)	(0.41)	(0.43)	(0.41)
F Professional	0.134		0.132		0.151		0.138	
	(0.34)		(0.34)		(0.36)		(0.35)	
F Technician, Ass. prof.	0.102	0.087	0.102	0.099	0.106	0.095	0.108	0.090
	(0.30)	(0.28)	(0.30)	(0.30)	(0.31)	(0.29)	(0.31)	(0.29)
F Clerk	0.063	0.048	0.068	0.047	0.064	0.052	0.069	0.053
	(0.24)	(0.21)	(0.25)	(0.21)	(0.24)	(0.22)	(0.25)	(0.22)
F Service worker, market sales	0.038	0.046	0.040	0.049	0.034	0.044	0.037	0.044
	(0.19)	(0.21)	(0.20)	(0.22)	(0.18)	(0.20)	(0.19)	(0.21)
F Skilled agr. and fishery	0.008	0.010	0.011	0.008	0.010	0.010	0.011	0.009
	(0.09)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)
F Crafts and rtld trade	0.193	0.266	0.196	0.264	0.180	0.258	0.188	0.256
	(0.40)	(0.44)	(0.40)	(0.44)	(0.38)	(0.44)	(0.39)	(0.44)
F Plant and machine operator	0.153	0.159	0.139	0.151	0.140	0.162	0.143	0.161
	(0.36)	(0.37)	(0.35)	(0.36)	(0.35)	(0.37)	(0.35)	(0.37)
F Elementary occupation	0.060	0.074	0.072	0.070	0.053	0.067	0.059	0.073
	(0.24)	(0.26)	(0.26)	(0.26)	(0.22)	(0.25)	(0.24)	(0.26)
F Born in England	0.814	0.726	0.809	0.725	0.818	0.728	0.815	0.736
	(0.39)	(0.45)	(0.39)	(0.45)	(0.39)	(0.44)	(0.39)	(0.44)
F Born in Scotland	0.086	0.102	0.086	0.099	0.077	0.090	0.082	0.091
	(0.28)	(0.30)	(0.28)	(0.30)	(0.27)	(0.29)	(0.27)	(0.29)
F Born in Wales	0.040		0.042		0.041		0.040	
	(0.20)		(0.20)		(0.20)		(0.20)	
F Born in N.I.	0.003	0.012	0.002	0.014	0.004	0.015	0.004	0.013
	(0.05)	(0.11)	(0.05)	(0.12)	(0.07)	(0.12)	(0.06)	(0.11)
F Born in Ireland	0.008	0.021	0.011	0.018	0.009	0.020	0.008	0.019
	(0.09)	(0.14)	(0.10)	(0.13)	(0.10)	(0.14)	(0.09)	(0.14)
F Born abroad	0.049	0.076	0.050	0.080	0.051	0.072	0.052	0.071
	(0.22)	(0.26)	(0.22)	(0.27)	(0.22)	(0.26)	(0.22)	(0.26)
Non-white	0.034	0.052	0.038	0.051	0.031	0.043	0.032	0.045
	(0.18)	(0.22)	(0.19)	(0.22)	(0.17)	(0.20)	(0.18)	(0.21)
F S-E no employees	0.100	0.058	0.101	0.058	0.094	0.048	0.085	0.049
	(0.30)	(0.23)	(0.30)	(0.23)	(0.29)	(0.21)	(0.28)	(0.22)
F S-E with employees	0.133	0.102	0.137	0.105	0.120	0.098	0.128	0.101
	(0.34)	(0.30)	(0.34)	(0.31)	(0.33)	(0.30)	(0.33)	(0.30)
F E, manager	0.260	0.208	0.235	0.213	0.270	0.229	0.254	0.220
	(0.44)	(0.41)	(0.42)	(0.41)	(0.44)	(0.42)	(0.44)	(0.41)
F E, foreman or supervisor	0.111	0.190	0.145	0.188	0.135	0.186	0.130	0.195
	(0.31)	(0.39)	(0.35)	(0.39)	(0.34)	(0.39)	(0.34)	(0.40)
F Employee	0.397		0.382		0.381		0.402	
	(0.49)		(0.49)		(0.49)		(0.49)	
F Age 30-42	0.001	0.535	0.113	0.530	0.029	0.495	0.020	0.515
	(0.03)	(0.50)	(0.32)	(0.50)	(0.17)	(0.50)	(0.14)	(0.50)
F Age 43-48	0.109	0.274	0.346	0.266	0.280	0.302	0.197	0.271
	(0.31)	(0.45)	(0.48)	(0.44)	(0.45)	(0.46)	(0.40)	(0.44)
F Age 49-53	0.510		0.292		0.475		0.335	
	(0.50)		(0.46)		(0.50)		(0.47)	
F Age 54-59	0.380	0.069	0.248	0.073	0.216	0.051	0.448	0.076
	(0.49)	(0.25)	(0.43)	(0.26)	(0.41)	(0.22)	(0.50)	(0.26)
N. of averaged LI obs.	9.267		1.000		8.578		10.577	
	(4.34)		(0.00)		(3.44)		(4.32)	
F ln(LI)	10.203		7.637		7.723		7.739	
	(0.50)		(0.62)		(0.70)		(0.57)	
Father's y of birth		1931.553		1930.806		1928.401		1929.955
		(9.20)		(9.02)		(7.69)		(8.86)
Predicted F ln(LI)		10.424		7.573		7.222		7.190
		(0.78)		(0.29)		(0.34)		(0.38)
N	628	1144	880	834	4614	8363	6378	11976

f: from the synthetic fathers; s: reported information of sons about real fathers
F stands for father
LI stands for Labour Income

TABLE 9
PSID: FIRST STAGE REGRESSIONS, INSTRUMENTS

	AV3f	AV3s	MIN40f	MIN40s	A55f	A55s	ALLf	ALLs
F Manager, Professional	0.476	0.350	0.334	0.377	0.440	0.339	0.426	0.347
	(0.50)	(0.48)	(0.47)	(0.48)	(0.50)	(0.47)	(0.49)	(0.48)
F Sales	0.090	0.073	0.138	0.064	0.091	0.082	0.101	0.077
	(0.29)	(0.26)	(0.35)	(0.25)	(0.29)	(0.27)	(0.30)	(0.27)
F Clerk	0.037	0.040	0.039	0.036	0.060	0.037	0.061	0.038
	(0.19)	(0.19)	(0.19)	(0.19)	(0.24)	(0.19)	(0.24)	(0.19)
F Services and military	0.039	0.048	0.064	0.048	0.050	0.048	0.054	0.046
	(0.19)	(0.21)	(0.25)	(0.21)	(0.22)	(0.21)	(0.23)	(0.21)
F Agr. and fishery	0.032	0.058	0.005	0.057	0.027	0.066	0.032	0.066
	(0.18)	(0.23)	(0.07)	(0.23)	(0.16)	(0.25)	(0.18)	(0.25)
F Crafts and rtld trade	0.163	0.201	0.040	0.186	0.171	0.215	0.164	0.208
	(0.37)	(0.40)	(0.20)	(0.39)	(0.38)	(0.41)	(0.37)	(0.41)
F Plant, machine operator	0.130	0.201	0.374	0.199	0.133	0.189	0.131	0.193
	(0.34)	(0.40)	(0.48)	(0.40)	(0.34)	(0.39)	(0.34)	(0.39)
Self-employed	0.194	0.021	0.065	0.022	0.077	0.021	0.128	0.022
	(0.40)	(0.14)	(0.25)	(0.15)	(0.27)	(0.14)	(0.33)	(0.15)
F Up to 5 yrs ed.	0.021	0.106	0.031	0.100	0.018	0.121	0.018	0.115
	(0.15)	(0.31)	(0.17)	(0.30)	(0.13)	(0.33)	(0.13)	(0.32)
F From 6 to 8 yrs ed.	0.050	0.105	0.126	0.095	0.045	0.112	0.049	0.107
	(0.22)	(0.31)	(0.33)	(0.29)	(0.21)	(0.32)	(0.22)	(0.31)
F High school some train	0.230	0.119	0.272	0.123	0.230	0.107	0.224	0.114
	(0.42)	(0.32)	(0.45)	(0.33)	(0.42)	(0.31)	(0.42)	(0.32)
F High school+some college	0.214	0.181	0.310	0.204	0.228	0.176	0.224	0.180
	(0.41)	(0.39)	(0.46)	(0.40)	(0.42)	(0.38)	(0.42)	(0.38)
F University	0.210	0.123	0.080	0.133	0.241	0.108	0.225	0.114
	(0.41)	(0.33)	(0.27)	(0.34)	(0.43)	(0.31)	(0.42)	(0.32)
Non-white	0.072	0.069	0.201	0.070	0.069	0.063	0.068	0.064
	(0.26)	(0.25)	(0.40)	(0.26)	(0.25)	(0.24)	(0.25)	(0.24)
F New England	0.058	0.046	0.006	0.048	0.024	0.049	0.021	0.048
	(0.23)	(0.21)	(0.08)	(0.21)	(0.15)	(0.22)	(0.14)	(0.21)
F Middle Atlantic	0.171	0.166	0.109	0.175	0.065	0.185	0.070	0.184
	(0.38)	(0.37)	(0.31)	(0.38)	(0.25)	(0.39)	(0.26)	(0.39)
F East North Central	0.153	0.180	0.019	0.180	0.049	0.195	0.052	0.187
	(0.36)	(0.38)	(0.14)	(0.38)	(0.22)	(0.40)	(0.22)	(0.39)
F South Atlantic	0.165	0.176	0.037	0.180	0.046	0.150	0.063	0.165
	(0.37)	(0.38)	(0.19)	(0.38)	(0.21)	(0.36)	(0.24)	(0.37)
F East South Central	0.083	0.083	0.133	0.076	0.023	0.084	0.033	0.081
	(0.28)	(0.28)	(0.34)	(0.26)	(0.15)	(0.28)	(0.18)	(0.27)
F West South Central	0.090	0.063	0.154	0.074	0.028	0.052	0.034	0.058
	(0.29)	(0.24)	(0.36)	(0.26)	(0.17)	(0.22)	(0.18)	(0.23)
F Mountain	0.029	0.045	0.003	0.044	0.007	0.043	0.011	0.043
	(0.17)	(0.21)	(0.06)	(0.21)	(0.08)	(0.20)	(0.11)	(0.20)
F Pacific	0.073	0.063	0.010	0.052	0.015	0.058	0.027	0.058
	(0.26)	(0.24)	(0.10)	(0.22)	(0.12)	(0.23)	(0.16)	(0.23)
F Age 30-42	0.000	0.480	0.021	0.461	0.000	0.473	0.000	0.467
	(0.00)	(0.50)	(0.14)	(0.50)	(0.02)	(0.50)	(0.01)	(0.50)
F Age 43-48	0.067	0.312	0.049	0.327	0.004	0.322	0.002	0.317
	(0.25)	(0.46)	(0.22)	(0.47)	(0.06)	(0.47)	(0.04)	(0.47)
F Age 54-59	0.047	0.067	0.533	0.076	0.412	0.043	0.704	0.068
	(0.21)	(0.25)	(0.50)	(0.26)	(0.49)	(0.20)	(0.46)	(0.25)
F year of birth	1945.925	1935.123	1945.400	1937.290	1947.097	1931.559	1946.196	1932.878
	(2.65)	(8.71)	(3.25)	(7.71)	(1.72)	(7.83)	(2.52)	(8.73)
F ln(LI)	10.896		10.111		10.855		10.795	
	(0.63)		(1.29)		(0.99)		(0.99)	
Predicted F ln(LI)		10.576		10.985		10.598		10.598
		(0.40)		(0.50)		(0.42)		(0.42)
N	817	849	871	945	4459	5148	5943	7488

f: from the synthetic fathers; s: reported information of sons about real fathers
F stands for father
LI stands for Labour Income

APPENDIX B: FIRST STAGE

For all countries, Model 0 indicates the selected model

(a) Germany

GSOEP is a huge dataset and many variables are available as an instrument for the same characteristic. For this reason as an explorative exercise, several models were tested and ranked according the comparability with other countries and the value of the adjusted R^2. Appendix B reports the four models that better satisfy these criteria.

It is worth mentioning that all of them are characterised by a high adjusted R^2. The four models have some common variables: four categorical dummies for age, the year of birth, a dummy to indicate whether the individual is Turk, an indicator for living in West Germany and six dummies variables for different levels of primary and secondary education.

As per educational attainment, the selected variable has six categories according to the different types of primary and secondary school. Additionally, two dummies are created from two separate variables: one for a university degree and one for vocational training. An alternative model was estimated without the indicator of training, on the assumption that the type of secondary school might already be a good predictor of future training. However, as this dummy was statistically significant in two other specifications, the model without it was excluded.

The four short-listed models are also characterised by having the same dummies for occupational position: they correspond to the nine[64] major groups of ISCO88. Two additional variables would have been available: a categorical variable following the Erikson and Goldthorpe Class Categories (IS88), and another occupational indicator with a reduced number of categories.

Indeed variations in the indicators of occupation constitute the main difference between the selected specification (Model 0 in Table 10) and the others (Model 1 to 3). Given that the models based IS88 do not have a better predictive power, ISCO88 was preferred for consistency with BHPS-USS. Nonetheless, IS88 is used to create the indicators for managerial positions and for the type of self-employment, similarly to the dummies included in the model based on British

[64] Originally, the categories are ten. However, the military workers are included in category number five, with protective services.

TABLE 10
GSEOP: MODEL 0, FIRST STAGE REGRESSIONS

	ALL	
Age 30-43	-0.216***	(0.032)
Age 44-48	-0.056***	(0.019)
Age 54-59	0.247***	(0.022)
Turkish national	-0.074**	(0.032)
F Secondary school	-0.008	(0.029)
F Intermediate school	0.126***	(0.037)
F Technical school	0.179***	(0.064)
F Upper secondary school	0.275***	(0.049)
F No degree	0.048*	(0.028)
Legislator, manager	0.007	(0.050)
Professional	-0.113***	(0.040)
Clerk	-0.019	(0.074)
Service and sales workers	-0.085	(0.071)
Skilled agr. and fishery	-0.328***	(0.102)
Crafts and rtld trade	-0.114	(0.070)
Plant and machine operators	-0.084	(0.073)
Elementary occupation	-0.288***	(0.078)
High skilled	0.416***	(0.071)
Low skilled	0.097	(0.064)
SE with employees	0.069	(0.076)
SE no employees	-0.174*	(0.100)
SE farmers	-0.323**	(0.161)
Labourers	-0.059	(0.071)
Voc. training	0.049**	(0.021)
University	0.269***	(0.036)
dadloc1989	0.685***	(0.024)
Father's y of birth	0.040***	(0.003)
Observations	4534	
R^2	0.473	

Clustered standard errors in parenthesis
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

data. The derived variables are included in Model 0 and 2. A series of dummies was also derived from the alternative occupational classification and included in Model 1. Finally Model 3 only includes a dummy for self-employment.

Furthermore, as the adjusted R^2 is similar across the models, the comparability with the other countries is the main criterion to exclude Model 1 and 3. Although both based on the IS88 classification, the dummies in Model 0 are seven (rather than ten in Model 2). This was done in order to increase the size of categories characterised by a too reduced number of observations, which directly affects the statistical significance of the coefficients. This applies to ALL, but it was a problem especially for smaller sample sizes. For this reason, Model 0 was preferred to Model 2. The selected specification fits the data reasonably well. Indeed, the adjusted R^2 is close to 0.5 for all samples (0.5941 for AV3, 0.5290 for MIN40, 0.4559 for A55, 0.4738 for ALL).

(b) Italy

With SHIW, the amount of publicly available information about own parents is limited. The main limitation is the lower the number of categories for each instrument, such as occupational position, state of origin or industry. For example, the other datasets provide detailed information about paternal occupation, through the full ISCO88 scale in BHPS-USS and GSOEP or a similar code of the National Census in PSID. The only available occupational variable in SHIW consists of eight categories (ten if considering "don't know"and "not employed"). Thus, whereas the models in the other datasets mainly differ by alternative categorisations of the same original variable, in SHIW a further reduction in the number of categories would be illogical.

Subsequently, the Italian model is created by selecting those instruments that guarantee the highest possible comparability with the other datasets. The survey construction, highlighted in Section III, considerably reduces the number of observations. Some dummies may be characterised by a very low percentage of observations. The relevant table in Appendix A shows that this is the case with self-employment, where the percentage of self-employed fathers with a positive labour income is very low: specifically, the fathers that are reported by their sons

TABLE 11
GSOEP:MODEL 1, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Age 30-43	-0.298	(0.115)	-0.105	(0.052)	-0.174	(0.040)	-0.222	(0.040)
Age 44-48	-0.177	(0.043)	-0.052	(0.041)	-0.034	(0.024)	-0.061	(0.025)
Age 54-59	0.224	(0.057)	0.267	(0.042)	0.185	(0.029)	0.243	(0.029)
Turkish national	-0.081	(0.055)	0.004	(0.050)	-0.077	(0.042)	-0.067	(0.043)
F Secondary school	0.023	(0.062)	0.060	(0.063)	-0.011	(0.042)	-0.029	(0.047)
F Intermediate school	0.148	(0.089)	0.245	(0.079)	0.150	(0.064)	0.119	(0.070)
F Technical school	0.111	(0.160)	0.407	(0.107)	0.149	(0.110)	0.175	(0.119)
F Upper secondary school	0.182	(0.134)	0.494	(0.090)	0.284	(0.092)	0.261	(0.097)
F No degree	0.065	(0.053)	0.080	(0.052)	0.038	(0.040)	0.054	(0.045)
Legislator, manager	0.219	(0.096)	0.229	(0.075)	0.201	(0.060)	0.241	(0.059)
Professional	0.156	(0.079)	0.010	(0.065)	0.163	(0.060)	0.136	(0.058)
Clerk	-0.080	(0.090)	-0.192	(0.092)	-0.061	(0.066)	-0.081	(0.070)
Service and sales workers	-0.197	(0.107)	-0.135	(0.088)	-0.097	(0.073)	-0.115	(0.072)
Skilled agr. and fishery	-0.336	(0.186)	-0.404	(0.121)	-0.419	(0.120)	-0.449	(0.122)
Crafts and rtld trade	-0.125	(0.073)	-0.074	(0.063)	-0.105	(0.053)	-0.128	(0.051)
Plant and machine operators	-0.126	(0.078)	-0.079	(0.066)	-0.079	(0.057)	-0.094	(0.054)
Elementary occupation	-0.259	(0.084)	-0.269	(0.080)	-0.248	(0.063)	-0.314	(0.061)
Workers	-0.183	(0.097)	-0.118	(0.075)	-0.038	(0.065)	-0.078	(0.062)
Foreman, supervisor	-0.039	(0.107)	0.036	(0.085)	0.133	(0.072)	0.092	(0.068)
White collar	0.032	(0.096)	0.060	(0.074)	0.134	(0.068)	0.102	(0.066)
Civil servant	-0.104	(0.103)	-0.055	(0.084)	-0.028	(0.073)	-0.041	(0.071)
dadloc1989	0.679	(0.056)	0.849	(0.047)	0.698	(0.050)	0.707	(0.046)
University	0.226	(0.085)	0.294	(0.067)	0.222	(0.071)	0.268	(0.066)
Voc. training	-0.021	(0.042)	0.009	(0.042)	0.045	(0.035)	0.038	(0.034)
Father’s y of birth	0.045	(0.007)	0.032	(0.005)	0.026	(0.005)	0.039	(0.005)
Observations	536		1330		3695		4534	
R^2	0.604		0.506		0.445		0.460	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parentheses

TABLE 12
GSOEP:MODEL 2, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Age 30-43	-0.338	(0.112)	-0.098	(0.051)	-0.164	(0.040)	-0.214	(0.040)
Age 44-48	-0.169	(0.043)	-0.048	(0.039)	-0.027	(0.024)	-0.057	(0.024)
Age 54-59	0.232	(0.055)	0.266	(0.042)	0.189	(0.029)	0.248	(0.028)
Turkish national	-0.072	(0.051)	-0.002	(0.050)	-0.080	(0.041)	-0.066	(0.041)
F Secondary school	0.069	(0.058)	0.058	(0.060)	0.011	(0.040)	-0.008	(0.045)
F Intermediate school	0.168	(0.088)	0.229	(0.075)	0.147	(0.064)	0.112	(0.068)
F Technical school	0.061	(0.151)	0.387	(0.103)	0.156	(0.104)	0.175	(0.110)
F Upper secondary school	0.204	(0.125)	0.475	(0.084)	0.295	(0.086)	0.265	(0.092)
F No degree	0.059	(0.052)	0.067	(0.050)	0.035	(0.039)	0.048	(0.043)
Legislator, manager	-0.151	(0.108)	0.021	(0.092)	-0.006	(0.077)	0.017	(0.075)
Professional	-0.146	(0.082)	-0.224	(0.075)	-0.062	(0.070)	-0.104	(0.067)
Clerk	-0.037	(0.167)	-0.132	(0.185)	-0.020	(0.120)	0.049	(0.135)
Service and sales workers	-0.326	(0.198)	0.051	(0.145)	-0.135	(0.117)	-0.047	(0.114)
Skilled agr. and fishery	-0.114	(0.258)	0.039	(0.317)	-0.184	(0.160)	-0.150	(0.161)
Crafts and rtld trade	-0.341	(0.188)	0.097	(0.128)	-0.179	(0.107)	-0.108	(0.101)
Plant and machine operators	-0.310	(0.193)	0.148	(0.142)	-0.115	(0.113)	-0.029	(0.107)
Elementary occupation	-0.399	(0.198)	0.012	(0.156)	-0.235	(0.117)	-0.201	(0.112)
High skilled	0.357	(0.128)	0.280	(0.119)	0.278	(0.104)	0.362	(0.108)
Low skilled	-0.074	(0.120)	-0.092	(0.101)	-0.015	(0.092)	0.047	(0.098)
Routine service	-0.187	(0.147)	-0.195	(0.137)	-0.097	(0.098)	-0.153	(0.101)
SE with employees	0.126	(0.178)	-0.188	(0.133)	-0.000	(0.121)	0.007	(0.121)
SE no employees	-0.214	(0.209)	-0.326	(0.181)	-0.302	(0.196)	-0.237	(0.182)
SE farmers	-0.698	(0.428)	-0.777	(0.427)	-0.625	(0.246)	-0.523	(0.233)
Skilled labourers	0.018	(0.189)	-0.394	(0.139)	-0.045	(0.118)	-0.105	(0.118)
semi- unskilled lab.	-0.056	(0.190)	-0.486	(0.149)	-0.121	(0.123)	-0.189	(0.122)
Farm labourers	-0.336	(0.269)	-0.592	(0.261)	-0.273	(0.169)	-0.321	(0.172)
University	0.258	(0.083)	0.280	(0.064)	0.232	(0.069)	0.270	(0.065)
Voc. training	-0.022	(0.041)	0.029	(0.040)	0.057	(0.032)	0.049	(0.032)
dadloc1989	0.673	(0.053)	0.833	(0.045)	0.670	(0.047)	0.679	(0.043)
Father's y of birth	0.046	(0.007)	0.031	(0.005)	0.027	(0.005)	0.040	(0.004)
Observations	536		1330		3695		4534	
R^2	0.631		0.530		0.457		0.474	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parentheses

TABLE 13
GSOEP:MODEL 3, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Age 30-43	-0.309	(0.108)	-0.089	(0.052)	-0.162	(0.040)	-0.211	(0.040)
Age 44-48	-0.168	(0.044)	-0.043	(0.041)	-0.026	(0.024)	-0.055	(0.025)
Age 54-59	0.219	(0.057)	0.273	(0.042)	0.185	(0.030)	0.243	(0.029)
Turkish national	-0.088	(0.056)	-0.008	(0.051)	-0.084	(0.041)	-0.076	(0.043)
F Secondary school	0.057	(0.059)	0.087	(0.063)	0.019	(0.041)	0.003	(0.046)
F Intermediate school	0.189	(0.085)	0.279	(0.078)	0.183	(0.063)	0.155	(0.068)
F Technical school	0.150	(0.162)	0.436	(0.107)	0.189	(0.114)	0.208	(0.121)
F Upper secondary school	0.201	(0.129)	0.514	(0.089)	0.306	(0.089)	0.283	(0.095)
F No degree	0.073	(0.053)	0.079	(0.052)	0.041	(0.040)	0.054	(0.045)
Legislator, manager	0.230	(0.098)	0.222	(0.075)	0.201	(0.059)	0.239	(0.059)
Professional	0.130	(0.078)	-0.014	(0.064)	0.131	(0.060)	0.104	(0.059)
Clerk	-0.090	(0.091)	-0.198	(0.093)	-0.076	(0.066)	-0.097	(0.071)
Service and sales workers	-0.264	(0.099)	-0.181	(0.086)	-0.165	(0.072)	-0.180	(0.070)
Skilled agr. and fishery	-0.439	(0.170)	-0.482	(0.116)	-0.481	(0.114)	-0.521	(0.117)
Crafts and rtld trade	-0.246	(0.063)	-0.173	(0.050)	-0.195	(0.045)	-0.225	(0.044)
Plant and machine operators	-0.260	(0.073)	-0.181	(0.053)	-0.179	(0.053)	-0.204	(0.051)
Elementary occupation	-0.382	(0.082)	-0.358	(0.076)	-0.332	(0.058)	-0.409	(0.057)
Father S-E	0.072	(0.087)	0.005	(0.068)	-0.062	(0.062)	-0.030	(0.060)
dadloc1989	0.673	(0.052)	0.844	(0.046)	0.686	(0.047)	0.697	(0.044)
University	0.257	(0.087)	0.308	(0.066)	0.235	(0.072)	0.279	(0.068)
Voc. training	-0.013	(0.042)	0.019	(0.041)	0.053	(0.035)	0.044	(0.033)
Father's y of birth	0.045	(0.007)	0.032	(0.005)	0.026	(0.005)	0.039	(0.005)
Observations	536		1330		3695		4534	
R^2	0.594		0.501		0.435		0.453	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parentheses

as self-employed are 2 to 3 percent in all the samples. As a further example, synthetic fathers in AV3 are older then those in other samples. Thus the dummy indicating individuals younger than forty-four has no observations.

