University of Konstanz
Department of Economics

Intergenerational poverty transmission in Europe: The role of education

Luna Bellani & Michela Bia

Working Paper Series
2016-02

Intergenerational poverty transmission in Europe: the role of education*

Luna Bellani§
University of Konstanz
Dondena, Bocconi University

Michela Bia‡
LISER

This Version: January 2016

Abstract

This paper examines the role of education as causal channel through which growing up poor affects the individual's economic outcomes as an adult. We contribute to the literature on intergenerational transmission in two ways. First, we apply a potential outcomes approach to quantify the impact of experiencing poverty while growing up and we provide a sensitivity analysis on the unobserved parental ability. Second, we analyze the role of individual human capital accumulation as an intermediate variable and we provide a sensitivity analysis on further possible unobserved confounders. The analysis is based on the module on intergenerational transmission of 2011 of the EU-SILC data, where retrospective questions about parental characteristics (such as education, age, occupation) were asked. We find that, on average, over the 27 European countries considered, growing up poor leads to an increase of 4 percentage points in the risk of being poor and to a decrease of 5% in the adult equivalent income. Moreover, we find that experiencing poverty during childhood will more likely translate into an exclusion from secondary education (of 12 percentage points on average) and that education plays indeed a substantial role accounting for almost 35% of the total effect on adult income.

Keywords: poverty, intergenerational transmission, potential outcome, causal mediation analysis, education

JEL classification codes: D31, I32, J62

*The authors thank Philippe Van Kerm, Guido Schwerdt, Alfonso Flores-Lagunes, Andreas Peichl and the participants to seminar at LISER, to the XX ECINEQ conference in Luxembourg and to the workshop on Public Economics and Inequality in Berlin for comments on this or earlier version of the paper. Bellani acknowledges financial support from an AFR grant (PDR 2011-1) from the Luxembourg Fonds National de la Recherche cofunded under the Marie Curie Actions of the European Commission (FP7-COFUND). Usual disclaimers apply.

§E-mail:luna.bellani@uni-konstanz.de

‡E-mail:michela.bia@liser.lu

1 Introduction

The impact of poverty during childhood on individuals economic outcomes later in life is a topic of active research and a major policy concern in many developed as well as developing countries.

In the US, children as a group are disproportionately represented among the poor: roughly one in five live in poverty compared with one in eight adults (US Census Bureau 2014). Moreover, according to Ratcliffe (2015) about two out of every five children spent at least one year in poverty before turning 18. Persistently poor children are then 43% less likely to finish college than poor children, and 13% less likely to finish high school.

In Europe the picture is similar. From the latest data provided by EUROSTAT, in 2011 in the EU-27 children were the population age group with the highest risk of poverty or social exclusion. The share of children living in a household at risk of poverty or social exclusion ranged from 16-18% in the Nordic countries, Slovenia and the Netherlands to 40-52% in Hungary, Latvia, Romania and Bulgaria. Moreover, 49.2% of children whose parents highest level of education was low were at risk of poverty compared to 7.5% of children whose parents highest level of education was high (Lopez Vilaplana, 2013).

In the European Union the education level of current adults is related to the level of education of their parents in all the Member States, with an average association index,[1] for being low educated, of 14.7 for adults having low educated parents. In Bulgaria and Croatia, this association index was more than 40, while in Norway, Estonia, Denmark and Finland, it was less than 5 (Grundiza and Lopez Vilaplana, 2013).

The economic literature on intergenerational transmission focuses typically on the estimates of the intergenerational elasticity in income or earnings of parents and their offspring. Fewer studies focus on poverty persistence across generations (see among others Mayer (1997); Shea (2000); Acemoglu and Pischke (2001); Ermisch et al. (2004)). These papers find significant impact of parental income or parental financial difficulties on children human capital accumulation and later labor market outcomes, in the range of 5% decrease in education given parental joblessness for Britain, and a 1.4 percentage point increase in the probability of attending college for an increase of income of 10% for the United States.

Blanden et al. (2007) analyze in detail the association between childhood family income and later adult

[1]This association index is calculated as an odds ratio and measures how strongly the low level of education of adults is related to the low level of education of parents compared to the high level of education of parents (Grundiza and Lopez Vilaplana, 2013)

earnings among sons, exploring the role of education, ability, non-cognitive skills and labor market experience in generating intergenerational persistence in the UK. They do so by decomposing the estimated mobility coefficient conditional on those mediating variables. They show that inequalities in achievements at age 16 and in post-compulsory education by family background are extremely important in determining the level of intergenerational mobility. In particular they find a dominant role of education in generating persistence. Cognitive and non-cognitive skills both work indirectly through influencing the level of education obtained, with the cognitive variables accounting for 20% of intergenerational persistence and non-cognitive variables accounting for 10%.