Table 14 illustrates the estimation results.

TABLE 14
SHIW: FIRST STAGE REGRESSIONS

	ALL	
No education	-0.514***	(0.111)
Primary ed.	-0.195***	(0.047)
Upper sec. ed.	0.156***	(0.048)
University ed.	0.464***	(0.074)
Production worker	-0.085*	(0.046)
Junior manager, officer	0.197***	(0.055)
Manager	0.436***	(0.087)
Self employed	-0.048	(0.051)
Industry	0.067*	(0.039)
Agriculture	-0.264***	(0.095)
Northern Italy	0.186***	(0.040)
Central Italy	0.149***	(0.046)
Age 30-43	-0.163	(0.108)
Age 44-48	-0.176***	(0.040)
Age 54-59	0.188***	(0.033)
Year of birth	0.035***	(0.004)
Observations	1516	
R^2	0.342	

Clustered standard errors in parenthesis
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Paternal earnings are regressed on five dummy variables for education, five indicators for job position, three indicators of occupational sector. Considering the limited amount of self-employed workers, the model merges the distinction between self-employed member of a profession, free-lance and entrepreneur into an indicator for self-employment, for a total of twenty-three variables. It is worth mentioning that the variables indicating the area where the individual lives (North, Centre, South) are not available in the second stage. Thus, following Piraino (2007); Mocetti (2007), they are approximated by the area where the son was born. The model is completed by four age categories and the year of birth.

A major difference with the other countries is that it is not possible to identify the race of the individuals. A question on ethnicity is not asked, and although the questionnaire include details about the country of birth, this information is not available for researchers external to the Bank of Italy. An attempt was done to include a dummy indicating a foreign nationality, by identifying the cases of which the region of birth is missing.The indicator was not statistically significant in any of the samples and the identification of the variable in the second stage would require that the father is born in the same place of the son (because no direct information on paternal place of birth is available). Thus, the variable was dropped.

(c) United Kingdom

BHPS-USS is a large dataset and several variables are available to indicate the same socioeconomic characteristic. In this sense, it shares many similarities with GSOEP. Particularly, the two datasets have a similar component, as they were part of the European Community Household Panel (ECHP) and of CNEF (administrated by Cornell University, the project provides cross-country equivalent variables from a series of major datasets including BHPS, GSOEP and PSID)[65]. Thus in this case as well several models were estimated and excluded on the basis of the same criteria followed for the case of German data.

Table 15 shows the results of the first stage regressions for the selected model (Model 0). As the table indicates, the log of labour income is regressed on five categorical variables for educational attainment, nine[66] variables on occupation (derived from ISCO88); five dummies providing additional information about the job status: if employed, whether manager, supervisor or simple employee; if self-employed, whether if with employers or not. Anagraphical characteristics include six indicator variables for the state of birth (derived by aggregating county and district of birth), a dummy variable for non white individuals[67], year of birth and four age categories.

The short-listed models, reported below, do not differ extensively. The main difference with Model 0 is the number of categories for each variable. In Model 1 the occupational categories are reduced from nine to six and self-employment is not accounted for. Model 2 abolishes

[65]CNEF however was not used for the current research, as it is not available for Italy.
[66]Like with GSOEP, soldiers and members of the army are included in category number five.
[67]An attempt was made to isolate Asians, but the subsample is too limited and the results were not statistically significant.

TABLE 15

BHPS-USS: MODEL 0, FIRST STAGE REGRESSIONS

	ALL	
No completed education	-0.216***	(0.046)
Some ed.	-0.123***	(0.041)
University ed.	0.076	(0.065)
Legislator, manager	-0.051	(0.058)
Technician, Ass. prof.	-0.074	(0.056)
Clerk	-0.287***	(0.067)
Service worker, market sales	-0.361***	(0.111)
Skilled agr. and fishery	-0.688***	(0.128)
Crafts and rtld trade	-0.154***	(0.055)
Plant and machine operator	-0.198***	(0.058)
Elementary occupation	-0.331***	(0.073)
Born in England	-0.037	(0.064)
Born in Scotland	-0.145*	(0.085)
Born in N.I.	0.057	(0.100)
Born in Ireland	0.168	(0.114)
Born abroad	0.030	(0.155)
Non white	-0.175	(0.124)
S-E no employees	0.180*	(0.096)
S-E with employees	-0.348***	(0.070)
E, manager	0.363***	(0.043)
E, foreman or supervisor	0.193***	(0.034)
Age 30-43	-0.136**	(0.065)
Age 44-48	-0.120***	(0.024)
Age 54-59	0.160***	(0.023)
Year of birth	0.039***	(0.004)
Observations	4989	
R^2	0.311	

Clustered standard errors in parenthesis

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

the distinction between self-employed with and without employees. Model 3 excludes any additional information about self-employment and managerial position.

The selected specification is the one that better conciliates the two following facts. On the one hand, a reduced categorisation would reduce variation; on the other hand, too many categories might not assign enough observations for each group in the second stage. Moreover, Model 0 has the highest Adjusted R^2 (0.3077 for ALL, 0.2924 for ALL55, 0.4141 for AV3, 0.2398 for MIN40). Finally, with Model 0 the hypothesis of no omitted variables is not rejected at ten percent confidence level for AV3 and MIN40.

TABLE 16
BHPS-USS: MODEL 1, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
No completed education	-0.211	(0.057)	-0.260	(0.054)	-0.235	(0.055)	-0.206	(0.048)
Some ed.	-0.149	(0.049)	-0.194	(0.052)	-0.132	(0.046)	-0.118	(0.042)
University ed.	0.123	(0.065)	0.099	(0.081)	0.102	(0.070)	0.094	(0.068)
Legislator, manager	0.049	(0.066)	0.058	(0.080)	0.030	(0.063)	-0.011	(0.059)
Technician, Ass. prof.	-0.173	(0.076)	-0.047	(0.083)	-0.040	(0.061)	-0.097	(0.057)
Clerk	-0.257	(0.087)	-0.409	(0.093)	-0.264	(0.072)	-0.315	(0.068)
Skilled agr. and fishery	-0.819	(0.186)	-0.742	(0.183)	-0.687	(0.157)	-0.747	(0.134)
Crafts and rtld trade	-0.157	(0.069)	-0.245	(0.080)	-0.192	(0.060)	-0.197	(0.057)
Plant and machine operator	-0.196	(0.073)	-0.330	(0.089)	-0.219	(0.062)	-0.239	(0.059)
Elementary occupation	-0.320	(0.097)	-0.414	(0.104)	-0.350	(0.077)	-0.363	(0.073)
Born in England	-0.073	(0.084)	-0.106	(0.057)	-0.113	(0.077)	-0.046	(0.069)
Born in Scotland	-0.108	(0.105)	-0.174	(0.087)	-0.257	(0.099)	-0.156	(0.089)
Born in N.I.	-0.323	(0.348)	-0.475	(0.139)	-0.167	(0.136)	0.013	(0.139)
Born in Ireland	-0.103	(0.190)	0.060	(0.189)	0.012	(0.094)	0.135	(0.112)
Born abroad	-0.018	(0.164)	-0.241	(0.165)	-0.095	(0.184)	0.020	(0.154)
Non white	-0.146	(0.120)	-0.143	(0.159)	-0.192	(0.141)	-0.174	(0.121)
Self-employed	-0.186	(0.113)	-0.042	(0.274)	0.258	(0.206)	0.320	(0.166)
E, manager	0.125	(0.136)	0.320	(0.278)	0.683	(0.218)	0.777	(0.182)
E, foreman or supervisor	-0.020	(0.144)	0.327	(0.277)	0.586	(0.215)	0.652	(0.180)
Employee	-0.169	(0.141)	0.119	(0.278)	0.420	(0.221)	0.465	(0.183)
o.Age 30-43	0.000	(.)	-0.092	(0.100)	-0.101	(0.068)	-0.125	(0.067)
Age 44-48	-0.188	(0.067)	-0.122	(0.060)	-0.105	(0.023)	-0.123	(0.024)
Age 54-59	0.029	(0.050)	0.199	(0.069)	0.083	(0.025)	0.161	(0.023)
Service worker, market sales	-0.293	(0.132)	-0.095	(0.091)	-0.293	(0.125)	-0.377	(0.110)
Year of birth	0.034	(0.006)	0.030	(0.009)	0.034	(0.005)	0.040	(0.004)
Observations	496		735		3558		4989	
R^2	0.447		0.287		0.274		0.287	

AV3, MIN40: Robust standard errors; A55: Clustered standard errors in parethesis

TABLE 17

BHPS-USS: MODEL 2, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
No completed education	-0.271	(0.057)	-0.344	(0.059)	-0.288	(0.055)	-0.259	(0.049)
Some ed.	-0.165	(0.050)	-0.219	(0.053)	-0.150	(0.047)	-0.136	(0.043)
University ed.	0.143	(0.073)	0.133	(0.079)	0.134	(0.076)	0.127	(0.076)
iscodadit1	0.036	(0.067)	0.124	(0.087)	0.036	(0.061)	0.059	(0.055)
iscodadit3	0.097	(0.090)	0.290	(0.110)	0.212	(0.066)	0.225	(0.061)
iscodadit4	0.289	(0.092)	0.340	(0.115)	0.280	(0.070)	0.279	(0.063)
iscodadit5	0.152	(0.090)	0.230	(0.110)	0.153	(0.074)	0.208	(0.069)
iscodadit6	-0.098	(0.136)	0.235	(0.306)	0.492	(0.235)	0.571	(0.194)
Born in England	-0.092	(0.081)	-0.100	(0.105)	-0.120	(0.079)	-0.060	(0.070)
Born in Scotland	-0.146	(0.110)	-0.180	(0.123)	-0.281	(0.105)	-0.185	(0.093)
Born in N.I.	-0.282	(0.296)	-0.411	(0.338)	-0.149	(0.116)	0.023	(0.123)
Born in Ireland	-0.110	(0.193)	-0.046	(0.276)	0.014	(0.100)	0.098	(0.103)
Born abroad	-0.050	(0.148)	-0.190	(0.178)	-0.109	(0.182)	-0.003	(0.150)
Non white	-0.123	(0.110)	-0.166	(0.165)	-0.182	(0.142)	-0.162	(0.119)
E, manager	0.069	(0.143)	0.406	(0.297)	0.759	(0.242)	0.850	(0.200)
E, foreman or supervisor	-0.084	(0.145)	0.332	(0.297)	0.644	(0.239)	0.708	(0.199)
Employee	-0.222	(0.143)	0.132	(0.293)	0.491	(0.247)	0.526	(0.203)
o.Age 30-43	0.000	(.)	-0.101	(0.100)	-0.116	(0.069)	-0.143	(0.069)
Age 44-48	-0.217	(0.072)	-0.115	(0.067)	-0.100	(0.024)	-0.120	(0.025)
Age 54-59	0.009	(0.052)	0.180	(0.075)	0.077	(0.025)	0.153	(0.023)
Year of birth	0.034	(0.006)	0.029	(0.009)	0.033	(0.005)	0.040	(0.004)
Observations	491		735		3558		4989	
R^2	0.411		0.246		0.255		0.266	

AV3, MIN40: Robust standard errors; A55: Clustered standard errors in parethesis

TABLE 18

BHPS-USS: MODEL 3, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
No completed education	-0.258	(0.055)	-0.305	(0.055)	-0.291	(0.056)	-0.260	(0.049)
Some ed.	-0.174	(0.052)	-0.210	(0.053)	-0.167	(0.048)	-0.152	(0.044)
University ed.	0.158	(0.066)	0.138	(0.077)	0.130	(0.068)	0.126	(0.065)
Legislator, manager	0.140	(0.062)	0.098	(0.075)	0.093	(0.056)	0.079	(0.054)
Technician, Ass. prof.	-0.189	(0.082)	-0.026	(0.084)	-0.046	(0.060)	-0.096	(0.058)
Clerk	-0.327	(0.084)	-0.361	(0.091)	-0.274	(0.072)	-0.335	(0.068)
Skilled agr. and fishery	-0.903	(0.196)	-0.748	(0.179)	-0.763	(0.206)	-0.827	(0.164)
Crafts and rtld trade	-0.223	(0.064)	-0.285	(0.074)	-0.253	(0.057)	-0.262	(0.055)
Plant and machine operator	-0.267	(0.069)	-0.330	(0.084)	-0.267	(0.061)	-0.298	(0.058)
Elementary occupation	-0.368	(0.092)	-0.406	(0.101)	-0.383	(0.076)	-0.404	(0.070)
Born in England	-0.069	(0.088)	-0.096	(0.060)	-0.096	(0.079)	-0.025	(0.070)
Born in Scotland	-0.076	(0.108)	-0.143	(0.089)	-0.226	(0.101)	-0.113	(0.090)
Born in N.I.	-0.444	(0.410)	-0.598	(0.174)	-0.226	(0.185)	-0.037	(0.187)
Born in Ireland	-0.045	(0.155)	0.010	(0.223)	-0.067	(0.103)	0.115	(0.108)
Born abroad	-0.016	(0.183)	-0.257	(0.182)	-0.054	(0.191)	0.068	(0.161)
Non white	-0.198	(0.140)	-0.140	(0.172)	-0.233	(0.149)	-0.216	(0.127)
o.Age 30-43	0.000	(.)	-0.031	(0.100)	-0.098	(0.071)	-0.118	(0.072)
Age 44-48	-0.213	(0.070)	-0.081	(0.059)	-0.096	(0.023)	-0.111	(0.024)
Age 54-59	0.023	(0.053)	0.171	(0.068)	0.080	(0.025)	0.156	(0.024)
Service worker, market sales	-0.295	(0.137)	-0.048	(0.092)	-0.295	(0.134)	-0.389	(0.116)
Year of birth	0.035	(0.006)	0.025	(0.009)	0.035	(0.005)	0.041	(0.004)
Observations	497		735		3558		4989	
R^2	0.402		0.243		0.224		0.233	

AV3, MIN40: Robust standard errors; A55: Clustered standard errors in parethesis *(d) United States*

Although several variables are available with PSID, and alternative models can be constructed, the six competing models are characterised by a low Adjusted R^2, for ALL ranging from 0.21 to 0.24. Differences among them are mainly based on different ways of simplifying the original variables. The only exception is the derived occupational status, which remains invariant across specifications. The occupational categories rely on the Census classification but they are very similar to those resulting from ISCO88. The final variable is composed by eight categories[68]. In order to obtain it, however, the National Census coding had to be rendered comparable across years. In fact, the variable is coded following either the classification of the 1970 Census of Population or that of the Census in 2000 (the latter from the 2003 wave). Thus, firstly I re-coded the two instruments in terms of the 1990 Census with the occupation cross-walk provided by the United States Census Bureau. Afterwards, the occupational category was reduced to nine main groups, and then further to eight, similarly to what has been done for GSOEP and BHPS-US, and as explained above. Considering the low Adjusted R^2, the decisive criterion used for the model selection for the United States are the results of the specification tests. Both the RESET and the link tests were considered. The results of both tests do not provide evidence for misspecification for three competing models (Models 0, 1 and 4).

Among these three, the selected model, reported in Table 19, is the one characterised by a higher adjusted R^2 (0.22 versus 0.21). This is also the model the the highest number of regressors. Particularly, it contains additional information about the industry, whereas the others do not.

As indicated in the table, it contains four categories for the industry the individual works in. Alternative models either do not contain that information at all (Models 1, 4) or contain a more detailed classification of sectors into nine groups (Models 2, 3, 5). The procedure to obtain these categories is similar to the one employed to simplify the occupational position.

[68]The ninth category is military, but the number of synthetic fathers working in military is zero. Thus, as already done for BHPS and GSOEP, the military workers are included in category number five, with protective services and police corps. The alternative would be to exclude individuals working in the army, but that would decrease the number of observations.

TABLE 19
PSID:: MODEL 0, FIRST STAGE REGRESSIONS

	ALL	
Manager, Professional	0.007	(0.088)
Sales	-0.234**	(0.106)
Clerk	-0.161*	(0.098)
Services and military	-0.562***	(0.108)
Agriculture and fishery	-0.671***	(0.141)
Crafts and rtld trade	-0.225**	(0.091)
Plant, machine operator	-0.363***	(0.094)
Self-employed	-0.288***	(0.057)
Age 30-42	-0.108**	(0.045)
Age 43-48	-0.056**	(0.026)
Age 54-59	-0.070**	(0.029)
Up to 5 yrs education	-0.291*	(0.155)
From 6 to 8 yrs education	-0.118	(0.093)
High school and some train	0.226***	(0.057)
High school+some college	0.385***	(0.063)
University degree	0.510***	(0.080)
Postgraduate	0.704***	(0.083)
Non-white	-0.151***	(0.040)
New England	0.298***	(0.079)
Middle Atlantic	-0.008	(0.056)
East North Central	0.036	(0.050)
South Atlantic	0.015	(0.042)
East South Central	0.015	(0.049)
West South Central	-0.151***	(0.054)
Mountain	-0.239***	(0.092)
Pacific	0.050	(0.059)
Agriculture and Mining	0.080	(0.092)
Construction and transport	-0.247***	(0.064)
Services and PA	-0.053	(0.037)
Year of birth	0.003	(0.005)
Observations	7918	
R^2	0.221	

Clustered standard errors in parethesis
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

A similar method was used as well in order to organize the state where the father grew up into macro-regions. Indeed, whereas until 1993 the states are coded from 1 to 51 (GSA code), for the following years FIPS coding is used (from 1 to 56). Thus, a sort of cross-walk was constructed in order to add for each observation the GSA and FIPS codes, the region as specified in PSID, and an additional classification in nine sub-regions created by the National Census Bureau. Given the low number of individuals living in Alaska and Hawaii, the number of GSA regions was reduced from five to four, and these two states included in West.

The selected model also contains seven dummies for educational attainment. This is common to all the other specifications except one, Model 1, where individuals with first and postgraduate degrees are merged into a unique indicator.

As per ethnicity, two choices were possible: a dummy variable indicating whether the race of the individual is not white; a three-category regressor specifying whether the individual is white, black or other. The models containing three indicators for ethnicity have a slightly higher adjusted R^2. Nonetheless none of these was selected for two main reasons: firstly, the improvement of the goodness of fit is only limited and it is counterbalanced by the fact that these models perform poorly when testing for misspecification; secondly, in the second stage the individuals are not asked about the race of their parents. We assume that the father has the same race as the son. Having a dummy variable rather than three categories might reduce the error. Although an error might arise, race is too important to be excluded when predicting labour earnings in America. Therefore, I chose to include the variable anyway.

Finally, and like for all other datasets, the model includes the year of birth, four age categories and a dummy variable for self-employment.

TABLE 20
PSID: MODEL 1, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Manager, Professional	0.075	(0.103)	0.092	(0.091)	-0.005	(0.083)	0.002	(0.087)
Sales	-0.078	(0.131)	-0.193	(0.124)	-0.186	(0.106)	-0.207	(0.105)
Clerk	0.041	(0.114)	-0.019	(0.105)	-0.102	(0.093)	-0.160	(0.097)
Services and military	-0.466	(0.121)	-0.426	(0.119)	-0.524	(0.106)	-0.563	(0.107)
Agriculture and fishery	-0.289	(0.153)	-0.670	(0.176)	-0.572	(0.133)	-0.551	(0.129)
Crafts and rtld trade	-0.073	(0.108)	-0.254	(0.110)	-0.257	(0.087)	-0.254	(0.092)
Plant, machine operator	-0.162	(0.110)	-0.339	(0.108)	-0.330	(0.090)	-0.356	(0.094)
Self-employed	-0.116	(0.065)	-0.369	(0.080)	-0.304	(0.062)	-0.325	(0.056)
Age 30-42	-0.205	(0.164)	0.459	(0.130)	-0.092	(0.046)	-0.102	(0.046)
Age 43-48	-0.084	(0.045)	0.330	(0.091)	-0.050	(0.027)	-0.055	(0.027)
Age 54-59	-0.060	(0.073)	-0.308	(0.113)	-0.071	(0.032)	-0.075	(0.029)
Up to 5 yrs ed.	-0.355	(0.078)	-0.423	(0.090)	-0.413	(0.094)	-0.375	(0.082)
From 6 to 8 yrs ed.	-0.231	(0.059)	-0.294	(0.072)	-0.281	(0.064)	-0.240	(0.058)
High school some train	0.191	(0.050)	0.166	(0.063)	0.150	(0.050)	0.160	(0.047)
High school+some college	0.338	(0.067)	0.233	(0.081)	0.269	(0.066)	0.286	(0.067)
University	0.526	(0.073)	0.420	(0.102)	0.505	(0.069)	0.482	(0.068)
Non-white	-0.118	(0.044)	-0.138	(0.062)	-0.159	(0.041)	-0.152	(0.040)
New England	0.198	(0.108)	0.378	(0.117)	0.271	(0.088)	0.299	(0.079)
Middle Atlantic	0.072	(0.067)	0.157	(0.091)	-0.040	(0.062)	-0.008	(0.056)
East North Central	0.064	(0.070)	0.177	(0.081)	0.053	(0.054)	0.043	(0.050)
South Atlantic	0.044	(0.062)	0.200	(0.081)	0.008	(0.045)	0.011	(0.043)
East South Central	-0.018	(0.073)	0.176	(0.090)	0.042	(0.054)	0.028	(0.049)
West South Central	-0.062	(0.074)	0.003	(0.092)	-0.171	(0.058)	-0.153	(0.055)
Mountain	-0.219	(0.115)	0.031	(0.117)	-0.248	(0.093)	-0.237	(0.093)
Pacific	0.092	(0.083)	0.114	(0.100)	0.043	(0.062)	0.044	(0.059)
Year of birth	0.001	(0.007)	-0.048	(0.013)	0.001	(0.006)	0.002	(0.005)
Observations	954		1315		6324		7918	
R^2	0.350		0.272		0.225		0.215	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parenthesis

TABLE 21
PSID: MODEL 2, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Manager, Professional	0.114	(0.101)	0.201	(0.094)	0.053	(0.040)	0.062	(0.039)
Sales	-0.135	(0.137)	-0.148	(0.132)	-0.190	(0.064)	-0.204	(0.059)
Clerk	-0.020	(0.111)	0.006	(0.108)	-0.119	(0.047)	-0.182	(0.047)
Services and military	-0.397	(0.121)	-0.342	(0.122)	-0.445	(0.055)	-0.475	(0.052)
Agriculture and fishery	-0.462	(0.249)	-0.347	(0.218)	-0.551	(0.111)	-0.529	(0.098)
Crafts and rtld trade	-0.074	(0.104)	-0.195	(0.111)	-0.237	(0.043)	-0.239	(0.041)
Plant, machine operator	-0.236	(0.108)	-0.345	(0.107)	-0.370	(0.046)	-0.407	(0.045)
Self-employed	-0.018	(0.066)	-0.253	(0.083)	-0.218	(0.041)	-0.237	(0.036)
Age 30-42	-0.247	(0.125)	0.504	(0.126)	-0.098	(0.039)	-0.099	(0.037)
Age 43-48	-0.069	(0.044)	0.348	(0.087)	-0.054	(0.024)	-0.054	(0.023)
Age 54-59	-0.024	(0.071)	-0.324	(0.112)	-0.059	(0.035)	-0.061	(0.025)
Up to 5 yrs education	-0.081	(0.152)	-0.153	(0.183)	-0.291	(0.173)	-0.229	(0.143)
From 6 to 8 yrs education	-0.152	(0.086)	-0.116	(0.107)	-0.109	(0.059)	-0.122	(0.050)
High school and some train	0.209	(0.056)	0.270	(0.069)	0.272	(0.035)	0.224	(0.031)
High school+some college	0.396	(0.067)	0.416	(0.088)	0.406	(0.040)	0.368	(0.036)
University degree	0.512	(0.080)	0.457	(0.106)	0.517	(0.047)	0.489	(0.043)
Postgraduate	0.775	(0.092)	0.688	(0.132)	0.803	(0.051)	0.735	(0.047)
Non-white	-0.100	(0.043)	-0.096	(0.057)	-0.139	(0.026)	-0.130	(0.023)
Grown in NE	0.096	(0.052)	0.141	(0.066)	0.090	(0.030)	0.101	(0.028)
Grown in NC	0.032	(0.047)	0.044	(0.054)	0.049	(0.025)	0.052	(0.023)
Grown in West	-0.010	(0.061)	-0.016	(0.069)	-0.006	(0.033)	-0.009	(0.031)
Agriculture	-0.052	(0.204)	-0.648	(0.180)	-0.205	(0.099)	-0.227	(0.086)
Mining	0.132	(0.105)	0.416	(0.136)	0.221	(0.074)	0.159	(0.074)
Construction	-0.368	(0.071)	-0.438	(0.083)	-0.347	(0.039)	-0.367	(0.036)
Trade	-0.289	(0.070)	-0.317	(0.078)	-0.297	(0.038)	-0.326	(0.034)
Transport and commerce	-0.001	(0.054)	-0.053	(0.081)	-0.015	(0.030)	-0.024	(0.027)
Finance and insurance	0.118	(0.084)	-0.112	(0.122)	-0.004	(0.052)	-0.026	(0.047)
Services	-0.344	(0.055)	-0.343	(0.062)	-0.315	(0.028)	-0.358	(0.028)
Public admin.	-0.122	(0.066)	-0.133	(0.083)	-0.111	(0.036)	-0.124	(0.032)
Year of birth	0.002	(0.007)	-0.052	(0.013)	0.002	(0.004)	0.002	(0.003)
Observations	954		1315		6324		7918	
R^2	0.397		0.298		0.245		0.239	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parenthesis