Although many contributions agree that growing up in a poor family is associated to the probability of falling below the poverty threshold in adulthood, the key contentious question for policy is whether this association is truly causal in the sense that poverty in childhood per se influences later outcomes or whether it is driven by other factors correlated with both childhood poverty and later outcomes, such as parenting styles, family structure, neighborhood influences, genetic transmissions, etc. Moreover, as suggested also by Blanden et al. (2007), it is relevant for policy to examine plausible causal channels through which being born poor affects the individual's economic and social status as an adult.

An important part of this poverty persistence is likely to be driven by the effect of parental background on cognitive skills acquired by their children in formal (and informal) education. Recent studies show that educational differences tend to persist across generations, and differences in such persistence explain a large share of the cross-country variation of intergenerational wage correlations (e.g. Solon (2004)).

Experiencing financial difficulties while growing up is not the only determinant of outcomes later in life. Because of the complexity of the process, different statistical techniques have been used, each of which relies on a different set of assumptions. In particular, siblings difference models and instrumental variables approaches have been applied to similar questions. However, while the first method does not guarantee that estimates are unbiased, since there may remain some child-specific factors contributing to potential bias and are based on a selected type of family that can in fact be different with respect to other factors affecting the outcome of the child, as well as their poverty status, the alternative method suffers from the difficulty to find an additional variable which determines childhood poverty status and which at the same time has no direct influence on the outcome variable, i.e. a good instrument, resulting in possible weak instrument bias and anyway providing only local average effect of childhood poverty.

Therefore, this paper's contribution to the literature is twofold: i) we apply a potential outcomes approach for causal inference to quantify the impact of experiencing financial difficulties while growing up and we provide a sensitivity analysis on the unobserved parental ability; ii) we analyze the channel of this poverty transmission, introducing individual human capital accumulation as an intermediate variable and we provide a sensitivity analysis on unobserved confounders also for this mediation analysis.

Our analysis is based on the module on intergenerational transmission of 2011 of the EU-SILC data,[2] where retrospective questions about parental characteristics (such as education, age, occupation) were asked.

We find that, controlling for unobserved confounders, e.g. child ability, being poor in childhood significantly decreases the level of income in adulthood, increasing the average probability of being poor. Moreover, our results reveal a significant role of human capital accumulation in this intergenerational transmission.

The remainder of the paper is organized as follows. Section 2 introduces the estimation strategy and in Section 3 the data used through the whole paper are described. In Section 4 we analyze both the average and the distributional impact of growing up poor, while in Section 5 we focus on the mechanism behind this impact, analyzing the role of education. Section 6 concludes.

2 Estimation strategy

As briefly reviewed in our introduction, standard parametric models rely on strong assumptions about parents and individuals behavior as well as about the mechanisms of poverty transmission.

In this paper we apply a different approach to this question and we follow the framework of potential outcomes approach for causal inference (Rubin, 1974, 1978), which considers a randomized experiment where (a) subjects are randomly selected from the target population; (b) a binary treatment is randomly allocated to the subjects; (c) there are no hidden versions of the treatment and there is no interference between units (Stable Unit Treatment Value Assumption - SUTVA) as the golden standard for estimating causal effects.

In our contest is not possible for obvious ethical reasons to design a randomized experiment or even to find and exploit a so-called "natural experiment", i.e. a policy change that affects the poverty status of one group of parents but not another, and for which the change is uncorrelated with the unobserved characteristics of the families concerned.[3] Therefore, in our study, the critical problem of non-random treatment assignment (assumption (b) above) implies that additional assumptions have to be made in order to estimate the causal effects of the treatment. An important identifying assumption is the selection

[2] http://ec.europa.eu/eurostat/web/income-and-living-conditions/overview.

[3] See as an example the use of school reform in the literature of interegenerational transmission of education. Refer to Piopiunik (2014) and the literature cited therein.

on observables (unconfoundedness) (Rosenbaum and Rubin, 1983).[4]