TABLE 22
PSID: MODEL 3, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Manager, Professional	0.113	(0.101)	0.201	(0.094)	0.053	(0.040)	0.062	(0.084)
Sales	-0.135	(0.137)	-0.148	(0.132)	-0.190	(0.064)	-0.203	(0.106)
Clerk	-0.021	(0.112)	0.003	(0.108)	-0.118	(0.047)	-0.182	(0.095)
Services and military	-0.398	(0.121)	-0.347	(0.121)	-0.444	(0.055)	-0.475	(0.101)
Agriculture and fishery	-0.461	(0.249)	-0.346	(0.219)	-0.550	(0.111)	-0.527	(0.151)
Crafts and rtld trade	-0.074	(0.104)	-0.193	(0.110)	-0.236	(0.043)	-0.238	(0.089)
Plant, machine operator	-0.236	(0.108)	-0.342	(0.107)	-0.369	(0.046)	-0.406	(0.093)
Self-employed	-0.018	(0.066)	-0.254	(0.083)	-0.219	(0.041)	-0.238	(0.057)
Age 30-42	-0.249	(0.125)	0.501	(0.126)	-0.100	(0.039)	-0.101	(0.040)
Age 43-48	-0.071	(0.044)	0.345	(0.087)	-0.055	(0.024)	-0.055	(0.025)
Age 54-59	-0.023	(0.071)	-0.322	(0.111)	-0.058	(0.035)	-0.060	(0.028)
Up to 5 yrs education	-0.088	(0.155)	-0.172	(0.182)	-0.305	(0.174)	-0.244	(0.142)
From 6 to 8 yrs education	-0.156	(0.085)	-0.124	(0.106)	-0.119	(0.059)	-0.129	(0.090)
High school and some train	0.208	(0.056)	0.268	(0.069)	0.272	(0.035)	0.223	(0.056)
High school+some college	0.396	(0.067)	0.414	(0.087)	0.406	(0.040)	0.368	(0.062)
University degree	0.512	(0.080)	0.457	(0.106)	0.517	(0.047)	0.488	(0.078)
Postgraduate	0.774	(0.092)	0.686	(0.132)	0.802	(0.051)	0.734	(0.082)
White	0.106	(0.045)	0.117	(0.062)	0.151	(0.026)	0.141	(0.043)
Other	0.038	(0.093)	0.141	(0.117)	0.102	(0.068)	0.101	(0.104)
Grown in NE	0.095	(0.052)	0.137	(0.066)	0.087	(0.030)	0.098	(0.053)
Grown in NC	0.030	(0.047)	0.038	(0.054)	0.046	(0.025)	0.049	(0.045)
Grown in West	-0.017	(0.063)	-0.038	(0.071)	-0.018	(0.033)	-0.020	(0.062)
Agriculture	-0.054	(0.203)	-0.648	(0.181)	-0.205	(0.099)	-0.228	(0.124)
Mining	0.128	(0.108)	0.419	(0.137)	0.215	(0.075)	0.153	(0.111)
Construction	-0.370	(0.070)	-0.441	(0.083)	-0.350	(0.039)	-0.370	(0.064)
Trade	-0.290	(0.070)	-0.319	(0.078)	-0.298	(0.038)	-0.328	(0.055)
Transport and commerce	-0.003	(0.055)	-0.057	(0.082)	-0.016	(0.030)	-0.026	(0.046)
Finance and insurance	0.117	(0.084)	-0.111	(0.123)	-0.004	(0.052)	-0.027	(0.075)
Services	-0.344	(0.055)	-0.344	(0.062)	-0.316	(0.028)	-0.359	(0.047)
Public admin.	-0.122	(0.066)	-0.132	(0.083)	-0.111	(0.036)	-0.124	(0.054)
Year of birth	0.002	(0.007)	-0.052	(0.013)	0.002	(0.004)	0.002	(0.004)
Observations	954		1315		6324		7918	
R^2	0.397		0.299		0.245		0.240	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parenthesis

TABLE 23
PSID: MODEL 4, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Manager, Professional	0.053	(0.105)	0.087	(0.091)	-0.014	(0.086)	-0.009	(0.091)
Sales	-0.093	(0.131)	-0.204	(0.124)	-0.188	(0.109)	-0.214	(0.109)
Clerk	-0.004	(0.113)	-0.021	(0.105)	-0.114	(0.095)	-0.173	(0.100)
Services and military	-0.478	(0.122)	-0.452	(0.116)	-0.518	(0.109)	-0.562	(0.110)
Agriculture and fishery	-0.344	(0.155)	-0.716	(0.176)	-0.595	(0.136)	-0.569	(0.132)
Crafts and rtld trade	-0.096	(0.109)	-0.256	(0.110)	-0.258	(0.090)	-0.257	(0.095)
Plant, machine operator	-0.181	(0.111)	-0.332	(0.108)	-0.327	(0.092)	-0.358	(0.097)
Self-employed	-0.102	(0.065)	-0.351	(0.080)	-0.297	(0.063)	-0.320	(0.056)
Age 30-42	-0.258	(0.155)	0.546	(0.130)	-0.096	(0.041)	-0.100	(0.041)
Age 43-48	-0.082	(0.046)	0.368	(0.090)	-0.054	(0.025)	-0.055	(0.025)
Age 54-59	-0.069	(0.073)	-0.357	(0.114)	-0.070	(0.032)	-0.074	(0.029)
Up to 5 yrs ed.	-0.360	(0.080)	-0.418	(0.091)	-0.425	(0.097)	-0.381	(0.085)
From 6 to 8 yrs ed.	-0.235	(0.059)	-0.287	(0.072)	-0.282	(0.064)	-0.240	(0.059)
High school some train	0.193	(0.050)	0.151	(0.062)	0.154	(0.050)	0.163	(0.048)
High school+some college	0.325	(0.067)	0.201	(0.081)	0.263	(0.067)	0.278	(0.067)
University	0.523	(0.072)	0.406	(0.102)	0.508	(0.068)	0.483	(0.068)
White	0.124	(0.047)	0.124	(0.064)	0.159	(0.046)	0.151	(0.045)
Other	0.034	(0.088)	0.114	(0.113)	0.074	(0.105)	0.072	(0.101)
Grown in NE	0.089	(0.053)	0.138	(0.066)	0.081	(0.058)	0.094	(0.055)
Grown in NC	0.019	(0.049)	0.036	(0.055)	0.042	(0.048)	0.044	(0.047)
Grown in West	-0.025	(0.067)	-0.038	(0.072)	-0.024	(0.066)	-0.023	(0.065)
Year of birth	0.000	(0.007)	-0.056	(0.013)	0.002	(0.006)	0.002	(0.005)
Observations	954		1315		6324		7918	
R^2	0.342		0.265		0.219		0.211	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parenthesis

TABLE 24

PSID: MODEL 5, FIRST STAGE REGRESSIONS

	AV3		MIN40		A55		ALL	
Manager, Professional	0.133	(0.101)	0.201	(0.093)	0.059	(0.078)	0.070	(0.081)
Sales	-0.123	(0.138)	-0.140	(0.134)	-0.195	(0.106)	-0.201	(0.104)
Clerk	0.020	(0.114)	0.003	(0.108)	-0.108	(0.090)	-0.170	(0.093)
Services and military	-0.390	(0.121)	-0.337	(0.123)	-0.455	(0.099)	-0.480	(0.098)
Agriculture and fishery	-0.400	(0.232)	-0.343	(0.216)	-0.543	(0.164)	-0.515	(0.149)
Crafts and rtld trade	-0.055	(0.104)	-0.189	(0.109)	-0.236	(0.082)	-0.236	(0.086)
Plant, machine operator	-0.217	(0.107)	-0.345	(0.107)	-0.370	(0.087)	-0.402	(0.090)
Self-employed	-0.034	(0.067)	-0.275	(0.083)	-0.229	(0.063)	-0.246	(0.057)
Age 30-42	-0.210	(0.130)	0.426	(0.126)	-0.094	(0.045)	-0.098	(0.045)
Age 43-48	-0.074	(0.043)	0.309	(0.088)	-0.051	(0.027)	-0.053	(0.026)
Age 54-59	-0.010	(0.071)	-0.279	(0.110)	-0.058	(0.031)	-0.059	(0.028)
Up to 5 yrs education	-0.108	(0.147)	-0.162	(0.177)	-0.271	(0.170)	-0.229	(0.137)
From 6 to 8 yrs education	-0.165	(0.083)	-0.141	(0.106)	-0.126	(0.105)	-0.137	(0.088)
High school and some train	0.201	(0.055)	0.269	(0.070)	0.269	(0.061)	0.219	(0.055)
High school+some college	0.386	(0.065)	0.427	(0.088)	0.398	(0.068)	0.361	(0.062)
University degree	0.518	(0.080)	0.489	(0.106)	0.520	(0.081)	0.493	(0.078)
Postgraduate	0.769	(0.093)	0.696	(0.131)	0.795	(0.086)	0.728	(0.082)
White	0.113	(0.044)	0.154	(0.067)	0.166	(0.042)	0.155	(0.041)
Other	0.087	(0.092)	0.194	(0.121)	0.159	(0.106)	0.145	(0.101)
New England	0.170	(0.107)	0.352	(0.121)	0.257	(0.088)	0.279	(0.079)
Middle Atlantic	0.067	(0.065)	0.139	(0.092)	-0.033	(0.061)	-0.007	(0.055)
East North Central	0.062	(0.068)	0.159	(0.079)	0.049	(0.053)	0.035	(0.048)
South Atlantic	0.027	(0.061)	0.175	(0.083)	0.002	(0.044)	0.000	(0.041)
East South Central	-0.057	(0.070)	0.133	(0.090)	0.021	(0.050)	-0.001	(0.046)
West South Central	-0.053	(0.071)	0.009	(0.091)	-0.161	(0.056)	-0.142	(0.054)
Mountain	-0.256	(0.111)	-0.048	(0.118)	-0.289	(0.093)	-0.269	(0.092)
Pacific	0.079	(0.080)	0.095	(0.101)	0.047	(0.060)	0.035	(0.056)
Agriculture	-0.062	(0.190)	-0.599	(0.179)	-0.182	(0.138)	-0.219	(0.123)
Mining	0.172	(0.101)	0.390	(0.132)	0.215	(0.100)	0.151	(0.103)
Construction	-0.373	(0.070)	-0.441	(0.083)	-0.349	(0.066)	-0.370	(0.064)
Trade	-0.287	(0.070)	-0.312	(0.078)	-0.287	(0.059)	-0.322	(0.055)
Transport and commerce	-0.002	(0.055)	-0.067	(0.082)	-0.015	(0.047)	-0.028	(0.045)
Finance and insurance	0.116	(0.085)	-0.117	(0.123)	0.007	(0.085)	-0.024	(0.074)
Services	-0.340	(0.055)	-0.336	(0.062)	-0.309	(0.049)	-0.353	(0.047)
Public admin.	-0.125	(0.066)	-0.125	(0.084)	-0.104	(0.059)	-0.123	(0.054)
Year of birth	0.003	(0.007)	-0.045	(0.013)	0.001	(0.006)	0.002	(0.005)
Observations	954		1315		6324		7918	
R^2	0.405		0.305		0.251		0.243	

AV3, MIN40: Robust standard errors; A55, ALL: Clustered standard errors in parenthesis

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

APPENDIX C: SECOND STAGE REGRESSIONS: RESIDUAL CORRELATIONS

The following tables report the regression of the residuals from the intergenerational equation (eq.6) on the instruments used in the first stage (the variables of eq.5). Although some of the instruments are statistically significant, the Adjusted R^2 indicates that the correlation is low. In fact its value for the sample *ALL* ranges from 0.01 (Italy) to 0.031 (US).

TABLE 25
GSOEP: RESIDUAL CORRELATIONS

	(1) AV3		(2) MIN40		(3) A55		(4) ALL	
F Age 30-43	0.005	[0.039]	0.001	[0.042]	0.018	[0.020]	0.010	[0.015]
F Age 44-48	0.000	[0.038]	-0.004	[0.040]	-0.000	[0.018]	0.000	[0.014]
F Age 54-59	-0.014	[0.055]	-0.005	[0.059]	-0.010	[0.035]	-0.004	[0.021]
F Turkish national	-0.116**	[0.050]	-0.203***	[0.062]	-0.249***	[0.029]	-0.150***	[0.024]
F Secondary school	-0.134**	[0.057]	-0.185***	[0.071]	-0.171***	[0.031]	-0.090***	[0.026]
F Intermediate school	-0.056	[0.066]	-0.142*	[0.079]	-0.072**	[0.033]	-0.018	[0.029]
F Technical school	0.107	[0.166]	-0.244	[0.241]	-0.511***	[0.154]	-0.221**	[0.107]
F Upper secondary school	-0.030	[0.073]	-0.238***	[0.091]	-0.126***	[0.039]	-0.064**	[0.031]
F No degree	-0.110*	[0.063]	-0.110	[0.084]	-0.179***	[0.036]	-0.087***	[0.029]
F Legislator, manager	0.125**	[0.055]	0.060	[0.072]	0.030	[0.030]	-0.024	[0.024]
F Professional	0.140**	[0.063]	0.238***	[0.078]	0.162***	[0.031]	0.107***	[0.027]
F Clerk	0.165**	[0.074]	0.192*	[0.111]	0.130***	[0.038]	0.069**	[0.032]
F Service worker, market sales	0.377***	[0.094]	0.159	[0.118]	0.217***	[0.048]	0.159***	[0.035]
F Skilled agr. and fishery	0.189*	[0.108]	-0.026	[0.131]	0.088*	[0.052]	0.058	[0.042]
F Crafts and ret. trade	0.221***	[0.081]	-0.038	[0.110]	0.053	[0.042]	0.009	[0.032]
F Plant and machine operator	0.167*	[0.085]	-0.142	[0.118]	-0.052	[0.046]	-0.104***	[0.034]
F Elementary occupation	0.295***	[0.092]	0.030	[0.121]	0.133***	[0.050]	0.089**	[0.036]
High skilled	-0.160*	[0.087]	-0.076	[0.109]	-0.179***	[0.040]	-0.164***	[0.035]
Low skilled	0.106	[0.069]	0.114	[0.101]	0.019	[0.036]	0.003	[0.031]
SE with employees	0.066	[0.097]	0.324***	[0.117]	0.150***	[0.044]	0.155***	[0.039]
SE no employees	-0.053	[0.093]	0.077	[0.109]	0.077*	[0.044]	-0.001	[0.037]
SE farmers	0.212*	[0.118]	0.346**	[0.145]	0.343***	[0.056]	0.229***	[0.047]
Labourers	-0.077	[0.088]	0.213**	[0.104]	0.044	[0.044]	0.044	[0.034]
F Vocational training	0.126***	[0.038]	0.130***	[0.041]	0.048***	[0.018]	0.050***	[0.015]
F University ed.	0.007	[0.065]	0.074	[0.073]	0.016	[0.032]	-0.013	[0.027]
dadloc1989	0.169***	[0.032]	0.114***	[0.033]	0.174***	[0.015]	0.179***	[0.014]
F year of birth	0.001	[0.002]	0.001	[0.002]	0.002*	[0.001]	0.002***	[0.001]
Constant	-2.119	[3.142]	-2.669	[3.646]	-3.288*	[1.739]	-3.377***	[1.247]
Observations	3232		4221		19563		27442	
Adjusted R^2	0.038		0.015		0.037		0.029	

Bootstrapped errors in brackets
F stands for father
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 26
SHIW: RESIDUAL CORRELATIONS

	(1) AV3		(2) MIN		(3) A55		(4) ALL	
F No education	-0.030	[0.045]	-0.032	[0.041]	0.005	[0.024]	0.003	[0.025]
F Primary education	0.015	[0.034]	-0.051*	[0.027]	-0.027	[0.017]	-0.021	[0.018]
F Upper sec. ed.	-0.002	[0.054]	-0.030	[0.042]	-0.041	[0.030]	-0.015	[0.028]
F University ed.	0.014	[0.107]	-0.174**	[0.077]	-0.098*	[0.053]	-0.092*	[0.053]
F Production worker	-0.115***	[0.029]	-0.052**	[0.022]	-0.047***	[0.017]	-0.057***	[0.013]
F Junior manager, officer	0.010	[0.068]	0.020	[0.060]	-0.004	[0.035]	-0.025	[0.037]
F Manager	0.037	[0.149]	-0.051	[0.111]	-0.115	[0.080]	-0.052	[0.070]
F Self-employed	0.128	[0.116]	0.039	[0.071]	0.091*	[0.052]	0.099**	[0.050]
F Industry	-0.043	[0.033]	0.025	[0.027]	0.002	[0.018]	0.004	[0.018]
F Agriculture	0.135***	[0.032]	0.082***	[0.026]	0.040**	[0.019]	0.078***	[0.019]
F Northern Italy	0.114***	[0.029]	0.070***	[0.023]	0.052***	[0.015]	0.084***	[0.015]
F Central Italy	0.031	[0.031]	0.008	[0.024]	-0.023	[0.017]	0.018	[0.017]
F Age 30-43	-0.008	[0.032]	0.004	[0.036]	-0.001	[0.020]	0.002	[0.018]
F Age 44-48	-0.004	[0.033]	0.008	[0.037]	0.001	[0.019]	0.001	[0.019]
F Age 54-59	0.007	[0.056]	0.005	[0.048]	0.004	[0.030]	0.002	[0.026]
Father's y of birth	0.001	[0.001]	0.000	[0.001]	0.000	[0.001]	0.000	[0.001]
Constant	-1.351	[2.571]	-0.117	[2.131]	-0.465	[1.523]	-0.307	[1.542]
Observations	1027		2740		5860		6860	
Adjusted R^2	0.056		0.006		0.006		0.010	

Bootstrapped errors in brackets
F stands for father
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 27

BHPS: RESIDUAL CORRELATIONS

	(1) AV3		(2) MIN		(3) A55		(4) ALL	
F No completed ed	0.018	[0.032]	0.040	[0.051]	-0.002	[0.016]	0.014	[0.014]
F Some ed.	0.053	[0.035]	0.019	[0.066]	0.030	[0.021]	0.053***	[0.016]
F University ed.	-0.008	[0.065]	0.137	[0.089]	-0.014	[0.030]	0.032	[0.023]
F Legislator, manager	-0.177***	[0.056]	-0.059	[0.084]	-0.144***	[0.027]	-0.106***	[0.023]
F Technician, Ass. prof.	-0.040	[0.070]	0.035	[0.085]	-0.068*	[0.036]	-0.026	[0.028]
F Clerk	-0.034	[0.074]	-0.010	[0.114]	-0.040	[0.036]	-0.021	[0.032]
F Service worker, market sales	-0.005	[0.082]	0.027	[0.115]	0.035	[0.036]	0.076**	[0.034]
F Skilled agr. and fishery	0.010	[0.103]	0.094	[0.162]	-0.017	[0.043]	-0.002	[0.040]
F Crafts and rtld trade	-0.090	[0.059]	-0.001	[0.087]	-0.130***	[0.030]	-0.086***	[0.024]
F Plant and machine operator	-0.082	[0.064]	0.052	[0.086]	-0.105***	[0.031]	-0.097***	[0.025]
F Elementary occupation	-0.109	[0.074]	-0.108	[0.116]	-0.162***	[0.034]	-0.117***	[0.029]
F Born in England	0.137***	[0.036]	0.173***	[0.052]	0.134***	[0.017]	0.123***	[0.016]
F Born in Scotland	0.163***	[0.046]	0.152**	[0.070]	0.153***	[0.022]	0.130***	[0.020]
F Born in N.I.	0.357***	[0.124]	0.497***	[0.183]	0.282***	[0.046]	0.179***	[0.042]
F Born in Ireland	0.408***	[0.117]	0.268	[0.197]	0.199***	[0.053]	0.324***	[0.047]
F Born abroad	0.206**	[0.093]	0.281**	[0.132]	0.090*	[0.053]	0.059	[0.043]
Non-white	0.068	[0.087]	-0.167	[0.199]	0.110*	[0.060]	0.089*	[0.046]
F S-E no employees	-0.025	[0.055]	-0.321**	[0.151]	0.012	[0.042]	-0.021	[0.032]
F S-E with employees	0.083	[0.061]	0.169*	[0.088]	0.183***	[0.033]	0.128***	[0.024]
F E, manager	0.123***	[0.048]	0.133*	[0.074]	0.071***	[0.026]	0.049***	[0.018]
F E, foreman or supervisor	-0.040	[0.034]	-0.040	[0.057]	0.010	[0.018]	-0.032**	[0.014]
F Age 30-42	0.027	[0.049]	0.019	[0.069]	0.036*	[0.022]	0.035*	[0.021]
F Age 43-48	0.016	[0.049]	0.005	[0.070]	0.014	[0.021]	0.019	[0.018]
F Age 54-59	0.027	[0.069]	0.038	[0.108]	-0.006	[0.041]	0.016	[0.030]
Father’s y of birth	-0.000	[0.002]	0.000	[0.003]	-0.002	[0.001]	-0.001	[0.001]
Constant	0.137	[3.610]	-0.779	[4.894]	3.181	[2.262]	1.969	[1.591]
Observations	1561		1409		10216		14363	
Adjusted R^2	0.025		0.022		0.018		0.017	

Bootstrapped errors in brackets

F stands for father

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 28
PSID: RESIDUAL CORRELATIONS

	(1) AV3		(2) MIN		(3) A55		(4) ALL	
F Manager, Professional	-0.004	[0.120]	-0.228*	[0.133]	0.007	[0.085]	0.088	[0.055]
F Sales	0.205	[0.134]	0.083	[0.174]	0.069	[0.087]	0.234***	[0.058]
F Clerk	-0.134	[0.136]	-0.232	[0.166]	-0.138	[0.110]	-0.004	[0.071]
F Services and military	0.273**	[0.135]	0.115	[0.150]	0.231**	[0.090]	0.340***	[0.061]
F Agr. and fishery	0.137	[0.181]	-0.092	[0.219]	0.172*	[0.099]	0.215***	[0.079]
F Crafts and rtld trade	-0.044	[0.115]	-0.119	[0.140]	0.043	[0.075]	0.108*	[0.055]
F Plant, machine operator	-0.039	[0.114]	-0.200	[0.139]	-0.000	[0.077]	0.105*	[0.054]
Self-employed	0.246*	[0.129]	0.085	[0.170]	0.284***	[0.070]	0.299***	[0.057]
F Age 30-42	-0.009	[0.078]	-0.012	[0.107]	0.007	[0.043]	-0.004	[0.031]
F Age 43-48	-0.011	[0.070]	-0.025	[0.095]	0.008	[0.036]	0.000	[0.030]
F Age 54-59	0.015	[0.091]	-0.020	[0.132]	0.005	[0.068]	0.006	[0.041]
F Up to 5 yrs ed.	-0.176	[0.182]	0.274	[0.261]	0.067	[0.099]	-0.088	[0.094]
F From 6 to 8 yrs ed.	0.014	[0.091]	0.077	[0.153]	0.105*	[0.055]	0.056	[0.039]
F High school and some train	-0.005	[0.078]	0.063	[0.113]	0.008	[0.045]	-0.005	[0.037]
F High school+some college	-0.015	[0.095]	0.094	[0.128]	0.124**	[0.055]	0.068	[0.043]
F University degree	0.006	[0.102]	0.156	[0.138]	0.167***	[0.062]	0.124***	[0.044]
F Postgraduate	-0.040	[0.111]	0.086	[0.154]	-0.026	[0.070]	-0.056	[0.051]
Non-white	-0.116*	[0.061]	-0.176**	[0.089]	-0.071*	[0.042]	-0.073**	[0.031]
F New England	0.103	[0.135]	-0.015	[0.138]	0.135**	[0.059]	0.123**	[0.053]
F Middle Atlantic	0.058	[0.071]	-0.026	[0.101]	0.058	[0.042]	0.092***	[0.032]
F East North Central	-0.178**	[0.073]	-0.266***	[0.091]	-0.255***	[0.033]	-0.158***	[0.031]
F South Atlantic	-0.113	[0.074]	-0.096	[0.088]	-0.148***	[0.043]	-0.089**	[0.036]
F East South Central	-0.147*	[0.085]	-0.244**	[0.112]	-0.322***	[0.051]	-0.185***	[0.038]
F West South Central	-0.015	[0.092]	0.046	[0.121]	-0.044	[0.054]	0.011	[0.046]
F Mountain	0.040	[0.119]	-0.203	[0.174]	0.028	[0.062]	0.037	[0.048]
F Pacific	-0.151	[0.113]	-0.126	[0.149]	-0.193***	[0.063]	-0.161***	[0.052]
F Agriculture and Mining	0.005	[0.130]	0.084	[0.161]	-0.030	[0.067]	-0.008	[0.058]
F Construction and transport	0.205***	[0.069]	0.091	[0.096]	0.128***	[0.038]	0.124***	[0.032]
F Services and PA	0.083	[0.052]	0.050	[0.065]	0.079***	[0.029]	0.070***	[0.021]
Father's y of birth	-0.001	[0.003]	-0.001	[0.003]	0.000	[0.002]	0.000	[0.001]
Constant	1.145	[6.229]	2.898	[6.663]	-0.498	[3.970]	-0.274	[2.627]
Observations	849		956		5178		7530	
Adjusted R^2	0.037		0.011		0.040		0.031	

Bootstrapped errors in brackets
F stands for father
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix D: Second stage regressions