Let us consider a set of N individuals, and denote each of them by subscript i: $i = 1, \ldots, N$. Let T_i indicate whether a child was growing up in a poor household, $T_i = 1$ (treated), or not, $T_i = 0$ (control). For each individual, we observe a vector of pre-treatment variables, X_i and the value of the outcome variable associated with the treatment, $Y_i(1)$ for being a poor child, $Y_i(0)$ for not being a poor child. The central assumption of our approach is that the "assignment to treatment" is unconfounded given the set of observable variables: $Y_i(0), Y_i(1) \perp T_i|X_i$. If the average treatment effect of interest is the "Average Treatment on the Treated" (ATT), the unconfoundedness assumption is then reduced to: $Y_i(0) \perp T_i|X_i$, where, within each cell defined by X, treatment assignment is random, and the outcome of controls are used to estimate the counterfactual outcome of treated in case of no treatment. Let $p(X)$ be the probability of growing in a poor household given the set of covariates X: $p(X) = Pr(T = 1|X = x) = E[T|X = x]$. Following Rosenbaum and Rubin (1983), treatment and potential outcomes are independent also conditional on p(X): $Y_i(0), Y_i(1) \perp T_i|p(X)$, thus, for a given propensity score value, exposure to treatment can be considered as random and thus poor and non poor children should be on average observationally identical. Therefore, we apply a propensity score matching method to select a control group of non-treated individuals (in this case non poor as a child) who are very similar to treated individuals conditional on a set of observable characteristics (parental characteristics, family composition, and other features fixed in childhood, such as the number of siblings or birth order) (unconfoundedness). The matched samples of poor and non-poor children will then be used to assess impacts on adulthood outcomes.

Formally, given the population of units i, if we know the propensity score $p(X_i)$, then the average effect of being poor on those exposed to poverty (ATT) can be written as follows:

$$\tau_t = E[Y_{1i} - Y_{0i}|T_i = 1] = E_{p(X_i)|T_i=1}[E[Y_{1i} - Y_{0i}|T_i = 1, p(X_i)]] =$$

$$E_{p(X_i)|T_i=1}[E[Y_{1i}|T_i = 1, p(X_i)]] - E_{p(X_i)|T_i=1}[E[Y_{0i}|T_i = 0, p(X_i)]]$$

As previously introduced, we analyze a mechanism behind this average effect. In order to do so we use a causal mediation analysis. The mediation analysis aims at quantifying the effect of a treatment that operates through a particular mechanism. In our study, this mechanism will be human capital accumulation.

Several other ways to conceptualize the mediatory role of an intermediate variable in the treatment - outcome relationship have been proposed in the causal inference literature. One of the most popular

[4]For a review of the statistical and econometric work focusing on estimating average treatment effects under this assumption, see Imbens (2004).

framework for identifying and estimating causal mechanisms is principal stratification (PRS).[5] PRS defines causal effects by comparing individuals with the same potential values of the post-treatment variable under each of the the treatment status (Frangakis and Rubin, 2002; Joffe et al., 2007). In particular, the use of PRS allows the introduction of direct versus indirect effects, analyzing the notion of causality when controlling for post-treatment variables (Mealli and Rubin, 2003; Rubin, 2004). Flores and Flores-Lagunes (2009) study more in detail the relationship of the concept of direct versus indirect effects with respect to the total average treatment effect, and they formally discuss the identification and estimation of causal mechanisms and net effects under different assumptions. More recently, Huber (2014) shows the identification of causal mechanism of a binary treatment variable under the unconfoundedness assumption, basing his estimation strategy on inverse weighting.

Acknowledging that in our specific context there might exist potential unobserved variables, that confound the outcome and mediator relationship even after controlling for a rich and high-quality set of information, as the one included in our EU-SILC data, in this paper we follow the procedure described in Imai et al. (2010), who also developed a method to assess the sensitivity of the estimated causal mediated effect to unobserved confounders. Formally, let $M_i(t)$ denote the potential value of the mediating variable for unit i with the treatment $T = t$, and let $Y_i(t, m)$ denote the potential outcome if $T = t$ and $M = m$. Under the framework of potential outcomes, the causal effect is the result of a comparison between the two potential results. This is considered as the basic problem of causal inference, and it is true also in mediation analyses, where the observed outcome $Y_i(T_i, M_i(T_i))$ depends on both the treatment status and the value of the mediator under the observed treatment level. Unlike the identification of the average treatment effect, identifying direct and indirect effects requires more stringent assumptions than random assignment. In this setting, an additional assumption is therefore required, the so-called sequential ignorability (SI):

$$Y_i(t', m), M_i(t) \perp T_i | X_i = x \tag{1}$$

$$Y_i(t', m) \perp M_i(t) | T_i = t, X_i = x \tag{2}$$

where $0 < Pr(M_i = m|T_i = t, X_i = x)$ and $0 < Pr(T = 1|X = x)$, for $t = 0$ and $t = 1$, is in the common support of X_i and M_i, respectively. Assumption 1 is the standard unconfoundedness assumption, where treatment assignment is assumed to be independent of potential outcomes and potential mediating variables, conditional on pre-tretament characteristics. Assumption 2 states that the mediator variable is now ignorable, given the observed treatment level and the observed characteristics,

[5] As we do not aim to completely review the vast literature on the topic, we refer here to some of the main works recently developed in the framework of Principal Stratification.

that is, among those individuals with the same poverty status and the same pretreatment observable characteristics, the level of education can be considered as if it were randomized. Given that such an assumption is not directly testable, we will thus provide sensitivity analyses to quantify the extent to which our empirical findings are robust to potential violations of the SI assumption. Our key quantity of interest is the change in the outcome (children outcomes later in life) corresponding to a change in the mediating variable from the level that would be observed under the control status $M_i(0)$ (higher education when not growing up in poverty), to the level that would be observed under the treatment status $M_i(1)$ (higher education when growing up in poverty), while holding the treatment variable constant at t (conditional on those growing up poor as a child)(Robins and Greenland, 1992).