Table 29

IGE calculated by simultaneously estimating the first and the second stage

	Germany		Italy		UK		US	
Father's ln(LI)	0.602***	[0.015]	0.450***	[0.028]	0.341***	[0.019]	0.609***	[0.029]
Age-40	0.000***	[0.018]	-0.057**	[0.022]	0.019	[0.029]	0.019**	[0.064]
$(\text{Age-40})^2$	0.023***	[0.002]	0.001	[0.005]	0.001	[0.003]	0.012**	[0.005]
$(\text{Age-40})^3$	-0.001***	[0.000]	0.000	[0.000]	-0.000	[0.000]	-0.001	[0.001]
LI(Age-40)	0.011***	[0.002]	0.016***	[0.003]	0.003	[0.003]	0.001	[0.006]
LI(Age-40)^2	-0.003***	[0.000]	-0.000	[0.001]	-0.000	[0.000]	-0.001**	[0.001]
LI(Age-40)^3	0.000***	[0.000]	-0.000	[0.000]	0.000	[0.000]	0.000	[0.000]
Year of birth	0.093***	[0.002]	0.061***	[0.001]	0.049***	[0.002]	-0.010***	[0.003]
F y of birth	-0.036***	[0.003]	-0.018***	[0.001]	-0.024***	[0.001]	0.010***	[0.003]
F Age 30-43	0.086***	[0.015]	0.030	[0.045]	0.070***	[0.017]	-0.079***	[0.029]
F Age 44-48	-0.088***	[0.013]	-0.132***	[0.031]	-0.028***	[0.012]	-0.016	[0.027]
F Age 54-59	-0.280***	[0.021]	-0.208***	[0.054]	-0.223***	[0.021]	0.166***	[0.044]
Controls	Yes		Yes		Yes		Yes	
Observations	27442		6860		14363		7530	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

TABLE 30
GSOEP, WEST AND EAST GERMANY: TS2SLS

	1 WG	1 EG	2 WG	2 EG	3 WG	3 EG
Father's ln(LI)	0.320***	0.529***	0.460***	0.588***	0.225***	0.328***
	[0.013]	[0.041]	[0.019]	[0.052]	[0.019]	[0.047]
Age-40	0.098***	0.108***	0.018	0.107*	-0.010	0.164***
	[0.002]	[0.004]	[0.028]	[0.063]	[0.028]	[0.055]
$(\text{Age-40})^2$	-0.001***	-0.000	0.034***	0.011	0.033***	0.015**
	[0.000]	[0.000]	[0.003]	[0.007]	[0.003]	[0.006]
$(\text{Age-40})^3$	-0.000	-0.000*	-0.001***	-0.000	-0.001***	-0.001**
	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
LI(Age-40)			0.009***	0.000	0.011***	-0.008
			[0.003]	[0.007]	[0.003]	[0.007]
LI(Age-40)^2			-0.004***	-0.001	-0.003***	-0.002**
			[0.000]	[0.001]	[0.000]	[0.001]
LI(Age-40)^3			0.000***	0.000	0.000***	0.000**
			[0.000]	[0.000]	[0.000]	[0.000]
Less than secondary education					-0.343***	-0.457***
					[0.021]	[0.107]
Secondary education					-0.241***	-0.228*
					[0.053]	[0.121]
University education					0.423***	0.652***
					[0.010]	[0.024]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22081	5361	22081	5361	22081	5361
Adjusted R^2	0.259	0.208	0.265	0.208	0.327	0.309

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 31
GSOEP: TS2SLS

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.552***	[0.030]	0.514***	[0.027]	0.583***	[0.013]
Age-40	0.092***	[0.005]	0.109***	[0.006]	0.103***	[0.002]
$(\text{Age-40})^2$	-0.002***	[0.000]	-0.002***	[0.000]	-0.001	[0.000]
$(\text{Age-40})^3$	0.000*	[0.000]	-0.000	[0.000]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	3232		4221		19563	
Adjusted R^2	0.257		0.212		0.248	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 32

GSOEP: TS2SLS, WITH INTERACTION

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.627***	[0.043]	0.577***	[0.034]	0.602***	[0.021]
Age-40	-0.018	[0.067]	-0.124*	[0.069]	-0.114***	[0.040]
$(\text{Age-40})^2$	0.024***	[0.008]	0.025***	[0.007]	0.027***	[0.010]
$(\text{Age-40})^3$	-0.001*	[0.001]	-0.000	[0.001]	-0.001	[0.001]
LI(Age-40)	0.012*	[0.007]	0.025***	[0.008]	0.023***	[0.004]
LI(Age-40)^2	-0.003***	[0.001]	-0.003***	[0.001]	-0.003***	[0.001]
LI(Age-40)^3	0.000*	[0.000]	0.000	[0.000]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	3232		4221		19563	
Adjusted R^2	0.263		0.218		0.249	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

TABLE 33

SHIW: IGE, BY REGION

	North	North, ed	Centre	Centre, ed	South	South, ed
Father's ln(LI)	0.356***	0.199***	0.346***	0.202***	0.420***	0.179***
	[0.053]	[0.062]	[0.072]	[0.074]	[0.043]	[0.045]
Age-40	-0.012	-0.009	-0.001	-0.001	-0.119**	-0.107**
	[0.044]	[0.043]	[0.102]	[0.103]	[0.051]	[0.049]
$(\text{Age-40})^2$	0.011	0.016	0.011	0.013	-0.006	-0.005
	[0.012]	[0.012]	[0.013]	[0.013]	[0.006]	[0.006]
$(\text{Age-40})^3$	-0.001	-0.001	-0.000	-0.000	0.001**	0.001*
	[0.001]	[0.001]	[0.001]	[0.001]	[0.000]	[0.000]
LI(Age-40)	0.010**	0.010**	0.010	0.010	0.022***	0.021***
	[0.005]	[0.005]	[0.012]	[0.012]	[0.006]	[0.006]
LI(Age-40)^2	-0.001	-0.002	-0.001	-0.001	0.001	0.001
	[0.001]	[0.001]	[0.002]	[0.002]	[0.001]	[0.001]
LI(Age-40)^3	0.000	0.000	0.000	0.000	-0.000**	-0.000*
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Educ1: Up to primary educ.		-0.198***		-0.337***		-0.390***
		[0.047]		[0.067]		[0.034]
Educ2: Lower sec educ.		-0.149***		-0.071**		-0.197***
		[0.020]		[0.030]		[0.021]
Educ4: Higher education		0.255***		0.197***		0.191***
		[0.047]		[0.043]		[0.035]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2723	2723	1294	1294	2843	2843
Adjusted R^2	0.330	0.361	0.383	0.406	0.356	0.402

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

TABLE 34
SHIW: TS2SLS

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.374***	[0.035]	0.590***	[0.039]	0.542***	[0.022]
Age-40	0.080***	[0.007]	0.077***	[0.007]	0.077***	[0.004]
$(\text{Age-40})^2$	-0.000	[0.001]	-0.000	[0.001]	-0.001	[0.001]
$(\text{Age-40})^3$	-0.000	[0.000]	-0.000	[0.000]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1027		2740		5860	
Adjusted R^2	0.403		0.371		0.371	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 35
SHIW: TS2SLS, WITH INTERACTION

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.329***	[0.040]	0.538***	[0.046]	0.449***	[0.035]
Age-40	0.042	[0.053]	-0.099	[0.062]	-0.062	[0.070]
$(\text{Age-40})^2$	-0.021**	[0.010]	-0.008	[0.008]	-0.021	[0.016]
$(\text{Age-40})^3$	0.002**	[0.001]	0.001	[0.001]	0.002	[0.001]
LI(Age-40)	0.005	[0.006]	0.021***	[0.007]	0.016**	[0.008]
$\text{LI}(\text{Age-40})^2$	0.002**	[0.001]	0.001	[0.001]	0.002	[0.002]
$\text{LI}(\text{Age-40})^3$	-0.000**	[0.000]	-0.000	[0.000]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1027		2740		5860	
Adjusted R^2	0.405		0.373		0.373	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 36
BHPS: TS2SLS

	(1) AV3		(2) MIN40		(3) A55	
Father's ln(LI)	0.345***	[0.047]	0.336***	[0.064]	0.362***	[0.021]
Age-40	0.029***	[0.009]	0.011	[0.015]	0.046***	[0.002]
$(\text{Age-40})^2$	0.001	[0.001]	-0.003***	[0.001]	-0.001	[0.001]
$(\text{Age-40})^3$	-0.000	[0.000]	0.000	[0.000]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1561		1409		10216	
Adjusted R^2	0.056		0.077		0.103	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 37

BHPS: TS2SLS, WITH INTERACTION

	(1) AV3		(2) MIN40		(3) A55	
Father's ln(LI)	0.337***	[0.059]	0.352***	[0.066]	0.352***	[0.028]
Age-40	0.117	[0.112]	-0.029	[0.242]	0.047	[0.051]
$(\text{Age-40})^2$	0.001	[0.015]	0.012	[0.029]	-0.013	[0.013]
$(\text{Age-40})^3$	-0.001	[0.002]	-0.001	[0.003]	0.001	[0.001]
LI(Age-40)	-0.009	[0.012]	0.004	[0.025]	0.000	[0.005]
$\text{LI}(\text{Age-40})^2$	-0.000	[0.001]	-0.001	[0.003]	0.001	[0.001]
$\text{LI}(\text{Age-40})^3$	0.000	[0.000]	0.000	[0.000]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1561		1409		10216	
Adjusted R^2	0.055		0.075		0.103	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

TABLE 38

PSID:TS2SLS

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.616***	[0.055]	0.623***	[0.066]	0.574***	[0.026]
Age-40	0.037**	[0.019]	-0.100***	[0.034]	0.023***	[0.004]
$(\text{Age-40})^2$	-0.001	[0.001]	-0.010*	[0.006]	-0.001	[0.001]
$(\text{Age-40})^3$	-0.000	[0.000]	0.000	[0.001]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	864		956		5178	
Adjusted R^2	0.146		0.157		0.093	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

TABLE 39

PSID: TS2SLS, WITH INTERACTION

	(1) AV3		(2) MIN		(3) A55	
Father's ln(LI)	0.628***	[0.091]	0.716***	[0.072]	0.639***	[0.041]
Age-40	0.510*	[0.282]	-1.063*	[0.597]	0.030	[0.092]
$(\text{Age-40})^2$	0.018	[0.034]	0.140	[0.166]	0.044*	[0.024]
$(\text{Age-40})^3$	-0.012*	[0.007]	0.020	[0.017]	-0.003	[0.002]
LI(Age-40)	-0.045*	[0.027]	0.090*	[0.054]	-0.001	[0.009]
$\text{LI}(\text{Age-40})^2$	-0.002	[0.003]	-0.014	[0.015]	-0.004*	[0.002]
$\text{LI}(\text{Age-40})^3$	0.001*	[0.001]	-0.002	[0.002]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	864		956		5178	
Adjusted R^2	0.148		0.173		0.094	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

APPENDIX E: IGE TREND

(a) IGE and business cycle

As a further attempt to understand the parameters behind cross-country and overtime differences, Figures 3 and 4 plot the level of IGE by year (not by cohort) versus the the unemployment rate and the detrended GDP as indicator of the business cycle. This is to investigate whether business cycles affect the trend of IGE in a significant way. Mobility is affected by education spending, and during a recession public expenditure might decrease. Thus, the expectation is that they might.

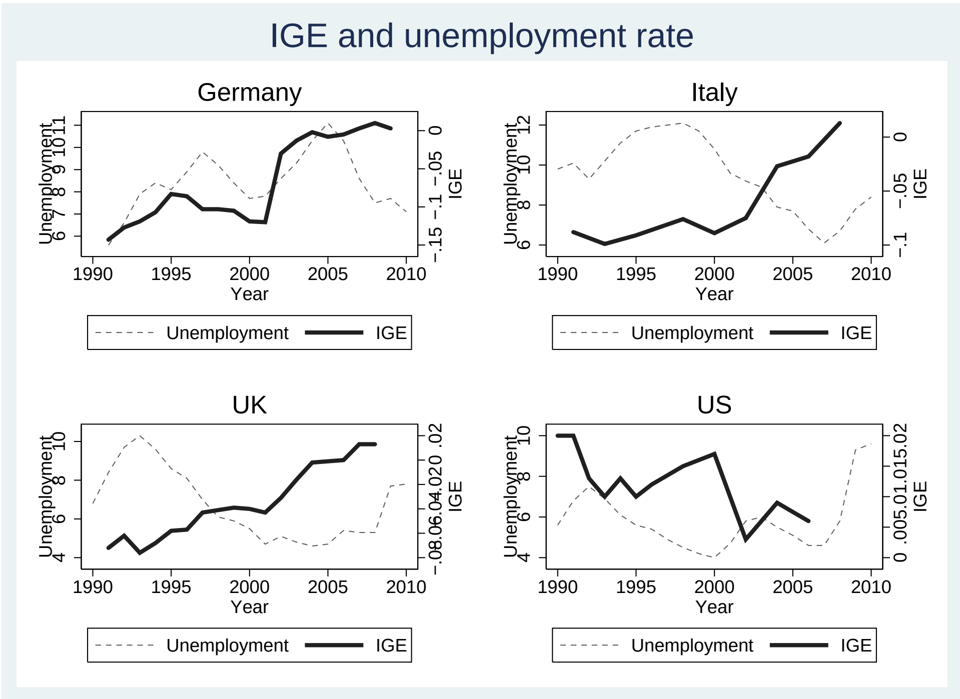

FIGURE 3
INTERGENERATIONAL ELASTICITY VERSUS UNEMPLOYMENT RATE BY YEAR

Overtime the IGE increases in Italy, Germany and in the United Kingdom whereas it is constant in the United States, thus confirming the findings of the analysis by cohort. Although a deeper analysis is needed in order to draw a conclusion on the cyclicality of IGE, the patterns highlighted in the figures suggest that business cycles might have a role in explaining IGE. For all countries, except for the United States where the latest data are for the year 2008, there is a decrease in IGE from 2008, the year when the Great Recession took a steep downward turn. Nonetheless, it is important to mention the limitations of this exercise: firstly, the changes in

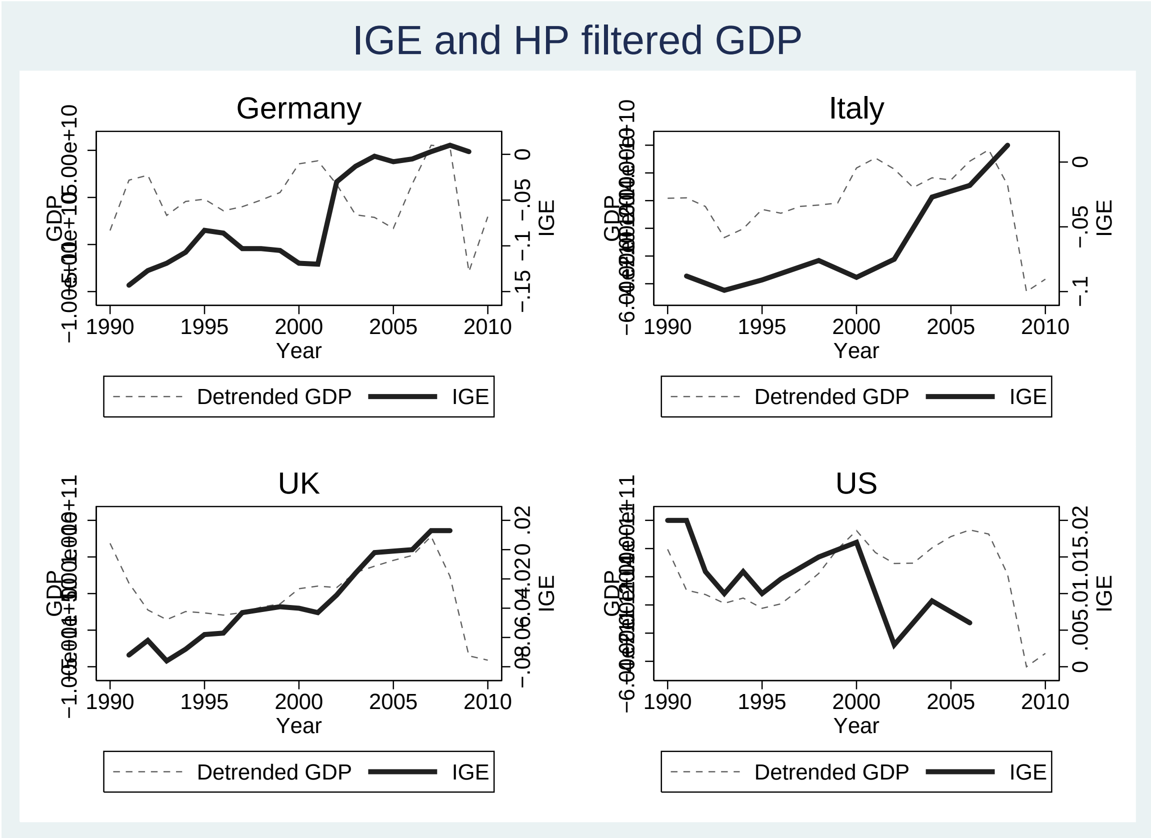

FIGURE 4
INTERGENERATIONAL ELASTICITY VERSUS DETRENDED GDP BY YEAR

IGE from one year to another are limited. Secondly, the number of available years is reduced, especially with SHIW and PSID that have biannual data; finally, the patterns are sometimes contradictory. For example, in Germany, IGE increases with unemployment from 1991 to 1994; then, from 1994 to 1997 the graph suggests a procyclical behaviour; and finally, from 1997 to 2007 it appears countercyclical again. In Italy as well a clear conclusion cannot be drawn: IGE behaves countercyclically until 1998 and then it appears mostly procyclical. In the United Kingdom, the pattern appears to indicate a quite clear positive correlation between IGE and business cycles, at least after 1992. The graph suggests that IGE in the United States might have a similar behaviour as in the United Kingdom. Nonetheless, it is important to remember that results based on PSID are characterised by little change in the slopes overtime; additionally, the differences among them are not statistically different from zero.

(b) Tables of IGE across cohorts

TABLE 40
TS2SLS, COHORTS WITH EDUCATION, ALL

	(1) Germany	(2) Germany	(3) Italy	(4) Italy	(5) UK	(6) UK	(7) US	(8) US
Father's LI*Cohort 1950	0.465***	0.331***	0.387***	0.213***	0.299***	0.168***	0.645***	0.382***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1951	0.461***	0.331***	0.394***	0.217***	0.310***	0.172***	0.626***	0.374***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1952	0.466***	0.334***	0.394***	0.219***	0.317***	0.181***	0.632***	0.375***
	[0.018]	[0.018]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.030]
Father's LI*Cohort 1953	0.483***	0.347***	0.398***	0.225***	0.308***	0.169***	0.623***	0.365***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1954	0.479***	0.344***	0.407***	0.231***	0.331***	0.196***	0.619***	0.363***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1955	0.518***	0.379***	0.423***	0.245***	0.329***	0.192***	0.638***	0.377***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.030]
Father's LI*Cohort 1956	0.514***	0.377***	0.426***	0.250***	0.319***	0.180***	0.620***	0.362***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1957	0.513***	0.374***	0.437***	0.259***	0.324***	0.187***	0.615***	0.359***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1958	0.538***	0.396***	0.443***	0.264***	0.340***	0.202***	0.629***	0.368***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1959	0.554***	0.412***	0.451***	0.272***	0.355***	0.209***	0.621***	0.366***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1960	0.561***	0.421***	0.455***	0.277***	0.342***	0.201***	0.624***	0.366***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1961	0.573***	0.431***	0.465***	0.287***	0.353***	0.213***	0.629***	0.374***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1962	0.579***	0.435***	0.478***	0.302***	0.363***	0.221***	0.615***	0.352***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.021]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1963	0.582***	0.441***	0.480***	0.303***	0.358***	0.214***	0.599***	0.337***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.021]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1964	0.603***	0.462***	0.479***	0.303***	0.375***	0.231***	0.624***	0.360***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.031]	[0.031]
Father's LI*Cohort 1965	0.615***	0.470***	0.484***	0.307***	0.379***	0.232***	0.622***	0.363***
	[0.016]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.030]
Father's LI*Cohort 1966	0.614***	0.469***	0.495***	0.318***	0.379***	0.233***	0.621***	0.359***
	[0.016]	[0.017]	[0.029]	[0.031]	[0.022]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1967	0.622***	0.475***	0.507***	0.331***	0.383***	0.239***	0.605***	0.348***
	[0.016]	[0.016]	[0.030]	[0.032]	[0.022]	[0.022]	[0.029]	[0.030]
Father's LI*Cohort 1968	0.642***	0.493***	0.523***	0.346***	0.388***	0.235***	0.607***	0.354***
	[0.017]	[0.017]	[0.030]	[0.032]	[0.022]	[0.023]	[0.030]	[0.030]
Father's LI*Cohort 1969	0.649***	0.502***	0.514***	0.336***	0.396***	0.248***	0.631***	0.371***
	[0.016]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.031]	[0.031]
Father's LI*Cohort 1970	0.654***	0.504***	0.526***	0.349***	0.390***	0.240***	0.633***	0.366***
	[0.016]	[0.017]	[0.030]	[0.032]	[0.022]	[0.022]	[0.030]	[0.031]
Father's LI*Cohort 1971	0.669***	0.522***	0.546***	0.367***	0.401***	0.256***	0.606***	0.344***
	[0.016]	[0.017]	[0.030]	[0.033]	[0.022]	[0.022]	[0.031]	[0.032]
Father's LI*Cohort 1972	0.678***	0.525***	0.528***	0.348***	0.428***	0.275***	0.608***	0.339***
	[0.017]	[0.017]	[0.028]	[0.027]	[0.022]	[0.022]	[0.031]	[0.031]
Father's LI*Cohort 1973	0.695***	0.543***	0.550***	0.371***	0.423***	0.268***	0.630***	0.370***
	[0.016]	[0.017]	[0.030]	[0.033]	[0.022]	[0.022]	[0.031]	[0.032]
Father's LI*Cohort 1974	0.704***	0.547***	0.567***	0.384***	0.423***	0.271***	0.619***	0.362***
	[0.017]	[0.017]	[0.030]	[0.033]	[0.024]	[0.023]	[0.033]	[0.033]
Father's LI*Cohort 1975	0.691***	0.542***	0.560***	0.380***	0.421***	0.265***	0.616***	0.354***
	[0.017]	[0.017]	[0.030]	[0.033]	[0.023]	[0.023]	[0.032]	[0.033]
Father's LI*Cohort 1976	0.705***	0.552***	0.568***	0.386***	0.444***	0.284***	0.649***	0.384***
	[0.017]	[0.017]	[0.034]	[0.037]	[0.025]	[0.025]	[0.031]	[0.032]
Father's LI*Cohort 1977	0.713***	0.548***	0.579***	0.397***	0.457***	0.294***	0.630***	0.366***
	[0.018]	[0.019]	[0.037]	[0.039]	[0.025]	[0.025]	[0.034]	[0.034]
Father's LI*Cohort 1978	0.680***	0.528***	0.585***	0.407***	0.438***	0.293***	0.624***	0.373***
	[0.020]	[0.020]	[0.036]	[0.038]	[0.043]	[0.042]	[0.039]	[0.040]
Father's LI*Cohort 1979	0.709***	0.552***			0.428***	0.290***		
	[0.021]	[0.021]			[0.028]	[0.028]		
Father's LI*Cohort 1980	0.719***	0.561***			0.460***	0.302***		
	[0.031]	[0.030]			[0.031]	[0.030]		
Age-40	-0.060**	-0.083***	-0.105***	-0.096***	-0.038	-0.032	0.008	-0.014
	[0.026]	[0.025]	[0.034]	[0.033]	[0.035]	[0.034]	[0.073]	[0.071]
$(\text{Age-40})^2$	0.020***	0.018***	0.003	0.004	0.003	0.003	0.014**	0.012**
	[0.003]	[0.003]	[0.007]	[0.007]	[0.003]	[0.003]	[0.006]	[0.006]
$(\text{Age-40})^3$	-0.001***	-0.001***	0.000	0.000	-0.000	-0.000	-0.001	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]
LI(Age-40)	0.018***	0.020***	0.021***	0.020***	0.009**	0.008**	0.002	0.004
	[0.003]	[0.003]	[0.004]	[0.004]	[0.004]	[0.004]	[0.007]	[0.007]
LI(Age-40)^2	-0.002***	-0.002***	-0.000	-0.000	-0.000	-0.000	-0.001**	-0.001**
	[0.000]	[0.000]	[0.001]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
LI(Age-40)^3	0.000***	0.000***	-0.000	-0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Less than secondary education		-0.273***		-0.331***		-0.335***		-0.407***
		[0.023]		[0.026]		[0.023]		[0.048]
Secondary education		-0.279***		-0.143***		-0.150***		-0.163***
		[0.051]		[0.013]		[0.012]		[0.024]
University education		0.395***		0.189***		0.272***		0.330***
		[0.010]		[0.025]		[0.014]		[0.024]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27442	27442	6860	6860	14363	14363	7530	7530
Adjusted R^2	0.280	0.329	0.370	0.400	0.134	0.196	0.112	0.170

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 41
GSOEP: TS2SLS, COHORTS

	(1) AV3		(2) MIN40		(3) A55	
Father's LI*Cohort 1950	0.456***	[0.046]	0.399***	[0.037]	0.440***	[0.022]
Father's LI*Cohort 1951	0.456***	[0.045]	0.401***	[0.037]	0.451***	[0.022]
Father's LI*Cohort 1952	0.468***	[0.045]	0.413***	[0.038]	0.449***	[0.022]
Father's LI*Cohort 1953	0.480***	[0.045]	0.432***	[0.037]	0.473***	[0.022]
Father's LI*Cohort 1954	0.476***	[0.044]	0.413***	[0.038]	0.462***	[0.022]
Father's LI*Cohort 1955	0.523***	[0.045]	0.456***	[0.038]	0.501***	[0.022]
Father's LI*Cohort 1956	0.514***	[0.045]	0.460***	[0.036]	0.504***	[0.022]
Father's LI*Cohort 1957	0.515***	[0.043]	0.455***	[0.037]	0.498***	[0.022]
Father's LI*Cohort 1958	0.544***	[0.044]	0.479***	[0.036]	0.526***	[0.022]
Father's LI*Cohort 1959	0.557***	[0.044]	0.486***	[0.036]	0.540***	[0.022]
Father's LI*Cohort 1960	0.566***	[0.044]	0.500***	[0.036]	0.548***	[0.022]
Father's LI*Cohort 1961	0.576***	[0.044]	0.499***	[0.036]	0.559***	[0.022]
Father's LI*Cohort 1962	0.583***	[0.045]	0.549***	[0.036]	0.567***	[0.022]
Father's LI*Cohort 1963	0.583***	[0.043]	0.547***	[0.036]	0.568***	[0.022]
Father's LI*Cohort 1964	0.607***	[0.043]	0.572***	[0.036]	0.590***	[0.021]
Father's LI*Cohort 1965	0.616***	[0.043]	0.575***	[0.035]	0.601***	[0.021]
Father's LI*Cohort 1966	0.619***	[0.043]	0.581***	[0.035]	0.612***	[0.021]
Father's LI*Cohort 1967	0.626***	[0.043]	0.590***	[0.036]	0.627***	[0.021]
Father's LI*Cohort 1968	0.644***	[0.043]	0.607***	[0.036]	0.643***	[0.021]
Father's LI*Cohort 1969	0.647***	[0.043]	0.595***	[0.036]	0.646***	[0.021]
Father's LI*Cohort 1970	0.657***	[0.044]	0.612***	[0.035]	0.649***	[0.021]
Father's LI*Cohort 1971	0.669***	[0.043]	0.618***	[0.035]	0.657***	[0.021]
Father's LI*Cohort 1972	0.671***	[0.044]	0.625***	[0.035]	0.655***	[0.021]
Father's LI*Cohort 1973	0.688***	[0.044]	0.638***	[0.035]	0.669***	[0.021]
Father's LI*Cohort 1974	0.696***	[0.045]	0.656***	[0.036]	0.676***	[0.022]
Father's LI*Cohort 1975	0.689***	[0.045]	0.653***	[0.039]	0.661***	[0.025]
Father's LI*Cohort 1976	0.700***	[0.045]	0.668***	[0.036]		
Father's LI*Cohort 1977	0.704***	[0.046]	0.677***	[0.037]		
Father's LI*Cohort 1978	0.672***	[0.053]	0.650***	[0.041]		
Father's LI*Cohort 1979			0.665***	[0.043]		
Father's LI*Cohort 1980			0.682***	[0.045]		
Age-40	-0.075	[0.075]	-0.217***	[0.078]	-0.174***	[0.041]
$(\text{Age-40})^2$	0.015	[0.010]	0.013*	[0.008]	0.014	[0.011]
$(\text{Age-40})^3$	-0.001	[0.001]	0.000	[0.001]	0.000	[0.001]
LI(Age-40)	0.019**	[0.008]	0.036***	[0.009]	0.029***	[0.004]
LI(Age-40)^2	-0.002	[0.001]	-0.001	[0.001]	-0.001	[0.001]
LI(Age-40)^3	0.000	[0.000]	-0.000	[0.000]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	3232		4221		19563	
Adjusted R^2	0.266		0.227		0.255	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 42
SHIW: TS2SLS, COHORTS

	(1) AV3		(2) MIN40		(3) A55	
Father's LI*Cohort 1950	0.223***	[0.042]	0.456***	[0.046]	0.367***	[0.036]
Father's LI*Cohort 1951	0.224***	[0.043]	0.469***	[0.047]	0.374***	[0.036]
Father's LI*Cohort 1952	0.228***	[0.042]	0.473***	[0.046]	0.375***	[0.036]
Father's LI*Cohort 1953	0.232***	[0.043]	0.472***	[0.046]	0.381***	[0.036]
Father's LI*Cohort 1954	0.244***	[0.042]	0.476***	[0.046]	0.389***	[0.035]
Father's LI*Cohort 1955	0.260***	[0.043]	0.488***	[0.046]	0.403***	[0.036]
Father's LI*Cohort 1956	0.259***	[0.043]	0.500***	[0.046]	0.407***	[0.036]
Father's LI*Cohort 1957	0.267***	[0.043]	0.505***	[0.046]	0.416***	[0.036]
Father's LI*Cohort 1958	0.286***	[0.043]	0.511***	[0.045]	0.420***	[0.036]
Father's LI*Cohort 1959	0.287***	[0.043]	0.513***	[0.046]	0.428***	[0.036]
Father's LI*Cohort 1960	0.291***	[0.043]	0.518***	[0.046]	0.433***	[0.035]
Father's LI*Cohort 1961	0.308***	[0.043]	0.534***	[0.046]	0.442***	[0.036]
Father's LI*Cohort 1962	0.321***	[0.043]	0.541***	[0.046]	0.453***	[0.036]
Father's LI*Cohort 1963	0.317***	[0.043]	0.554***	[0.045]	0.456***	[0.035]
Father's LI*Cohort 1964	0.315***	[0.042]	0.551***	[0.045]	0.458***	[0.035]
Father's LI*Cohort 1965	0.323***	[0.042]	0.552***	[0.045]	0.462***	[0.035]
Father's LI*Cohort 1966	0.336***	[0.041]	0.570***	[0.045]	0.477***	[0.035]
Father's LI*Cohort 1967	0.339***	[0.041]	0.585***	[0.045]	0.488***	[0.036]
Father's LI*Cohort 1968	0.369***	[0.042]	0.600***	[0.046]	0.511***	[0.036]
Father's LI*Cohort 1969	0.363***	[0.044]	0.584***	[0.047]	0.496***	[0.036]
Father's LI*Cohort 1970	0.355***	[0.043]	0.596***	[0.046]	0.507***	[0.035]
Father's LI*Cohort 1971	0.398***	[0.045]	0.612***	[0.045]	0.532***	[0.037]
Father's LI*Cohort 1972	0.374***	[0.044]	0.613***	[0.046]	0.530***	[0.036]
Father's LI*Cohort 1973	0.378***	[0.042]	0.616***	[0.046]	0.526***	[0.036]
Father's LI*Cohort 1974	0.382***	[0.047]	0.642***	[0.046]	0.541***	[0.035]
Father's LI*Cohort 1975	0.386***	[0.045]	0.621***	[0.046]	0.520***	[0.036]
Father's LI*Cohort 1976	0.350***	[0.048]	0.643***	[0.050]		
Father's LI*Cohort 1977			0.642***	[0.050]		
Father's LI*Cohort 1978			0.656***	[0.049]		
Age-40	0.008	[0.060]	-0.186***	[0.069]	-0.126*	[0.073]
$(\text{Age-40})^2$	-0.033***	[0.011]	-0.009	[0.009]	-0.022	[0.017]
$(\text{Age-40})^3$	0.002***	[0.001]	0.001*	[0.001]	0.002	[0.001]
LI(Age-40)	0.010	[0.007]	0.031***	[0.008]	0.023***	[0.008]
$\text{LI}(\text{Age-40})^2$	0.004***	[0.001]	0.001	[0.001]	0.003	[0.002]
$\text{LI}(\text{Age-40})^3$	-0.000***	[0.000]	-0.000*	[0.000]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1027		2740		5860	
Adjusted R^2	0.404		0.374		0.375	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 43
BHPS: TS2SLS, COHORTS

	(1) AV3		(2) MIN40		(3) A55	
Father's LI*Cohort 1950	0.334***	[0.061]	0.329***	[0.069]	0.304***	[0.030]
Father's LI*Cohort 1951	0.336***	[0.060]	0.309***	[0.068]	0.319***	[0.029]
Father's LI*Cohort 1952	0.339***	[0.061]	0.315***	[0.068]	0.325***	[0.028]
Father's LI*Cohort 1953	0.333***	[0.061]	0.302***	[0.068]	0.315***	[0.029]
Father's LI*Cohort 1954	0.355***	[0.062]	0.334***	[0.069]	0.333***	[0.028]
Father's LI*Cohort 1955	0.353***	[0.061]	0.321***	[0.068]	0.335***	[0.028]
Father's LI*Cohort 1956	0.343***	[0.061]	0.308***	[0.066]	0.322***	[0.028]
Father's LI*Cohort 1957	0.348***	[0.060]	0.340***	[0.066]	0.322***	[0.028]
Father's LI*Cohort 1958	0.358***	[0.061]	0.349***	[0.067]	0.339***	[0.028]
Father's LI*Cohort 1959	0.367***	[0.062]	0.355***	[0.067]	0.355***	[0.029]
Father's LI*Cohort 1960	0.352***	[0.061]	0.336***	[0.066]	0.344***	[0.028]
Father's LI*Cohort 1961	0.360***	[0.061]	0.346***	[0.066]	0.354***	[0.028]
Father's LI*Cohort 1962	0.369***	[0.060]	0.363***	[0.065]	0.354***	[0.028]
Father's LI*Cohort 1963	0.362***	[0.060]	0.362***	[0.064]	0.356***	[0.028]
Father's LI*Cohort 1964	0.376***	[0.061]	0.380***	[0.066]	0.372***	[0.027]
Father's LI*Cohort 1965	0.380***	[0.060]	0.381***	[0.067]	0.374***	[0.028]
Father's LI*Cohort 1966	0.387***	[0.060]	0.399***	[0.066]	0.372***	[0.027]
Father's LI*Cohort 1967	0.378***	[0.060]	0.396***	[0.068]	0.378***	[0.028]
Father's LI*Cohort 1968	0.384***	[0.060]	0.415***	[0.066]	0.387***	[0.028]
Father's LI*Cohort 1969	0.397***	[0.060]	0.413***	[0.067]	0.389***	[0.028]
Father's LI*Cohort 1970	0.393***	[0.061]	0.401***	[0.061]	0.379***	[0.027]
Father's LI*Cohort 1971	0.402***	[0.060]	0.407***	[0.066]	0.395***	[0.028]
Father's LI*Cohort 1972	0.428***	[0.061]	0.435***	[0.066]	0.416***	[0.029]
Father's LI*Cohort 1973	0.426***	[0.061]	0.425***	[0.067]	0.394***	[0.031]
Father's LI*Cohort 1974	0.429***	[0.062]	0.439***	[0.066]	0.407***	[0.032]
Father's LI*Cohort 1975	0.420***	[0.062]	0.450***	[0.065]	0.387***	[0.030]
Father's LI*Cohort 1976	0.454***	[0.062]	0.448***	[0.071]		
Father's LI*Cohort 1977	0.454***	[0.064]	0.481***	[0.070]		
Father's LI*Cohort 1978			0.520***	[0.073]		
Father's LI*Cohort 1979			0.486***	[0.071]		
Father's LI*Cohort 1980			0.555***	[0.078]		
Age-40	0.067	[0.125]	-0.247	[0.287]	-0.009	[0.056]
$(\text{Age-40})^2$	0.013	[0.016]	0.028	[0.031]	-0.014	[0.013]
$(\text{Age-40})^3$	-0.002	[0.002]	-0.002	[0.003]	0.001	[0.001]
LI(Age-40)	-0.005	[0.013]	0.026	[0.030]	0.006	[0.006]
LI(Age-40)^2	-0.001	[0.002]	-0.003	[0.003]	0.001	[0.001]
LI(Age-40)^3	0.000	[0.000]	0.000	[0.000]	-0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	1561		1409		10216	
Adjusted R^2	0.058		0.078		0.112	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
LI stands for Labour Income
Controls: the father's and son's years of birth and paternal age

TABLE 44

PSID: TS2SLS, COHORTS

	(1) AV3		(2) MIN40		(3) A55	
Father's LI*Cohort 1950	0.698***	[0.090]	0.753***	[0.075]	0.663***	[0.041]
Father's LI*Cohort 1951	0.672***	[0.090]	0.720***	[0.074]	0.646***	[0.041]
Father's LI*Cohort 1952	0.668***	[0.090]	0.712***	[0.075]	0.652***	[0.042]
Father's LI*Cohort 1953	0.659***	[0.090]	0.720***	[0.073]	0.642***	[0.041]
Father's LI*Cohort 1954	0.650***	[0.091]	0.698***	[0.071]	0.638***	[0.041]
Father's LI*Cohort 1955	0.658***	[0.091]	0.726***	[0.074]	0.656***	[0.041]
Father's LI*Cohort 1956	0.642***	[0.089]	0.725***	[0.073]	0.640***	[0.040]
Father's LI*Cohort 1957	0.636***	[0.092]	0.722***	[0.074]	0.638***	[0.041]
Father's LI*Cohort 1958	0.644***	[0.090]	0.715***	[0.072]	0.642***	[0.040]
Father's LI*Cohort 1959	0.639***	[0.093]	0.729***	[0.074]	0.645***	[0.040]
Father's LI*Cohort 1960	0.644***	[0.093]	0.719***	[0.076]	0.646***	[0.041]
Father's LI*Cohort 1961	0.656***	[0.093]	0.712***	[0.073]	0.653***	[0.040]
Father's LI*Cohort 1962	0.642***	[0.095]	0.714***	[0.075]	0.646***	[0.041]
Father's LI*Cohort 1963	0.627***	[0.094]	0.703***	[0.078]	0.617***	[0.041]
Father's LI*Cohort 1964	0.650***	[0.095]	0.720***	[0.077]	0.656***	[0.041]
Father's LI*Cohort 1965	0.640***	[0.095]	0.723***	[0.074]	0.648***	[0.041]
Father's LI*Cohort 1966	0.645***	[0.095]	0.711***	[0.077]	0.648***	[0.041]
Father's LI*Cohort 1967	0.628***	[0.093]	0.706***	[0.074]	0.627***	[0.040]
Father's LI*Cohort 1968	0.634***	[0.095]	0.697***	[0.073]	0.629***	[0.041]
Father's LI*Cohort 1969	0.649***	[0.095]	0.715***	[0.074]	0.654***	[0.041]
Father's LI*Cohort 1970	0.653***	[0.095]	0.705***	[0.075]	0.654***	[0.041]
Father's LI*Cohort 1971	0.627***	[0.096]	0.683***	[0.077]	0.644***	[0.042]
Father's LI*Cohort 1972	0.631***	[0.097]	0.674***	[0.076]	0.628***	[0.043]
Father's LI*Cohort 1973	0.647***	[0.097]	0.687***	[0.077]	0.650***	[0.043]
Father's LI*Cohort 1974	0.649***	[0.096]	0.677***	[0.078]		
Father's LI*Cohort 1975			0.671***	[0.077]		
Father's LI*Cohort 1976			0.713***	[0.077]		
Father's LI*Cohort 1977			0.695***	[0.083]		
Father's LI*Cohort 1978			0.702***	[0.086]		
Age-40	0.472	[0.296]	-1.135*	[0.687]	0.021	[0.093]
$(\text{Age-40})^2$	0.024	[0.034]	0.120	[0.171]	0.046*	[0.024]
$(\text{Age-40})^3$	-0.011	[0.008]	0.020	[0.018]	-0.003*	[0.002]
LI(Age-40)	-0.042	[0.029]	0.093	[0.061]	0.000	[0.009]
$\text{LI}(\text{Age-40})^2$	-0.003	[0.003]	-0.012	[0.015]	-0.004*	[0.002]
$\text{LI}(\text{Age-40})^3$	0.001	[0.001]	-0.002	[0.002]	0.000	[0.000]
Controls	Yes		Yes		Yes	
Observations	864		956		5178	
Adjusted R^2	0.147		0.164		0.101	

Bootstrapped errors in brackets

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

Appendix F: Conditional quantile regressions

Table 45
2SQR, ALL

	(1) Germany		(2) Italy		(3) UK		(4) US	
q10								
Father's ln(LI)	0.606***	[0.026]	0.471***	[0.052]	0.205***	[0.041]	0.430***	[0.055]
q15								
Father's ln(LI)	0.576***	[0.023]	0.453***	[0.040]	0.283***	[0.037]	0.487***	[0.050]
q25								
Father's ln(LI)	0.581***	[0.019]	0.402***	[0.032]	0.345***	[0.026]	0.517***	[0.032]
q40								
Father's ln(LI)	0.603***	[0.017]	0.392***	[0.024]	0.380***	[0.022]	0.547***	[0.030]
q50								
Father's ln(LI)	0.605***	[0.015]	0.370***	[0.029]	0.384***	[0.020]	0.583***	[0.026]
q75								
Father's ln(LI)	0.610***	[0.017]	0.396***	[0.026]	0.376***	[0.021]	0.623***	[0.025]
q85								
Father's ln(LI)	0.588***	[0.018]	0.463***	[0.030]	0.402***	[0.025]	0.605***	[0.031]
q90								
Father's ln(LI)	0.579***	[0.023]	0.480***	[0.033]	0.389***	[0.032]	0.631***	[0.049]
q95								
Father's ln(LI)	0.559***	[0.036]	0.487***	[0.054]	0.297***	[0.050]	0.712***	[0.075]
Controls	Yes		Yes		Yes		Yes	
Observations	27442		6860		14363		7530	

Bootstrapped errors in brackets
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
q stands for quantile
LI stands for Labour Income
Controls: the son‘s normalised age, its square and cube and its interactions with paternal LI, the father's and son's years of birth and paternal age

The test on the difference between pairs of coefficients supports the pattern highlighted in the above Table. Specifically, for the United States the Wald tests on pairs of slopes suggest that some of them are statistically different, particularly the pairs of slopes belonging to distant quantiles (for example, $\rho(0.25)$ versus $\rho(0.95)$, $\rho(0.40)$ versus $\rho(0.95)$). Instead the coefficients between the 50^{th} and the 90^{th} percentiles are not. For Italy, all pairs of coefficients are significantly different from each other, at least at 10 percent level of significance. The only exception is the pair $\rho(0.25)$ and $\rho(0.95)$ for which it is not possible to reject the null hypothesis that $\rho(0.75) - \rho(0.95) = 0$. For the United Kingdom, the coefficients indicate an increasing sensitiveness of the offspring' labour income to paternal earnings, at least until the median. This is confirmed by Wald tests. At 10 percent significance level the median is also different from the extremes: for example $\rho(0.25)$ and $\rho(0.95)$ seem statistically different from $\rho(0.50)$ but not different from each other. Although the $\rho(0.95)$ indicates a lower sensitiveness to paternal income for richer sons, the result is not completely supported by the results of the test: in fact, $\rho(0.75)$ and $\rho(0.95)$ are not statistically different from each other. Finally, the coefficients indicate that the elasticity in Germany is similar at different earnings levels: from the 10th to the 95^{th} percentile the average elasticity ranges from 0.57 to 0.61. The test statistics confirm that the slopes are not significantly different from each other.

Appendix G: Generalised Ordered logits

This appendix reports the matrices of probabilities computed on the generalised ordered logit of eq.10. As already mentioned with reference to eq.10, the control variables are those of eq.6. Consequently, beside a categorical variable for paternal income class, the other covariates of the logit are: the son's age normalised at 40, its square and cube; the interaction between normalised age (squared, and cubed) and paternal income; father's and son's years of birth; finally, a categorical variable for paternal age.

TABLE 46

GENERALISED ORDERED AND ORDERED LOGIT TRANSITION MATRICES

	Generalised Ordered logit TM					Ordered logit TM				
	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
Germany										
1.Quintile	*0.290*	0.235	0.168	0.157	0.151	*0.262*	0.235	0.210	0.169	0.123
	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.002)	(0.003)	(0.003)	(0.003)
2.Quintile	0.210	0.218	0.208	0.183	0.181	0.211	0.207	0.207	0.195	0.181
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)
3.Quintile	0.196	0.211	0.221	0.199	0.173	0.209	0.202	0.203	0.195	0.191
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
4.Quintile	0.168	0.198	0.203	0.209	0.223	0.166	0.194	0.207	0.209	0.225
	(0.005)	(0.006)	(0.005)	(0.006)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)
5.Quintile	0.125	0.151	0.204	0.243	*0.279*	0.115	0.165	0.204	0.231	*0.284*
	(0.003)	(0.005)	(0.006)	(0.005)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.006)
Observations	27442	27442	27442	27442	27442	27442	27442	27442	27442	27442
Italy										
1.Quintile	*0.335*	0.258	0.191	0.128	0.087	*0.330*	0.248	0.205	0.137	0.080
	(0.012)	(0.012)	(0.011)	(0.008)	(0.007)	(0.010)	(0.006)	(0.005)	(0.005)	(0.007)
2.Quintile	0.228	0.211	0.218	0.184	0.160	0.218	0.211	0.224	0.196	0.151
	(0.012)	(0.014)	(0.010)	(0.008)	(0.010)	(0.007)	(0.005)	(0.004)	(0.006)	(0.006)
3.Quintile	0.199	0.217	0.209	0.203	0.172	0.206	0.196	0.216	0.206	0.176
	(0.011)	(0.011)	(0.009)	(0.010)	(0.007)	(0.008)	(0.004)	(0.004)	(0.007)	(0.006)
4.Quintile	0.150	0.167	0.207	0.239	0.237	0.149	0.179	0.208	0.233	0.244
	(0.008)	(0.008)	(0.010)	(0.008)	(0.010)	(0.006)	(0.004)	(0.004)	(0.007)	(0.007)
5.Quintile	0.112	0.121	0.191	0.247	*0.329*	0.090	0.140	0.200	0.246	*0.324*
	(0.011)	(0.008)	(0.009)	(0.008)	(0.011)	(0.006)	(0.005)	(0.005)	(0.008)	(0.011)
Observations	6860	6860	6860	6860	6860	6860	6860	6860	6860	6860
United Kingdom										
1.Quintile	*0.244*	0.226	0.202	0.189	0.139	*0.246*	0.219	0.204	0.182	0.149
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.009)	(0.004)	(0.004)	(0.005)	(0.005)
2.Quintile	0.206	0.222	0.195	0.207	0.170	0.216	0.202	0.200	0.198	0.185
	(0.008)	(0.008)	(0.010)	(0.006)	(0.006)	(0.007)	(0.004)	(0.004)	(0.004)	(0.006)
3.Quintile	0.227	0.216	0.199	0.176	0.181	0.224	0.206	0.199	0.192	0.179
	(0.008)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)	(0.004)	(0.004)	(0.003)	(0.007)
4.Quintile	0.162	0.165	0.205	0.220	0.247	0.153	0.175	0.198	0.223	0.251
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)	(0.003)	(0.004)	(0.003)	(0.006)
5.Quintile	0.159	0.171	0.200	0.208	*0.263*	0.143	0.161	0.204	0.230	*0.249*
	(0.007)	(0.006)	(0.008)	(0.008)	(0.008)	(0.005)	(0.004)	(0.003)	(0.004)	(0.008)
Observations	14363	14363	14363	14363	14363	14363	14363	14363	14363	14363
United States										
1.Quintile	*0.316*	0.274	0.181	0.137	0.091	*0.324*	0.256	0.207	0.113	0.083
	(0.012)	(0.011)	(0.011)	(0.008)	(0.009)	(0.016)	(0.006)	(0.006)	(0.006)	(0.006)
2.Quintile	0.248	0.234	0.226	0.192	0.099	0.250	0.235	0.216	0.152	0.121
	(0.010)	(0.010)	(0.010)	(0.011)	(0.008)	(0.009)	(0.006)	(0.004)	(0.006)	(0.007)
3.Quintile	0.189	0.216	0.264	0.192	0.139	0.194	0.210	0.217	0.202	0.179
	(0.010)	(0.009)	(0.012)	(0.010)	(0.009)	(0.008)	(0.006)	(0.005)	(0.006)	(0.006)
4.Quintile	0.098	0.173	0.200	0.253	0.276	0.140	0.172	0.199	0.251	0.258
	(0.009)	(0.010)	(0.011)	(0.011)	(0.012)	(0.005)	(0.006)	(0.004)	(0.006)	(0.008)
5.Quintile	0.150	0.101	0.133	0.220	*0.396*	0.092	0.126	0.161	0.283	*0.358*
	(0.009)	(0.007)	(0.008)	(0.011)	(0.012)	(0.005)	(0.005)	(0.005)	(0.006)	(0.013)
Observations	7530	7530	7530	7530	7530	7530	7530	7530	7530	7530

Bootstrapped standard errors in parenthesis

Note: Columns indicate quintiles of sons; rows of fathers

TABLE 47
GSOEP: GENERALIZED ORDERED LOGIT TRANSITION MATRICES

	Without top and bottom quintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.212***	0.199***	0.211***	0.177***	0.201***
	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)
2.Quintile	0.218***	0.193***	0.183***	0.212***	0.194***
	(0.010)	(0.011)	(0.008)	(0.009)	(0.009)
3.Quintile	0.192***	0.229***	0.213***	0.196***	0.170***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
4.Quintile	0.202***	0.186***	0.194***	0.204***	0.214***
	(0.007)	(0.011)	(0.006)	(0.008)	(0.008)
5.Quintile	0.167***	0.204***	0.204***	0.203***	0.221***
	(0.007)	(0.010)	(0.007)	(0.010)	(0.007)
Observations	10147	10147	10147	10147	10147

	Without top and bottom deciles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.225***	0.204***	0.183***	0.186***	0.202***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
2.Quintile	0.213***	0.222***	0.194***	0.185***	0.187***
	(0.005)	(0.007)	(0.008)	(0.007)	(0.006)
3.Quintile	0.217***	0.197***	0.213***	0.201***	0.171***
	(0.007)	(0.009)	(0.007)	(0.007)	(0.007)
4.Quintile	0.201***	0.199***	0.184***	0.202***	0.214***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.007)
5.Quintile	0.131***	0.194***	0.234***	0.213***	0.229***
	(0.005)	(0.006)	(0.007)	(0.007)	(0.006)
Observations	17896	17896	17896	17896	17896

	Without top and bottom vintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.254***	0.221***	0.177***	0.167***	0.181***
	(0.005)	(0.006)	(0.005)	(0.004)	(0.005)
2.Quintile	0.213***	0.217***	0.202***	0.187***	0.181***
	(0.005)	(0.006)	(0.006)	(0.004)	(0.004)
3.Quintile	0.213***	0.207***	0.214***	0.196***	0.169***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)
4.Quintile	0.186***	0.194***	0.197***	0.200***	0.222***
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)
5.Quintile	0.121***	0.175***	0.217***	0.238***	0.249***
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
Observations	22340	22340	22340	22340	22340

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 48

SHIW: GENERALIZED ORDERED LOGIT TRANSITION MATRICES

	Without top and bottom quintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.224***	0.235***	0.220***	0.164***	0.157***
	(0.017)	(0.016)	(0.016)	(0.020)	(0.018)
2.Quintile	0.237***	0.194***	0.215***	0.183***	0.171***
	(0.017)	(0.018)	(0.017)	(0.018)	(0.018)
3.Quintile	0.213***	0.206***	0.188***	0.216***	0.176***
	(0.016)	(0.018)	(0.019)	(0.015)	(0.014)
4.Quintile	0.172***	0.190***	0.195***	0.203***	0.240***
	(0.020)	(0.017)	(0.021)	(0.018)	(0.019)
5.Quintile	0.170***	0.181***	0.182***	0.232***	0.235***
	(0.015)	(0.018)	(0.016)	(0.019)	(0.018)
Observations	2539	2539	2539	2539	2539
	Without top and bottom deciles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.292***	0.202***	0.218***	0.164***	0.124***
	(0.016)	(0.013)	(0.014)	(0.012)	(0.008)
2.Quintile	0.220***	0.205***	0.211***	0.179***	0.184***
	(0.014)	(0.013)	(0.012)	(0.012)	(0.011)
3.Quintile	0.210***	0.211***	0.194***	0.200***	0.184***
	(0.013)	(0.015)	(0.011)	(0.012)	(0.011)
4.Quintile	0.185***	0.175***	0.193***	0.222***	0.225***
	(0.012)	(0.012)	(0.015)	(0.011)	(0.013)
5.Quintile	0.141***	0.160***	0.200***	0.235***	0.264***
	(0.010)	(0.010)	(0.012)	(0.013)	(0.015)
Observations	4463	4463	4463	4463	4463
	Without top and bottom vintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.295***	0.234***	0.201***	0.154***	0.115***
	(0.012)	(0.014)	(0.009)	(0.011)	(0.008)
2.Quintile	0.220***	0.222***	0.207***	0.173***	0.178***
	(0.012)	(0.012)	(0.012)	(0.009)	(0.009)
3.Quintile	0.216***	0.214***	0.200***	0.197***	0.173***
	(0.011)	(0.014)	(0.012)	(0.008)	(0.009)
4.Quintile	0.153***	0.192***	0.204***	0.218***	0.232***
	(0.010)	(0.010)	(0.012)	(0.011)	(0.012)
5.Quintile	0.126***	0.155***	0.197***	0.235***	0.286***
	(0.009)	(0.009)	(0.011)	(0.012)	(0.013)
Observations	5598	5598	5598	5598	5598

Bootstrapped standard errors in parenthesis

Note: Columns indicate quintiles of sons; rows of fathers

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 49
BHPS: GENERALIZED ORDERED LOGIT TRANSITION MATRICES

	Without top and bottom quintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.223***	0.207***	0.199***	0.186***	0.185***
	(0.011)	(0.012)	(0.013)	(0.011)	(0.015)
2.Quintile	0.216***	0.195***	0.205***	0.177***	0.207***
	(0.008)	(0.013)	(0.010)	(0.011)	(0.011)
3.Quintile	0.242***	0.205***	0.183***	0.197***	0.173***
	(0.012)	(0.009)	(0.010)	(0.012)	(0.011)
4.Quintile	0.170***	0.216***	0.191***	0.220***	0.203***
	(0.009)	(0.013)	(0.012)	(0.012)	(0.012)
5.Quintile	0.149***	0.177***	0.223***	0.219***	0.232***
	(0.011)	(0.011)	(0.012)	(0.011)	(0.010)
Observations	5185	5185	5185	5185	5185
	Without top and bottom deciles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.223***	0.220***	0.191***	0.181***	0.184***
	(0.008)	(0.010)	(0.010)	(0.009)	(0.010)
2.Quintile	0.232***	0.200***	0.199***	0.200***	0.169***
	(0.010)	(0.010)	(0.008)	(0.009)	(0.008)
3.Quintile	0.232***	0.221***	0.195***	0.193***	0.159***
	(0.009)	(0.010)	(0.011)	(0.010)	(0.009)
4.Quintile	0.165***	0.177***	0.212***	0.218***	0.228***
	(0.008)	(0.008)	(0.010)	(0.008)	(0.010)
5.Quintile	0.146***	0.183***	0.203***	0.209***	0.259***
	(0.007)	(0.011)	(0.009)	(0.010)	(0.010)
Observations	9123	9123	9123	9123	9123
	Without top and bottom vintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.226***	0.217***	0.199***	0.187***	0.171***
	(0.010)	(0.007)	(0.009)	(0.007)	(0.008)
2.Quintile	0.217***	0.218***	0.191***	0.208***	0.165***
	(0.009)	(0.008)	(0.006)	(0.009)	(0.006)
3.Quintile	0.236***	0.215***	0.201***	0.176***	0.172***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)
4.Quintile	0.163***	0.171***	0.213***	0.215***	0.238***
	(0.008)	(0.008)	(0.009)	(0.008)	(0.010)
5.Quintile	0.158***	0.179***	0.197***	0.212***	0.254***
	(0.008)	(0.008)	(0.010)	(0.010)	(0.006)
Observations	11590	11590	11590	11590	11590

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 50

PSID: GENERALIZED ORDERED LOGIT TRANSITION MATRICES

	Without top and bottom quintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.257***	0.210***	0.195***	0.181***	0.156***
	(0.014)	(0.019)	(0.011)	(0.014)	(0.017)
2.Quintile	0.213***	0.169***	0.186***	0.225***	0.208***
	(0.017)	(0.011)	(0.015)	(0.017)	(0.019)
3.Quintile	0.195***	0.249***	0.251***	0.161***	0.145***
	(0.016)	(0.015)	(0.018)	(0.014)	(0.018)
4.Quintile	0.203***	0.172***	0.165***	0.238***	0.223***
	(0.013)	(0.012)	(0.014)	(0.018)	(0.016)
5.Quintile	0.146***	0.188***	0.200***	0.196***	0.270***
	(0.014)	(0.014)	(0.018)	(0.018)	(0.015)
Observations	2942	2942	2942	2942	2942
	Without top and bottom deciles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.317***	0.228***	0.192***	0.144***	0.119***
	(0.013)	(0.011)	(0.013)	(0.008)	(0.010)
2.Quintile	0.224***	0.228***	0.223***	0.211***	0.114***
	(0.010)	(0.012)	(0.015)	(0.011)	(0.009)
3.Quintile	0.192***	0.231***	0.243***	0.184***	0.150***
	(0.012)	(0.012)	(0.013)	(0.012)	(0.009)
4.Quintile	0.122***	0.177***	0.189***	0.246***	0.266***
	(0.010)	(0.016)	(0.009)	(0.014)	(0.014)
5.Quintile	0.149***	0.131***	0.158***	0.213***	0.349***
	(0.008)	(0.010)	(0.010)	(0.011)	(0.014)
Observations	6118	6118	6118	6118	6118
	Without top and bottom vintiles				
	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
1.Quintile	0.317***	0.207***	0.199***	0.164***	0.112***
	(0.015)	(0.012)	(0.013)	(0.009)	(0.009)
2.Quintile	0.225***	0.224***	0.223***	0.201***	0.126***
	(0.013)	(0.015)	(0.016)	(0.015)	(0.012)
3.Quintile	0.184***	0.226***	0.246***	0.173***	0.171***
	(0.011)	(0.016)	(0.013)	(0.012)	(0.012)
4.Quintile	0.150***	0.196***	0.185***	0.234***	0.234***
	(0.011)	(0.010)	(0.011)	(0.017)	(0.011)
5.Quintile	0.130***	0.140***	0.147***	0.227***	0.356***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.016)
Observations	4903	4903	4903	4903	4903

Bootstrapped standard errors in parenthesis

Note: Columns indicate quintiles of sons; rows of fathers

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 51
GENERALIZED ORDERED LOGIT TRANSITION MATRICES

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Germany										
1st	0.219	0.123	0.127	0.118	0.091	0.073	0.058	0.066	0.068	0.057
	(0.008)	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)	(0.005)	(0.006)	(0.004)
2nd	0.125	0.123	0.124	0.096	0.088	0.081	0.093	0.094	0.099	0.076
	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.004)	(0.005)	(0.005)
3rd	0.096	0.100	0.106	0.100	0.106	0.113	0.088	0.095	0.102	0.095
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
4th	0.109	0.116	0.114	0.114	0.101	0.096	0.089	0.092	0.084	0.085
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)	(0.005)
5th	0.081	0.111	0.106	0.114	0.104	0.111	0.107	0.097	0.089	0.081
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.004)	(0.006)	(0.004)
6th	0.083	0.109	0.103	0.107	0.122	0.108	0.103	0.090	0.091	0.086
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)
7th	0.083	0.094	0.104	0.096	0.085	0.093	0.101	0.112	0.107	0.125
	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)	(0.006)
8th	0.069	0.087	0.083	0.113	0.123	0.104	0.105	0.106	0.095	0.116
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)
9th	0.062	0.064	0.073	0.097	0.108	0.116	0.119	0.115	0.125	0.120
	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.005)	(0.006)
10th	0.071	0.052	0.059	0.074	0.085	0.096	0.127	0.127	0.139	0.170
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)
Observations	27442	27442	27442	27442	27442	27442	27442	27442	27442	27442
Italy										
1st	0.226	0.158	0.148	0.129	0.088	0.089	0.065	0.027	0.038	0.032
	(0.012)	(0.011)	(0.009)	(0.011)	(0.012)	(0.010)	(0.010)	(0.006)	(0.007)	(0.006)
2nd	0.122	0.166	0.142	0.097	0.112	0.093	0.091	0.071	0.049	0.056
	(0.014)	(0.012)	(0.009)	(0.014)	(0.013)	(0.011)	(0.011)	(0.011)	(0.010)	(0.009)
3rd	0.104	0.117	0.112	0.104	0.138	0.104	0.101	0.078	0.073	0.069
	(0.013)	(0.012)	(0.015)	(0.009)	(0.011)	(0.008)	(0.008)	(0.010)	(0.007)	(0.010)
4th	0.120	0.118	0.091	0.109	0.111	0.083	0.110	0.080	0.104	0.072
	(0.013)	(0.011)	(0.011)	(0.012)	(0.009)	(0.011)	(0.011)	(0.010)	(0.010)	(0.007)
5th	0.127	0.097	0.126	0.095	0.083	0.111	0.110	0.079	0.099	0.074
	(0.012)	(0.011)	(0.011)	(0.011)	(0.009)	(0.013)	(0.011)	(0.009)	(0.012)	(0.009)
6th	0.085	0.093	0.097	0.113	0.118	0.104	0.123	0.098	0.090	0.078
	(0.010)	(0.009)	(0.014)	(0.011)	(0.011)	(0.013)	(0.012)	(0.011)	(0.010)	(0.011)
7th	0.080	0.089	0.091	0.088	0.095	0.111	0.126	0.119	0.105	0.095
	(0.012)	(0.011)	(0.013)	(0.010)	(0.008)	(0.012)	(0.014)	(0.014)	(0.014)	(0.011)
8th	0.069	0.059	0.077	0.082	0.089	0.118	0.122	0.114	0.131	0.139
	(0.013)	(0.009)	(0.009)	(0.013)	(0.011)	(0.011)	(0.016)	(0.012)	(0.012)	(0.012)
9th	0.050	0.073	0.072	0.069	0.087	0.107	0.120	0.140	0.127	0.156
	(0.009)	(0.009)	(0.012)	(0.010)	(0.010)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)
10th	0.061	0.037	0.059	0.047	0.089	0.100	0.094	0.141	0.152	0.221
	(0.010)	(0.009)	(0.008)	(0.007)	(0.010)	(0.010)	(0.011)	(0.011)	(0.014)	(0.010)
Observations	6860	6860	6860	6860	6860	6860	6860	6860	6860	6860

	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
UK										
1st	0.110	0.112	0.105	0.122	0.113	0.103	0.103	0.102	0.073	0.056
	(0.008)	(0.009)	(0.006)	(0.008)	(0.009)	(0.010)	(0.010)	(0.007)	(0.007)	(0.007)
2nd	0.146	0.122	0.127	0.096	0.096	0.094	0.081	0.091	0.087	0.061
	(0.009)	(0.008)	(0.009)	(0.008)	(0.007)	(0.006)	(0.005)	(0.007)	(0.007)	(0.005)
3rd	0.092	0.113	0.123	0.118	0.109	0.098	0.103	0.094	0.074	0.076
	(0.006)	(0.009)	(0.009)	(0.008)	(0.009)	(0.008)	(0.008)	(0.007)	(0.006)	(0.006)
4th	0.084	0.124	0.116	0.092	0.094	0.090	0.102	0.110	0.086	0.102
	(0.006)	(0.007)	(0.008)	(0.006)	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
5th	0.124	0.118	0.113	0.122	0.109	0.102	0.095	0.087	0.065	0.065
	(0.009)	(0.009)	(0.006)	(0.008)	(0.009)	(0.007)	(0.009)	(0.007)	(0.006)	(0.006)
6th	0.113	0.101	0.100	0.095	0.097	0.091	0.090	0.086	0.080	0.148
	(0.007)	(0.009)	(0.006)	(0.007)	(0.008)	(0.007)	(0.009)	(0.007)	(0.008)	(0.012)
7th	0.095	0.101	0.086	0.087	0.091	0.108	0.111	0.109	0.106	0.104
	(0.007)	(0.009)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)	(0.006)	(0.008)
8th	0.066	0.062	0.074	0.082	0.101	0.109	0.098	0.121	0.150	0.136
	(0.008)	(0.006)	(0.005)	(0.006)	(0.007)	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)
9th	0.111	0.091	0.075	0.094	0.094	0.084	0.102	0.085	0.139	0.124
	(0.006)	(0.007)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.007)
10th	0.059	0.057	0.080	0.091	0.096	0.123	0.115	0.113	0.137	0.129
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.010)	(0.008)	(0.010)	(0.008)	(0.009)
Observations	14363	14363	14363	14363	14363	14363	14363	14363	14363	14363
United States										
1st	0.163	0.197	0.165	0.112	0.086	0.069	0.047	0.065	0.052	0.042
	(0.012)	(0.012)	(0.012)	(0.013)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.007)
2nd	0.141	0.130	0.144	0.130	0.118	0.088	0.078	0.081	0.059	0.032
	(0.011)	(0.011)	(0.016)	(0.013)	(0.012)	(0.010)	(0.008)	(0.010)	(0.008)	(0.009)
3rd	0.134	0.143	0.141	0.122	0.113	0.111	0.084	0.083	0.042	0.027
	(0.016)	(0.013)	(0.011)	(0.012)	(0.012)	(0.011)	(0.010)	(0.009)	(0.010)	(0.007)
4th	0.119	0.098	0.105	0.101	0.117	0.113	0.135	0.083	0.060	0.068
	(0.011)	(0.010)	(0.012)	(0.009)	(0.012)	(0.012)	(0.013)	(0.008)	(0.010)	(0.008)
5th	0.069	0.112	0.104	0.109	0.154	0.146	0.106	0.101	0.064	0.034
	(0.012)	(0.010)	(0.014)	(0.012)	(0.013)	(0.013)	(0.010)	(0.010)	(0.008)	(0.007)
6th	0.122	0.076	0.095	0.124	0.113	0.114	0.088	0.090	0.088	0.091
	(0.013)	(0.015)	(0.013)	(0.010)	(0.010)	(0.013)	(0.008)	(0.009)	(0.010)	(0.014)
7th	0.048	0.048	0.078	0.111	0.085	0.104	0.148	0.119	0.142	0.116
	(0.008)	(0.007)	(0.010)	(0.012)	(0.010)	(0.010)	(0.013)	(0.012)	(0.011)	(0.014)
8th	0.049	0.053	0.057	0.097	0.095	0.120	0.103	0.134	0.148	0.144
	(0.008)	(0.008)	(0.011)	(0.009)	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.014)
9th	0.068	0.071	0.052	0.053	0.061	0.042	0.102	0.128	0.192	0.231
	(0.010)	(0.011)	(0.009)	(0.009)	(0.007)	(0.011)	(0.010)	(0.011)	(0.017)	(0.016)
10th	0.095	0.066	0.054	0.041	0.058	0.103	0.102	0.112	0.155	0.214
	(0.012)	(0.009)	(0.012)	(0.009)	(0.009)	(0.012)	(0.008)	(0.013)	(0.014)	(0.012)
Observations	7530	7530	7530	7530	7530	7530	7530	7530	7530	7530

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers

Appendix H: Education

A further exercise is to analyse how and whether the estimations change is by augmenting the models with four dummy variables for educational attainment (from no secondary education to university degree).

The selected countries have different educational systems. Italy and Germany provide (almost) free public education, including university. Private universities are available but in limited number, and although prestigious, the majority of students attend public institutes. In Italy this also due to the fact that student loans are not available. The families finance the tuition and maintenance fees of the students without a scholarship. The United Kingdom introduced university fees in 1997, although they were capped at 1,000 pounds until 2004; after 2004, they increased to 3,000 pounds. Before that period, university was free and full-time students received maintenance grants. Since 1990, British students have been able to apply for a student loan, repayable after graduation. In the United States, private investment in human capital is more significant. Education is decentralized at the state level, and whereas compulsory education is free, higher education is financed privately. Students and their parents can access private or governmental loans, though, and the returns to tertiary education are higher than in other nations.

(a) TS2SLS

When four educational dummies are added to eq.6[69], β decreases, as expected (Table 52).

The highest reduction occurs in Britain and in the United States (by 41 percent with respect to the β in Table 1), followed by Italy (around 37 percent) and Germany (23 percent). The reduction is consistent with the prediction of model of Becker and Tomes, according to which one of the main channels through which parental earnings affect the income of their children is through investments in their human capital. Nonetheless, it might be interesting to further investigate why the role of education is lower in Germany with respect to Italy where higher

[69] For this estimation and all the following models where education is included, these are the variables used. The excluded category, the same for all the models and all datasets is the dummy indicating whether the individual has an educational level corresponding to further education after secondary school (or upper secondary school) but lower than higher education.

TABLE 52
TS2SLS, WITH EDUCATION

	(1) Germany		(2) Italy		(3) UK		(4) US	
Father's ln(LI)	0.465***	[0.016]	0.288***	[0.032]	0.202***	[0.022]	0.358***	[0.030]
Age-40	-0.033	[0.024]	-0.047	[0.031]	0.013	[0.032]	-0.004	[0.070]
$(\text{Age-40})^2$	0.023***	[0.002]	0.004	[0.007]	0.002	[0.003]	0.011*	[0.006]
$(\text{Age-40})^3$	-0.001***	[0.000]	0.000	[0.000]	-0.000	[0.000]	-0.001	[0.001]
LI(Age-40)	0.014***	[0.003]	0.015***	[0.004]	0.003	[0.004]	0.003	[0.007]
LI(Age-40)^2	-0.002***	[0.000]	-0.000	[0.001]	-0.000	[0.000]	-0.001**	[0.001]
LI(Age-40)^3	0.000***	[0.000]	-0.000	[0.000]	0.000	[0.000]	0.000	[0.000]
Less than secondary ed.	-0.279***	[0.022]	-0.328***	[0.026]	-0.320***	[0.023]	-0.416***	[0.048]
Secondary education	-0.297***	[0.051]	-0.144***	[0.013]	-0.151***	[0.012]	-0.163***	[0.024]
University education	0.393***	[0.010]	0.187***	[0.025]	0.276***	[0.014]	0.325***	[0.023]
Controls	Yes		Yes		Yes		Yes	
Observations	27442		6860		14363		7530	
Adjusted R^2	0.325		0.398		0.188		0.161	

Bootstrapped errors in brackets
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
LI stands for Labour Income
Further education is the excluded educational dummy
Controls: the father's and son's years of birth and paternal age

education is also inexpensive. Although beyond the scope of this research, a possible strategy might be to compare these results with the predictions of Solon's framework (2004) and assess other possible driving factors, such as the quality of education or the social background. As mentioned in section I, Blanden (2013) tests the prediction of Becker and Tomes (1986) and Solon (2004). His finding suggests that mobility is positively related to cross-sectional equality and education spending, and negatively correlated with the returns to education, as suggested by the theory. Blanden also motivates the lower value of β in West Germany by underlining the lower returns to education with respect to other countries, such as Italy. This might explain why the introduction of dummies for education with GSOEP reduces the IGE by a lower amount.

(b) IGE trend by Cohort

As expected, when the four dummies for education are added the absolute value of each coefficient diminishes (Figure 5). Nonetheless the conclusions are the same as in the absence of educational dummies: a positive trend characterises German, Italian and British data, whereas the differences among the slope coefficients estimated on PSID are not significant. The Wald tests are in line with those computed on the model without the indicators of education: overall

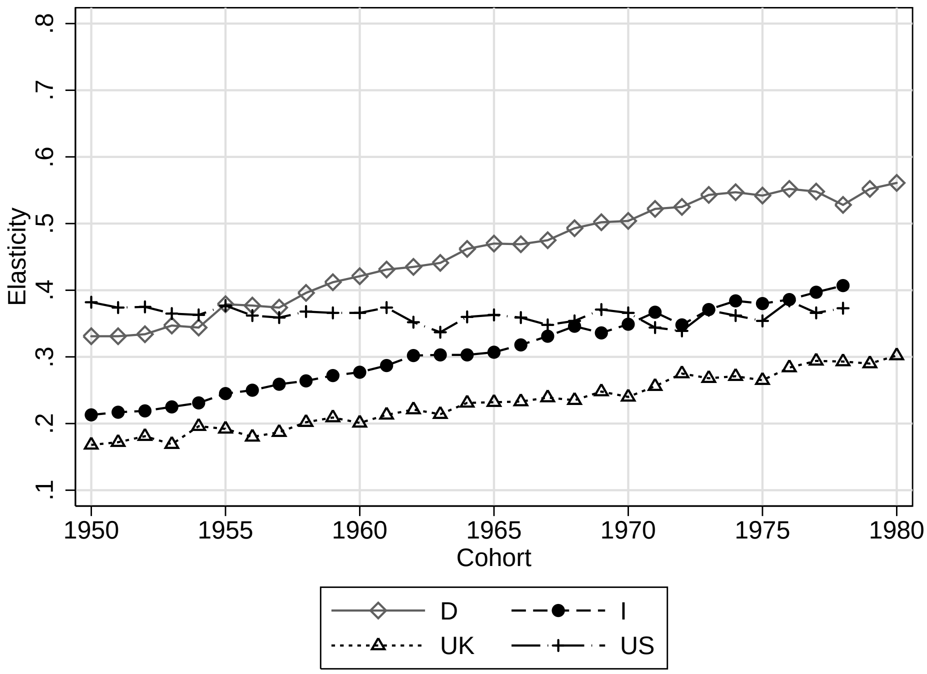

FIGURE 5
INTERGENERATIONAL ELASTICITY BY COHORT WITH EDUCATION

β_c are statistically different from each other, although the power of the test diminishes for closer and later cohorts. Furthermore, Figure 5[70] suggests that the positive IGE trend is not mitigated by the introduction of a variable for education. On the contrary, Table 40 suggests that it has widened the gap between β_{1950} and β_{1980} (or β_{1978} with SHIW). In Germany, the intergenerational elasticity raises from 0.331 for the 1950 cohort to 0.561 for the 1980 cohort. In Italy, β_c increases from 0.213 for $c = 1950$ to 0.407 for $c = 1978$, suggesting an increase of 89 percent. In the United Kingdom, similarly to the Italian and German cases, education appears to enhance the increment of IGE: β augments from 0.168 for β_{1950} to 0.302 for β_{1980}.

Figure 6 shows the percentage decrease in IGE when education is included. It suggests that apart from the United States, where the elasticity decreases between forty and forty-five percent overtime, in the other countries the effect of education on the IGE diminishes for the later cohorts[71].

[70]For additional details, refer to Table 40 in Appendix E.

[71]The findings are somewhat different in Blanden et al. (2005), although the method is different. The author decompose the effect of education on IGE on the 1958 (NCDS) and 1970 (BCS). The findings suggest that conditioning on education has a greater effect on reducing persistence for the younger cohort.

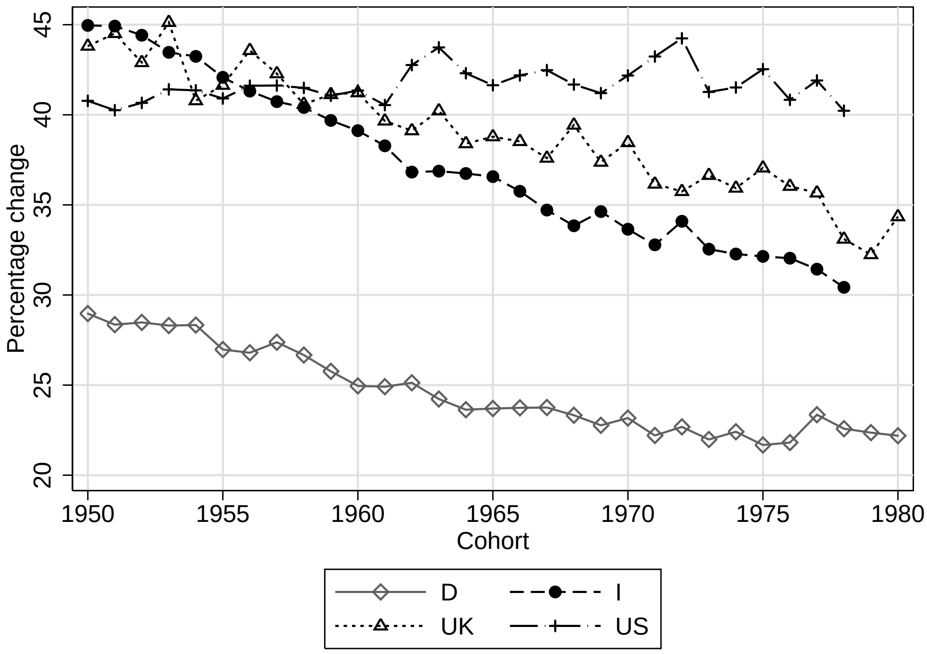

FIGURE 6
PERCENTAGE DECREASE IN IGE BY COHORT WHEN DUMMIES FOR EDUCATION ARE INCLUDED

Explaining the mechanism through which earnings are transmitted across generations goes beyond the scope of this article and as highlighted at the beginning of this document any hypothesis about education at this stage is pure speculation. Nonetheless, the fact that despite the widespread of education the IGE is augmenting and that the percentage increase is higher when education is accounted for suggests the need of a thorough reflection on the role of education on IGE. As mentioned in the previous section, the correlations in Blanden (2013) confirm the theoretical predictions of Becker and Tomes (1986) and Solon (2004): mobility decreases with cross-sectional inequality and the returns to education, whereas it increases with education spending. This does not clarify, however, why the trend in IGE increases when education is included. Indeed, as Blanden mentions the expansion of educational opportunities and of the number of graduates has two effects: returns to education should decrease because of the augmented labour supply; so should the heritability of education, provided more opportunity to access education. Both phenomena should positively affect mobility. Actually, these two events occur. For example, in Italy, according to the OECD, the proportion of graduates has increased from ten percent among the 55-64 year-old people to twenty percent among those aged 25-34.

The higher number of graduates has supposedly reduced the returns to education, as suggested by the OECD 2010 statistics. In 2010, Italian workers aged 25-34 and with a tertiary degree earned only 9 percent more than high school graduates in the same age group, whereas the same gap for the workers aged between 55 and 64 is 96 percent. But the IGE trend is positive in Italy. In the United Kingdom, as well, returns to education appear to have reduced: Blanden et al. (2007) uncover a fall in returns to education after age 16 between the 1958 and 1970 cohorts. Despite this, however, the authors find other factors that might have influenced the decrease in mobility: higher inequality of access to higher education. Using data from BCS, the authors suggest that cognitive and non-cognitive skills, labour market attachment and unequal opportunities to access higher education contribute to explain the extent of IGE.

(c) Transition matrices

This section reports the transition matrices derived from a logit of eq.10, augmented by the four educational dummies. Tables 53 to 56 present the probabilities predicted at two levels of educational attainment: less than secondary school, and when the individual has at least an university degree. Considering that education is one of the main channels through which the paternal income affects the son's future level of earnings, the goal of this exercise is to analyse how the probability changes when education is accounted for. The expectation is that the paternal quintile is less important and that the probabilities are more homogeneous. Before observing the results, however, it is important to keep in mind the possible limitations of this operation. This is why this task has only an explorative goal. It is also true that the aim is not to estimate the exact variation, but to compare the effects of the introduction of this variable across countries. In this context, we assume that the statistical drawbacks of this exercise have similar consequences on the four datasets.

To a certain extent the differences between the probabilities at the extreme elements of the main and minor diagonals are reduced. However, the reduction is limited. And the position with respect to the paternal quintile still plays a role: the probability of being in a given quintile decreases with the distance from the paternal quintile. The findings suggest that Italy and the United States are the countries where the level of education plays the weakest role in promoting equality of opportunities. In the former, the impact of education is even lower than in the latter. In the United States the likelihood of being in the first quintile for an individual without a secondary school level augments from 30.3 percent if $i = 5$ to 50.1 percent if $i = 1$. In Italy, the gap is wider and the likelihood doubles: from 26.4 for $i = 5$ to 52.6 percent $i = 1$. Similarly, although it never reaches the level of 20 percent, the probability of being in the fifth quintile increases with the father's quintile. Furthermore, whereas in Germany and in the United Kingdom an individual with university degree has at least 48 percent chance of being in the top two quintiles independently of the paternal income class, the same likelihood is applicable to the Italian and American samples only for $i \geq 3$. In the United States, however, a university degree reduces the probability of being in the first quintile in a significant way: the highest value is 17 percent for $q_{f,1}$; it was 50 percent with no secondary school and 30 percent when not conditioning on education. Instead, in Italy the results suggest that despite a degree the chance of being in the lowest quintile if the father is in that quantile is still elevated: 25.5 percent.

Thus, the educational dummies do not have the same effect on the predicted probabilities. It might be interesting to uncover the mechanisms that drive the differences in the impact of education: a possible way to go might be to analyse these results in the light of Solon (2004) and particularly in assessing whether the differing returns to education might play a role. For example, according to the OECD the difference between the returns to education of a graduate student and a high school graduate has decreased overtime in Italy. That might explain why the social origin still plays an important role even for graduate students. Another possible explanation might also be that the dummies are not suitable to capture differences in education. For example, the quality of education is not accounted for: in all countries, the sons of richer parents might afford attending private schools, or public universities with higher fees. Finally, some omitted variables that are positively correlated with paternal income, such as job networks, might contribute to explain the results.

Although the introduction of education reduces the persistence patterns, this is true to a certain extent. In fact, conditioning on the same level of education, the paternal quintile still plays a determinant role in affecting the probability of the sons of being in a given income quintile, especially in Italy and the United States.

TABLE 53
GSOEP: GENERALISED ORDERED LOGIT TRANSITION MATRICES

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Less than secondary school					
1.Quintile	0.393***	0.244***	0.197***	0.099***	0.066***
	(0.012)	(0.011)	(0.013)	(0.010)	(0.007)
2.Quintile	0.290***	0.242***	0.249***	0.132***	0.088***
	(0.010)	(0.011)	(0.013)	(0.012)	(0.009)
3.Quintile	0.269***	0.235***	0.266***	0.145***	0.084***
	(0.010)	(0.011)	(0.012)	(0.011)	(0.009)
4.Quintile	0.238***	0.228***	0.254***	0.160***	0.120***
	(0.010)	(0.009)	(0.013)	(0.013)	(0.010)
5.Quintile	0.188***	0.189***	0.278***	0.194***	0.150***
	(0.008)	(0.011)	(0.015)	(0.016)	(0.014)
University education					
1.Quintile	0.181***	0.196***	0.142***	0.196***	0.285***
	(0.006)	(0.008)	(0.007)	(0.008)	(0.008)
2.Quintile	0.126***	0.169***	0.161***	0.204***	0.341***
	(0.006)	(0.005)	(0.006)	(0.009)	(0.008)
3.Quintile	0.116***	0.164***	0.173***	0.227***	0.320***
	(0.006)	(0.004)	(0.006)	(0.009)	(0.006)
4.Quintile	0.099***	0.148***	0.156***	0.232***	0.365***
	(0.004)	(0.005)	(0.006)	(0.008)	(0.009)
5.Quintile	0.079***	0.113***	0.155***	0.255***	0.397***
	(0.004)	(0.004)	(0.005)	(0.007)	(0.007)
Observations	27442	27442	27442	27442	27442

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers

TABLE 54
SHIW: GENERALISED ORDERED LOGIT TRANSITION MATRICES

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Less than secondary school					
1.Quintile	0.526***	0.243***	0.138***	0.062***	0.030***
	(0.019)	(0.023)	(0.013)	(0.009)	(0.006)
2.Quintile	0.414***	0.237***	0.190***	0.101***	0.058***
	(0.026)	(0.024)	(0.019)	(0.015)	(0.012)
3.Quintile	0.374***	0.259***	0.192***	0.114***	0.061***
	(0.021)	(0.025)	(0.020)	(0.016)	(0.013)
4.Quintile	0.309***	0.222***	0.220***	0.157***	0.092***
	(0.020)	(0.023)	(0.021)	(0.023)	(0.018)
5.Quintile	0.264***	0.183***	0.226***	0.184***	0.142***
	(0.021)	(0.027)	(0.030)	(0.027)	(0.024)
University education					
1.Quintile	0.255***	0.247***	0.164***	0.161***	0.172***
	(0.023)	(0.018)	(0.019)	(0.014)	(0.017)
2.Quintile	0.169***	0.186***	0.166***	0.193***	0.286***
	(0.014)	(0.016)	(0.019)	(0.018)	(0.014)
3.Quintile	0.144***	0.190***	0.156***	0.206***	0.305***
	(0.014)	(0.015)	(0.018)	(0.017)	(0.020)
4.Quintile	0.108***	0.141***	0.148***	0.221***	0.383***
	(0.008)	(0.014)	(0.014)	(0.019)	(0.020)
5.Quintile	0.080***	0.097***	0.131***	0.219***	0.472***
	(0.009)	(0.012)	(0.015)	(0.019)	(0.021)
Observations	6860	6860	6860	6860	6860

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers

TABLE 55
BHPS: GENERALISED ORDERED LOGIT TRANSITION MATRICES

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Less than secondary school					
1.Quintile	0.487***	0.214***	0.151***	0.085***	0.063***
	(0.021)	(0.016)	(0.018)	(0.009)	(0.010)
2.Quintile	0.450***	0.221***	0.150***	0.097***	0.082***
	(0.021)	(0.015)	(0.020)	(0.010)	(0.011)
3.Quintile	0.475***	0.206***	0.150***	0.080***	0.088***
	(0.023)	(0.014)	(0.018)	(0.010)	(0.014)
4.Quintile	0.394***	0.186***	0.181***	0.113***	0.125***
	(0.023)	(0.015)	(0.019)	(0.012)	(0.018)
5.Quintile	0.398***	0.193***	0.175***	0.100***	0.135***
	(0.027)	(0.018)	(0.023)	(0.013)	(0.020)
University education					
1.Quintile	0.149***	0.160***	0.206***	0.240***	0.245***
	(0.010)	(0.011)	(0.010)	(0.011)	(0.012)
2.Quintile	0.125***	0.157***	0.189***	0.244***	0.285***
	(0.007)	(0.009)	(0.010)	(0.011)	(0.011)
3.Quintile	0.140***	0.155***	0.198***	0.211***	0.296***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
4.Quintile	0.101***	0.114***	0.190***	0.245***	0.350***
	(0.006)	(0.006)	(0.009)	(0.008)	(0.012)
5.Quintile	0.099***	0.114***	0.183***	0.227***	0.378***
	(0.006)	(0.010)	(0.011)	(0.012)	(0.013)
Observations	14363	14363	14363	14363	14363

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers

TABLE 56
PSID: GENERALISED ORDERED LOGIT TRANSITION MATRICES

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Less than secondary school					
1.Quintile	0.501***	0.293***	0.132***	0.040***	0.035***
	(0.022)	(0.016)	(0.017)	(0.007)	(0.007)
2.Quintile	0.434***	0.295***	0.179***	0.056***	0.036***
	(0.028)	(0.022)	(0.018)	(0.008)	(0.008)
3.Quintile	0.370***	0.307***	0.218***	0.052***	0.052***
	(0.026)	(0.028)	(0.020)	(0.011)	(0.012)
4.Quintile	0.213***	0.315***	0.253***	0.103***	0.116***
	(0.020)	(0.027)	(0.023)	(0.017)	(0.022)
5.Quintile	0.303***	0.211***	0.218***	0.090***	0.177***
	(0.022)	(0.025)	(0.027)	(0.019)	(0.034)
University education					
1.Quintile	0.170***	0.253***	0.228***	0.174***	0.175***
	(0.019)	(0.014)	(0.020)	(0.014)	(0.018)
2.Quintile	0.134***	0.203***	0.258***	0.233***	0.171***
	(0.014)	(0.017)	(0.018)	(0.014)	(0.016)
3.Quintile	0.097***	0.163***	0.270***	0.217***	0.253***
	(0.011)	(0.014)	(0.019)	(0.016)	(0.018)
4.Quintile	0.055***	0.126***	0.182***	0.260***	0.378***
	(0.006)	(0.010)	(0.011)	(0.018)	(0.016)
5.Quintile	0.103***	0.089***	0.130***	0.219***	0.459***
	(0.010)	(0.009)	(0.010)	(0.012)	(0.011)
Observations	7530	7530	7530	7530	7530

Bootstrapped standard errors in parenthesis
Note: Columns indicate quintiles of sons; rows of fathers

Appendix I: Additional robustness checks: outlying observations and the role of self-employment

The following two sections controls if the results are robust to an additional series of sample or model modifications. Specifically, section (a) investigates the effect of excluding the outlying observations, whereas section (b) explicitly accounts for self-employment.

(a) Outliers

Lee and Solon (2009) and Hertz (2007) exclude from their investigations the individuals with an income lower than 700 USD or higher than 700,000 USD in 2000 USD. The level of earnings is not a sample selection criterion for this analysis. Consequently, the sample includes some individuals that report an annual income lower than one USD. The reason is to avoid setting up an additional rule to select the observations included in the analysis. Particularly because it would be based on an arbitrary decision about the labour income threshold. Nonetheless, it may be worth to check the impact of these observations on the estimations. Especially considering that the low figures might result from a reporting or measurement error. Thus, this sections performs the same analysis of the main document on a reduced sample: the individuals who earn extremely low or extremely high amounts are excluded. Specifically, it excludes those reporting less than 1,200 in constant 2005 USD. Whereas the lower band is common to all countries, the upper threshold is based on the income distribution of each subsample (synthetic fathers or sons) and of each dataset. Considering that the goal of this exercise is to control for the outlying observations without affecting the characteristics of the sample, the criterion is rather conservative: the analysis excludes individuals whose earnings are in the top 0.125^{th} percentile of the income distribution. The number of excluded observation is reported in Table 57.

The results are presented in the tables below. Table 58 illustrates the value of the elasticity in each country, with and without controlling for education. A comparison with tables 1 and 52 suggests that the extremes observations have a limited impact on the slope coefficients. For the first model, without education, the IGE decreases in Italy (from 0.459 to 0.435), in

TABLE 57
EXCLUDED OBSERVATIONS PER DADTASET AND SUBSAMPLE

	Subsamples	Labour income < 1200	Above 0.125th percentile	Total
GSOEP	Synthetic Fathers	7	5	12
	Sons	134	34	168
SHIW	Synthetic Fathers	2	1	3
	Sons	13	10	23
BHPS	Synthetic Fathers	15	6	21
	Sons	28	17	45
PSID	Synthetic Fathers	47	9	56
	Sons	37	9	46

Generated by Excel2LaTeX

the United States (from 0.614 to 0.601), and in Germany (0.604 to 0.589), whereas it slightly increases the United Kingdom (0.343 to 0.359). When the dummies for education are included the coefficients do not change significantly in Germany, in the United Kingdom and United States. In Italy the slopes diminishes from 0.288 to 0.261.

Similar results to the main analysis emerge from the investigation of the IGE trend. Although the magnitude of the coefficients is slightly different, and in line with Table 58, the trend remains unchanged. In Germany, Italy and the United Kingdom the impact of paternal income on the offspring's earnings increased overtime, whereas in the United States it is not possible to detect a statistically significant trend[72].

Finally, figure 8 illustrates the leading and the minor diagonals of the generalized ordered logit transition matrices. Similarly to the other part of the analysis, the outlying observations do not affect the results. The shape of the graph is the same as Figure 2. The probabilities do not change by a large amount. The only variation worth being mentioned is the $\Pr(y = 1|q_{f,1})$ calculated on PSID: it augments from 31.6 to 33.2 percent.This is apparent by analysing Table .

Overall, the estimates do not change in a relevant way. Nonetheless, this exercise reveals that the same criterion may have different consequences on different datasets. Thus, particular care needs to be devoted to the choice of the threshold, if one decided to exclude the outliers.

[72] Although not reported, the output table is available. It also includes analysis of the trend with the four dummies for education. The results are similar to those of the main analysis.

TABLE 58

TS2SLS, BASELINE MODEL WITHOUT AND WITH EDUCATION

	(1) Germany	(2) Germany	(3) Italy	(4) Italy	(5) UK	(6) UK	(7) US	(8) US
Father's ln(LI)	0.589***	0.454***	0.435***	0.209***	0.359***	0.209***	0.601***	0.357***
	[0.016]	[0.016]	[0.026]	[0.020]	[0.020]	[0.020]	[0.025]	[0.026]
Age-40	0.005	-0.026	-0.069**	0.004	0.011	0.004	0.019	-0.005
	[0.023]	[0.022]	[0.030]	[0.033]	[0.034]	[0.033]	[0.066]	[0.063]
$(\text{Age-40})^2$	0.022***	0.021***	-0.005	0.001	0.000	0.001	0.011**	0.010**
	[0.003]	[0.002]	[0.004]	[0.003]	[0.003]	[0.003]	[0.005]	[0.005]
$(\text{Age-40})^3$	-0.001***	-0.001***	0.001**	-0.000	0.000	-0.000	-0.000	-0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]
LI(Age-40)	0.010***	0.013***	0.017***	0.004	0.004	0.004	0.001	0.003
	[0.003]	[0.002]	[0.003]	[0.004]	[0.004]	[0.004]	[0.006]	[0.006]
$\text{LI}(\text{Age-40})^2$	-0.002***	-0.002***	0.001	-0.000	-0.000	-0.000	-0.001**	-0.001**
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
$\text{LI}(\text{Age-40})^3$	0.000***	0.000***	-0.000***	0.000	-0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Less than secondary ed.		-0.241***		-0.329***		-0.329***		-0.324***
		[0.021]		[0.021]		[0.021]		[0.035]
Secondary ed.		-0.288***		-0.152***		-0.152***		-0.143***
		[0.048]		[0.010]		[0.010]		[0.022]
University		0.379***		0.262***		0.262***		0.317***
		[0.010]		[0.013]		[0.013]		[0.023]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27274	27274	6837	14318	14318	14318	7484	7484
Adjusted R^2	0.296	0.347	0.402	0.220	0.151	0.220	0.121	0.184

Bootstrapped errors in brackets

${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$

LI stands for Labour Income

Controls: the father's and son's years of birth and paternal age

Further education is the excluded educational dummy

TABLE 59
ORDERED LOGIT TRANSITION MATRICES, NO EXTREMES

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Germany					
1.Quintile	0.288	0.236	0.168	0.157	0.154
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
2.Quintile	0.206	0.218	0.205	0.186	0.187
	(0.004)	(0.006)	(0.004)	(0.004)	(0.005)
3.Quintile	0.202	0.208	0.223	0.196	0.169
	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)
4.Quintile	0.167	0.198	0.205	0.207	0.218
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
5.Quintile	0.124	0.151	0.206	0.243	0.272
	(0.004)	(0.004)	(0.005)	(0.007)	(0.004)
Observations	27274	27274	27274	27274	27274
Italy					
1.Quintile	0.331	0.266	0.193	0.128	0.082
	(0.010)	(0.012)	(0.010)	(0.009)	(0.006)
2.Quintile	0.227	0.205	0.215	0.186	0.167
	(0.011)	(0.012)	(0.012)	(0.008)	(0.008)
3.Quintile	0.201	0.221	0.203	0.205	0.170
	(0.009)	(0.009)	(0.010)	(0.014)	(0.009)
4.Quintile	0.147	0.169	0.207	0.243	0.233
	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)
5.Quintile	0.111	0.125	0.192	0.242	0.330
	(0.007)	(0.009)	(0.011)	(0.013)	(0.011)
Observations	6837	6837	6837	6837	6837
United Kingdom					
1.Quintile	0.238	0.231	0.208	0.188	0.135
	(0.008)	(0.007)	(0.006)	(0.008)	(0.006)
2.Quintile	0.212	0.221	0.189	0.206	0.172
	(0.006)	(0.008)	(0.006)	(0.009)	(0.008)
3.Quintile	0.232	0.214	0.195	0.178	0.181
	(0.009)	(0.007)	(0.007)	(0.006)	(0.008)
4.Quintile	0.157	0.165	0.211	0.219	0.247
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)
5.Quintile	0.160	0.169	0.199	0.208	0.265
	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
Observations	14318	14318	14318	14318	14318
United States					
1.Quintile	0.332	0.271	0.173	0.132	0.092
	(0.011)	(0.013)	(0.010)	(0.008)	(0.008)
2.Quintile	0.224	0.237	0.235	0.207	0.097
	(0.008)	(0.011)	(0.009)	(0.011)	(0.007)
3.Quintile	0.190	0.224	0.265	0.184	0.138
	(0.009)	(0.011)	(0.009)	(0.009)	(0.008)
4.Quintile	0.106	0.164	0.201	0.251	0.277
	(0.009)	(0.010)	(0.013)	(0.012)	(0.008)
5.Quintile	0.148	0.104	0.134	0.219	0.396
	(0.007)	(0.008)	(0.005)	(0.013)	(0.013)
Observations	7484	7484	7484	7484	7484

Bootstrapped standard errors in parenthesis
Notes: Columns indicate quintiles of sons; rows of fathers
The covariates of the logit are: the son's age normalised at 40, its square and cube; the interaction between normalised age (squared, and cubed) and paternal income; father's and son's years of birth; a categorical variable for paternal age.

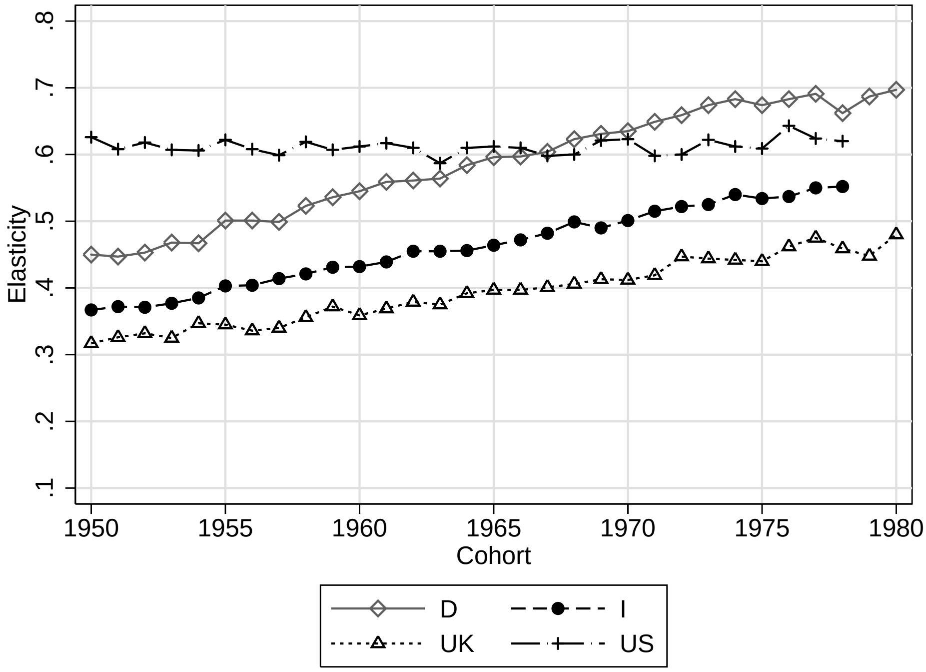

FIGURE 7
IGE TREND ON A REDUCED SAMPLE

(b) Self-employment

An additional robustness check consists in considering the role of self-employment. The sample includes all types of workers: employed, self-employed, full-time and part-time workers. The decision to include all these categories in the main analysis relates to the fact that the type of job might also be affected by the social origin. Thus, excluding for example part-time workers and self-employed might deprive our analysis of important characteristics. Nonetheless, in this section we explore how and whether these characteristics influence the results. Whereas data limitation on SHIW prevent us from discriminating between full- and part-timers, we can account for self-employment. For the purpose of this research, this concern is also the most relevant: it is of particular interest to control for the possibility that a direct channel of labour income transmission between the father and his offspring is through business inheritance, if the father is self-employed. If this happened often, the IGE might be overestimated. Additionally, labour income from self-employment might be more susceptible of measurement error. This might be the case of Italy where it is a common practice to under-report the income from self-employment to pay lower taxes.

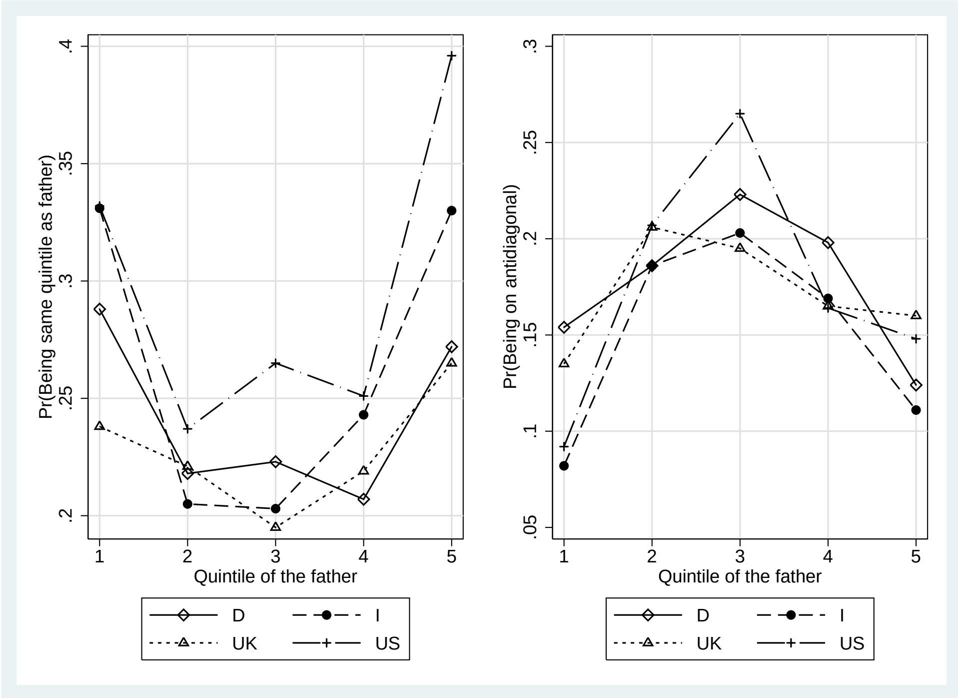

FIGURE 8
ORDERED LOGIT TRANSITION MATRIX ON A REDUCED SAMPLE: PROBABILITIES ON THE MAIN DIAGONAL

For the reasons mentioned above, the original model of eq. 6 is augmented by four variables. Eq. 13 includes one dummy variable to control if the individual is self-employed and a second dummy equal to 1 if the father was self-employed. Additionally, as we are interested in investigating the effect of self-employment through the paternal labour income two interaction terms are included: the interaction of the above dummies with the log of paternal earnings.

$$
\begin{aligned}
y_{st} = {} & \delta_0 + \beta \hat{y}_f + \lambda_1 self + \lambda_2 paself + \lambda_3 self\ \hat{y}_f + \lambda_4 paself\ \hat{y}_f \\
& + \delta_1 (age_s - 40) + \delta_2 (age_s - 40)^2 + \delta_3 (age_s - 40)^3 \\
& + \delta_4 (age_s - 40)\ \hat{y}_f + \delta_5\ (age_s - 40)^2\ \hat{y}_f \\
& + \delta_6\ (age_s - 40)^3\ \hat{y}_f + \delta_7 birthy_s + \delta_8 birthy_f \\
& + \sum_{j=1}^{4} \mu_j dadcat_{fj} + \varepsilon_t
\end{aligned}
\tag{13}
$$

Thus, β indicates the intergenerational elasticity of an employed individual with an employed father; $\beta + \lambda_3$ the IGE of a self-employed individual whose father was not self-employed;

$\beta + \lambda_4$ the IGE of an employed individual with a self-employed father; finally, $\beta + \lambda_3 + \lambda_4$ the elasticity of a self-employed individual with a self-employed father.

It must be underlined that less than 10 percent of the sons report a self-employed father in all datasets, whereas the percentage of self-employed individuals is between 12 and 25 percent, according to the dataset[73].

[73]The summary statistics are provided in Appendix A.

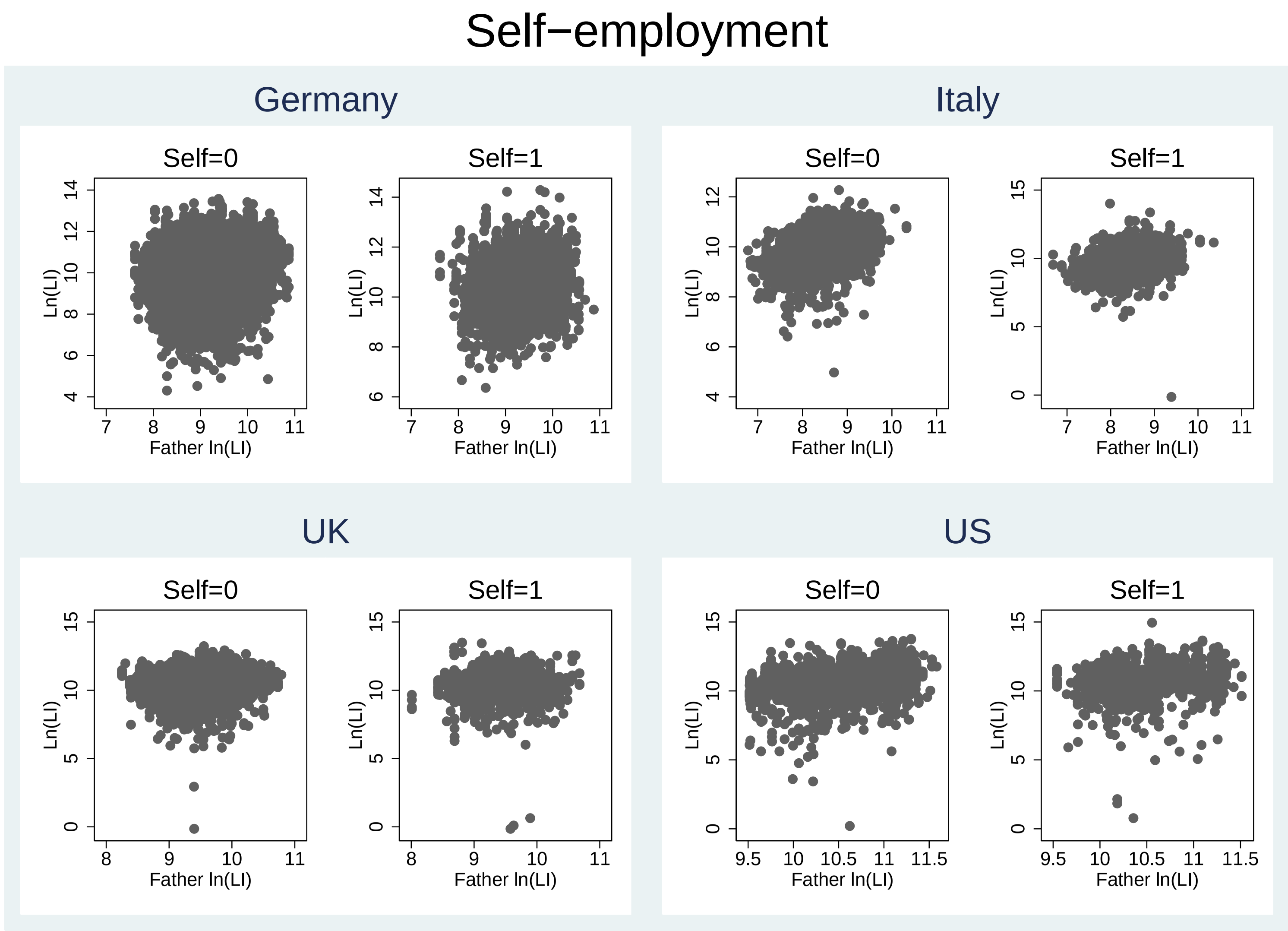

FIGURE 9
DISTRIBUTION OF LN(LI) BY TYPE OF EMPLOYMENT

Figure 9 illustrates the distribution of the log of labour income according to self-employment. For most datasets the distribution does not change drastically, except for Italy where the variance increases and a clear outlier can be detected among the self-employed. The goal of this exercise is to check to what extent self-employed affect the findings of the previous sections.

TABLE 60
TS2SLS, BASELINE MODEL WITHOUT AND WITH EDUCATION

	(1) Germany	(2) Germany	(3) Italy	(4) Italy	(5) UK	(6) UK	(7) US	(8) US
Father's ln(LI)	0.642***	0.504***	0.434***	0.272***	0.402***	0.242***	0.652***	0.394***
	[0.018]	[0.017]	[0.029]	[0.034]	[0.019]	[0.020]	[0.029]	[0.029]
Age-40	-0.007	-0.043*	-0.062*	-0.048	0.008	0.001	0.011	-0.012
	[0.025]	[0.024]	[0.032]	[0.031]	[0.032]	[0.031]	[0.067]	[0.064]
$(\text{Age-40})^2$	0.024***	0.023***	0.002	0.004	0.002	0.002	0.014**	0.013**
	[0.003]	[0.003]	[0.007]	[0.006]	[0.003]	[0.003]	[0.006]	[0.006]
$(\text{Age-40})^3$	-0.001***	-0.001***	0.000	0.000	-0.000	-0.000	-0.001	-0.001
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.001]	[0.001]
self	0.309	0.102	-0.476	-0.441	-0.100	-0.021	0.488	0.373
	[0.343]	[0.352]	[0.317]	[0.310]	[0.481]	[0.479]	[1.019]	[1.008]
Father self	1.683***	1.931***	0.039	0.020	2.272***	1.312***	8.388***	5.747***
	[0.216]	[0.205]	[1.364]	[1.294]	[0.368]	[0.349]	[1.220]	[0.992]
Father's ln(LI)*Self	-0.032	-0.013	0.059	0.054	-0.019	-0.024	-0.074	-0.063
	[0.037]	[0.038]	[0.038]	[0.037]	[0.051]	[0.051]	[0.097]	[0.096]
Father's ln(LI)*F Self	-0.168***	-0.199***	0.007	0.002	-0.234***	-0.136***	-0.779***	-0.526***
	[0.024]	[0.022]	[0.157]	[0.149]	[0.039]	[0.037]	[0.117]	[0.095]
LI(Age-40)	0.011***	0.015***	0.016***	0.015***	0.004	0.004	0.001	0.004
	[0.003]	[0.003]	[0.004]	[0.004]	[0.003]	[0.003]	[0.006]	[0.006]
LI(Age-40)^2	-0.003***	-0.002***	-0.000	-0.000	-0.000	-0.000	-0.001**	-0.001**
	[0.000]	[0.000]	[0.001]	[0.001]	[0.000]	[0.000]	[0.001]	[0.001]
LI(Age-40)^3	0.000***	0.000***	-0.000	-0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Less than secondary ed.		-0.275***		-0.328***		-0.305***		-0.423***
		[0.020]		[0.028]		[0.026]		[0.047]
Secondary education		-0.302***		-0.144***		-0.140***		-0.166***
		[0.050]		[0.012]		[0.012]		[0.021]
University education		0.389***		0.182***		0.270***		0.314***
		[0.011]		[0.027]		[0.015]		[0.022]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27442	27442	6860	6860	14363	14363	7530	7530
Adjusted R^2	0.280	0.328	0.369	0.398	0.145	0.202	0.119	0.184

Bootstrapped errors in brackets
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
LI stands for Labour Income
Further education is the excluded educational dummy
Controls: the father's and son's years of birth and paternal age

The estimation results, reported in Table 60, suggest that being self-employed has no effect on the IGE. Not only the magnitude of each λ_3 is limited (ranging from -0.074 to 0.059), but none of them is statistically significant. The intercept λ_1 is not significant either. On the contrary, a self-employed father has a significant effect: it positively affects the intercept of the earnings equation, whereas it has a negative indirect effect through paternal income. The exception is Italy, where both coefficients are not statistically significant. Overall, except for the Italian case, Table 60 suggests that when the self-employed fathers are included in the sample the amount of IGE decreases. In other words, the value of IGE for employed individuals with employed fathers increases in the three datasets. In the United Kingdom, the IGE of an employed person whose father was employed when the individual was 14 of age increases to 0.40 compared to 0.34 in Table 1. For the United States, the estimated β increases from 0.61 to 0.65. Finally, in Germany the IGE is larger by 4 percentage points (from 0.60 to 0.64).

It appears that business inheritance does not play a prevailing role on these samples. If that had been the case, we would have noticed an increase in the IGE for the sons of the self-employed fathers. It is important to keep in mind one limitation of the TS2SLS. One of the conditions for consistency of the estimators is that they should be identically distributed across the two samples. Instead, for all datasets the percentage of self-employed synthetic fathers is more elevated than the percentage of fathers that are reported as self-employed by the individuals of the main sample. Obviously, this drawback is common to all the other estimation results. However, this model adds one additional potential source of inconsistency to the baseline equation (eq. 6).

Despite the higher values of β when self-employment is not considered, the IGE for employed individuals with employed fathers has a similar trend to the case when no distinction on the basis of self-employment is made. Indeed, figure 10 illustrates that the IGE trend is consistent with that uncovered in figure 1.

In line with the results of Table 60, the magnitude of the slope coefficients increases for all countries except for Italy, where the values do not change significantly with respect to the main analysis. The most apparent consequence is that the IGE trend and the magnitude of the slopes in the United Kingdom and in Italy are very similar to each other. For the United States, the lack of statistical proof against the existence of a trend that was uncovered in the main analysis is confirmed. Similarly, in Germany the positive trend in IGE is reaffirmed.

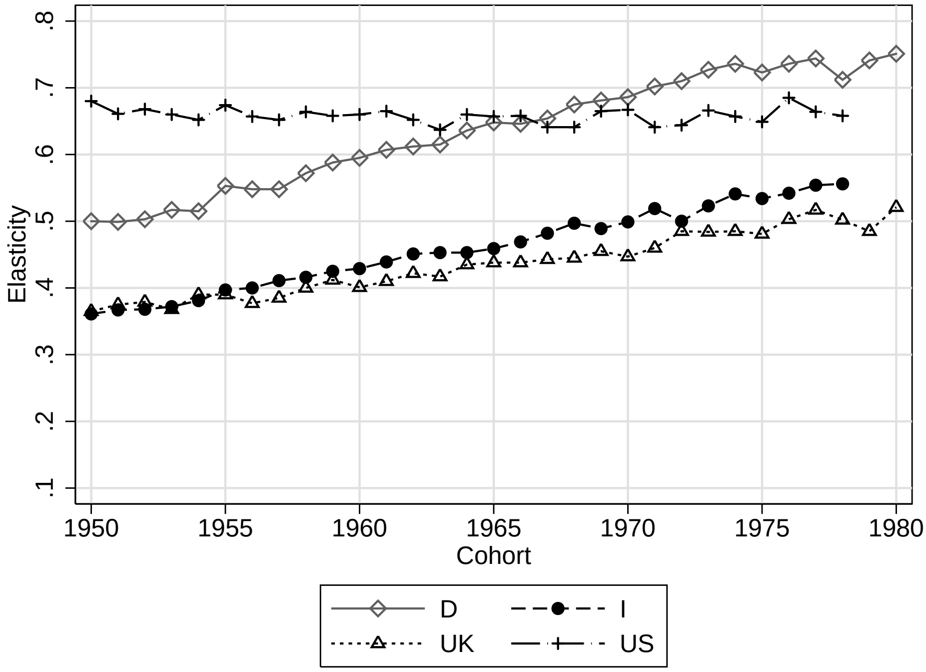

FIGURE 10
IGE TREND BY COHORT OF EMPLOYED INDIVIDUALS WITH EMPLOYED FATHERS

The final step of this exercise consists in estimating the ordered logistic regression. Figures 11 and 12 illustrate the probabilities in the main and minor diagonal for a generalised ordered logit based on eq. 13. Both figures contain the probabilities for two cases: on the panel to the left, all individuals are considered; the panel to the right illustrates the probabilities for the employed individuals with employed fathers[74].

As expected, the panel to the left suggests that the probability for the entire sample remains the same after augmenting the model with self-employment. Instead, when only employed workers are considered, Table 62, the level of persistence at the bottom still remains high, whereas the persistence at the top decreases. This is the case for Italy and Germany, although in both cases the $\Pr(y = 5|q_{f,5}) > 0.25$. In the United Kingdom and in the United States the elements of the main diagonal do not change in a relevant way. The only exception is an increase in the persistence at the first quintile with BHPS-USS, from 24.1 to 24.9 percent. For employed individuals, Figure 12 shows that the probability of being at the opposite quintile of the father's decreases: for all datasets $\Pr(y = 1|q_{f,5})$ and $\Pr(y = 5|q_{f,1})$ are lower than 15 percent. Overall,

[74]The number of self-employed fathers by quintile is too little for certain quintiles and in certain datasets to produce reliable predictions. For this reason they are not reported.

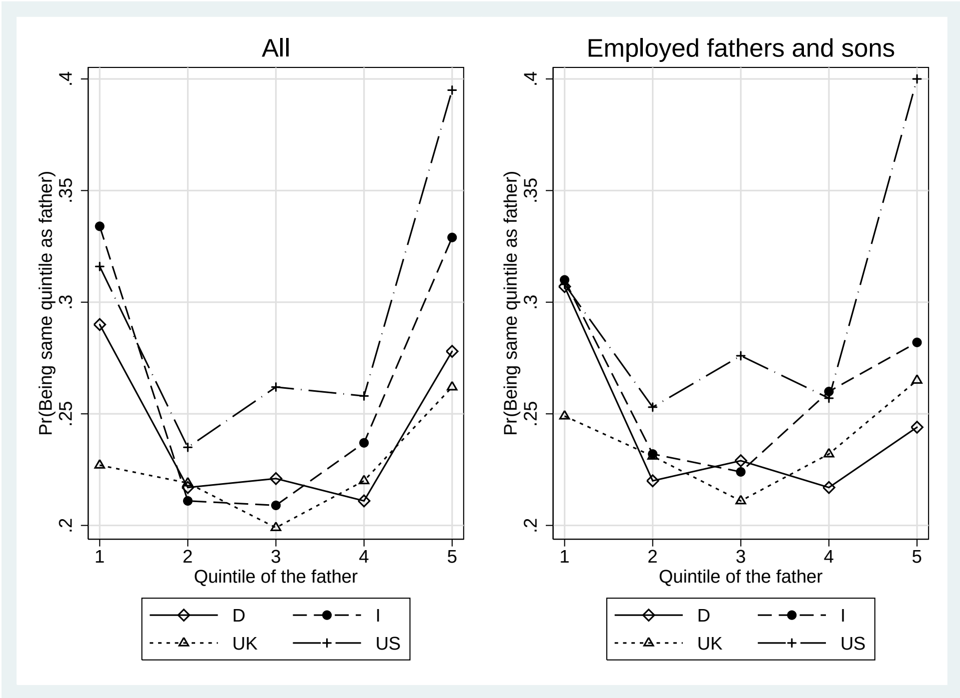

FIGURE 11
ORDERED LOGIT TRANSITION MATRIX ON A REDUCED SAMPLE: PROBABILITIES ON THE MAIN DIAGONAL

the matrix behaves in a similar way as the one in the main analysis. The extremes and the neighbouring elements are characterised by higher probabilities. Additionally, and it is even more so the case with Table 61 the probabilities off the main diagonal are lower than those on the diagonal or near the extremes.

To conclude, this section suggests that self-employment might have an impact on the IGE estimations. The findings indicate that the status of father is more relevant than that of the individual. Particularly, paternal self-employment may reduce IGE. Indeed, the β coefficients for employed individuals with employed fathers increase. The same occurs with the coefficients of the intergenerational elasticity across cohorts, although the patterns remains unchanged. Additionally, whereas the probabilities at the bottom remain as high or they increase, the persistence at the top diminishes for Germany and Italy. Finally, the contrast between the higher probabilities along the main diagonal and the lower values on the minor diagonal is more apparent for all countries.

TABLE 61
GENERALISED ORDERED LOGIT TRANSITION MATRICES, CONTROLLING FOR SELF-EMPLOYMENT

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Germany					
1.Quintile	0.290	0.234	0.168	0.156	0.151
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
2.Quintile	0.210	0.217	0.208	0.183	0.181
	(0.005)	(0.006)	(0.005)	(0.005)	(0.004)
3.Quintile	0.196	0.211	0.221	0.198	0.174
	(0.005)	(0.005)	(0.006)	(0.004)	(0.005)
4.Quintile	0.168	0.197	0.204	0.211	0.221
	(0.006)	(0.004)	(0.004)	(0.006)	(0.005)
5.Quintile	0.124	0.150	0.204	0.243	0.278
	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	27442	27442	27442	27442	27442
Italy					
1.Quintile	0.334	0.261	0.191	0.127	0.087
	(0.012)	(0.011)	(0.012)	(0.007)	(0.007)
2.Quintile	0.229	0.211	0.218	0.184	0.158
	(0.010)	(0.009)	(0.009)	(0.010)	(0.011)
3.Quintile	0.200	0.214	0.209	0.206	0.171
	(0.009)	(0.012)	(0.012)	(0.012)	(0.010)
4.Quintile	0.149	0.166	0.209	0.237	0.239
	(0.009)	(0.008)	(0.011)	(0.010)	(0.009)
5.Quintile	0.109	0.123	0.191	0.247	0.329
	(0.009)	(0.006)	(0.010)	(0.011)	(0.010)
Observations	6860	6860	6860	6860	6860
United Kingdom					
1.Quintile	0.241	0.227	0.204	0.187	0.141
	(0.009)	(0.008)	(0.008)	(0.006)	(0.007)
2.Quintile	0.210	0.219	0.194	0.208	0.169
	(0.008)	(0.007)	(0.008)	(0.007)	(0.005)
3.Quintile	0.228	0.217	0.199	0.175	0.181
	(0.006)	(0.007)	(0.006)	(0.006)	(0.007)
4.Quintile	0.162	0.165	0.206	0.220	0.247
	(0.008)	(0.007)	(0.008)	(0.005)	(0.007)
5.Quintile	0.156	0.172	0.201	0.208	0.262
	(0.006)	(0.007)	(0.007)	(0.009)	(0.007)
Observations	14363	14363	14363	14363	14363
United States					
1.Quintile	0.316	0.276	0.182	0.136	0.089
	(0.010)	(0.011)	(0.008)	(0.010)	(0.008)
2.Quintile	0.247	0.235	0.226	0.191	0.100
	(0.015)	(0.010)	(0.010)	(0.009)	(0.007)
3.Quintile	0.192	0.215	0.262	0.191	0.140
	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)
4.Quintile	0.099	0.173	0.196	0.258	0.274
	(0.046)	(0.011)	(0.011)	(0.011)	(0.010)
5.Quintile	0.147	0.101	0.137	0.220	0.395
	(0.009)	(0.007)	(0.007)	(0.011)	(0.010)
Observations	7530	7530	7530	7530	7530

Bootstrapped standard errors in parenthesis
Notes: Columns indicate quintiles of sons; rows of fathers
The covariates of the logit are: the son's age normalised at 40, its square and cube; the interaction between normalised age (squared, and cubed) and paternal income; father's and son's years of birth; a categorical variable for paternal age; a dummy if the son is self-employed and its interaction with paternal income; a dummy if the father is self-employed and its interaction with paternal income;

TABLE 62
GENERALISED ORDERED LOGIT TRANSITION MATRICES, FOR EMPLOYED INDIVIDUALS WITH EMPLOYED FATHERS

	1.Quintile	2.Quintile	3.Quintile	4.Quintile	5.Quintile
Germany					
1.Quintile	0.307	0.245	0.168	0.147	0.134
2.Quintile	0.208	0.220	0.215	0.187	0.170
3.Quintile	0.198	0.216	0.229	0.200	0.157
4.Quintile	0.172	0.204	0.216	0.217	0.191
5.Quintile	0.123	0.153	0.221	0.256	0.244
Observations	27442	27442	27442	27442	27442
Italy					
1.Quintile	0.310	0.288	0.205	0.132	0.066
2.Quintile	0.215	0.232	0.232	0.195	0.126
3.Quintile	0.180	0.228	0.224	0.225	0.143
4.Quintile	0.136	0.178	0.223	0.260	0.203
5.Quintile	0.093	0.133	0.210	0.283	0.282
Observations	6860	6860	6860	6860	6860
United Kingdom					
1.Quintile	0.249	0.224	0.212	0.193	0.121
2.Quintile	0.196	0.231	0.205	0.213	0.155
3.Quintile	0.197	0.223	0.211	0.187	0.181
4.Quintile	0.141	0.169	0.212	0.232	0.249
5.Quintile	0.138	0.171	0.207	0.219	0.265
Observations	14363	14363	14363	14363	14363
United States					
1.Quintile	0.308	0.299	0.191	0.133	0.069
2.Quintile	0.227	0.253	0.241	0.191	0.089
3.Quintile	0.171	0.222	0.276	0.198	0.134
4.Quintile	0.088	0.176	0.197	0.257	0.282
5.Quintile	0.115	0.100	0.149	0.235	0.400
Observations	7530	7530	7530	7530	7530

Predicted probabilities after generalised ordered logit for employed sons with employed fathers.
Notes: Columns indicate quintiles of sons; rows of fathers
The covariates of the logit are: the son's age normalised at 40, its square and cube; the interaction between normalised age (squared, and cubed) and paternal income; father's and son's years of birth; a categorical variable for paternal age; a dummy if the son is self-employed and its interaction with paternal income; a dummy if the father is self-employed and its interaction with paternal income;

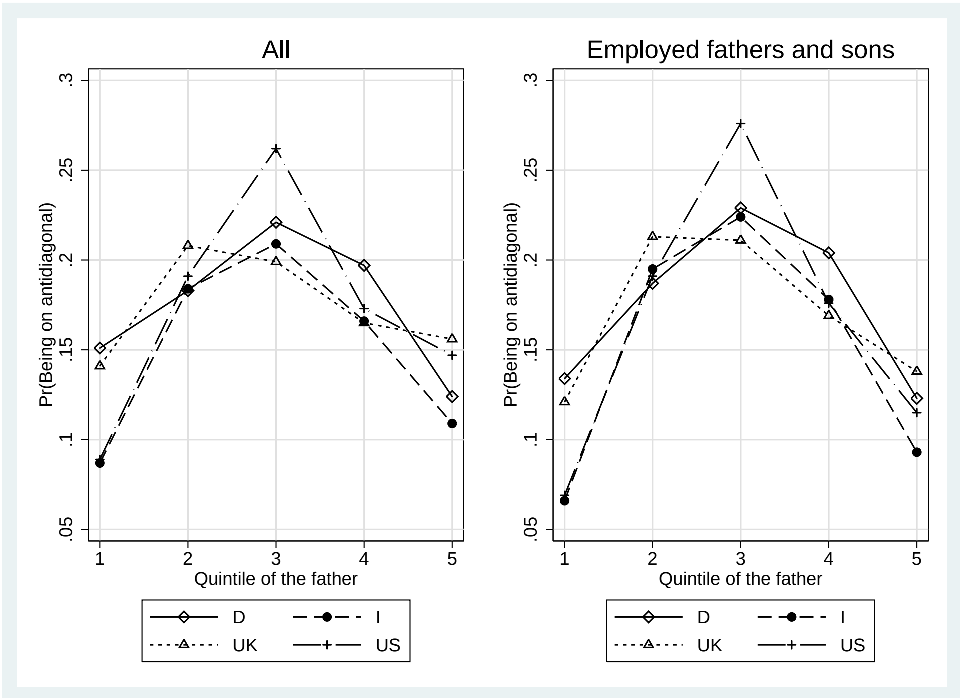

FIGURE 12
ORDERED LOGIT TRANSITION MATRIX ON A REDUCED SAMPLE: PROBABILITIES ON THE MAIN DIAGONAL