In particular, we are interested in the average causal mediation effect (ACME) defined as: $\bar{\tau}_t = E[Y_i(t, M_i(1)) - Y_i(t, M_i(0))]$. Similarly, we can define the average direct effect (ADE) as follows: $\bar{\gamma}_t = E[Y_i(1, M_i(t)) - Y_i(0, M_i(t))]$.

When both the mediating variable and the outcome variable are continuous, the estimated mediation effect is equivalent to fitting two regressions:

$$M_i = \alpha_2 + \beta_2 T_i + \xi' X_i + \epsilon_{i2} \tag{3}$$

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi' X_i + \epsilon_{i3} \tag{4}$$

with the ACME equal to the product of the coefficient on the treatment variable, β_2 and the coefficient on the mediating variable γ as long as the linearity assumtpion is satisfied. However, when outcome variables are binary and mediating variables are continuous, or when the mediator is binary and a non linear model is applied, the product of β_2 and γ does not lead to the estimation of the ACME (Imai et al. (2010)). In our study the mediating variable is binary (achieving at least secondary education) and a probit model is used:

$$M_i = \mathbf{1}\{M_i^* > 0\}$$

where

$$M_i^* = \alpha_2 + \beta_2 T_i + \xi' X_i + \epsilon_{i2}$$

The outcome variable is continuous (equivalized disposable income) and a linear regression model is implemented:

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi' X_i + \epsilon_{i3}$$

The error terms are independently and identically distributed (iid) following a standard normal distribution and a normal distribution with $Var(\epsilon_{i3}) = \sigma_3^2$ for ϵ_{i2} and ϵ_{i2}, respectively:

$$\epsilon_{i2} \sim \mathcal{N}(0,1) \qquad \epsilon_{i3} \sim \mathcal{N}(0,\sigma_3^2)$$

and we assume a bivariate normal distribution of the error terms:

$$(\epsilon_{i2}, \epsilon_{i3}) \sim \mathcal{N}(0, \rho\sigma_3^2)$$

where ρ is the correlation between the two error terms. [6] The ACME is computed as the average difference in predicted disposable income under the treatment across the levels of high school graduation with and without having experienced poverty (Hicks and Tingley, 2011).

Following this approach, we provide in section 5 information on the extent to which a causal effect of growing up in poverty on later outcomes occurs together with a causal effect of growing up poor on the intermediate outcome human capital status and investigate how sensitive our results are to the existence of unobserved confounders.

3 Data

The analysis is based on data from the the European Union Statistics on Income and Living Conditions (EU-SILC), which provides comparable, cross-sectional data on income, poverty, social exclusion and living conditions in the European Union.[7] For the specific purpose of this paper we will use the module on intergenerational transmission of 2011, where retrospective questions about parental characteristics (such as education, age, occupation) referring to the period in which the interviewee was a young teenager (between the age of 14 and 16) were asked to each household member aged over 24 and less than 66.[8] We further restrict our sample to the individual in working age between 35 to 55 to maintain a degree of homogeneity in the period of the life cycle in which the outcomes of interests are measured.

More in details, our treatment variable is constructed based on the financial situation of the household (vary bad or bad), while the variables we used as pre-treatment are the country of residence and of birth, the gender, the year and quarter of birth, the number of adults in the household, number of persons in the household in work, number of siblings, family composition, year and country of birth, highest level of education, main activity, main occupation and citizenship of the father and of the mother and tenancy

[6]This will allow us to run a sensitivity analysis by deriving mediation effects as a function of ρ, estimating the ACME under a series of ρ values different from zero.

[7]Refer to chapter 2 in Atkinson et al. (ming) for a detailed description of this database.

[8]This module was also asked in 2005, but given that the questions related to our main treatment are not comparable and the 2011 module provides more background variables on the parents, we decided here to focus on this last one. For an analysis of the 2005 module refer to Bellani and Bia (ming).

status. The outcome in adulthood that we are interested in are the log of the equivalized income,[9] and the probability of being poor (defined as having an equivalized income lower than 60% of the median in his/her country in that year). As already mentioned at the end of Section 2, as *intermediate* outcome occurring after treatment, that has an impact on the final outcome of interest, we are analyzing the probability of having at least completed secondary education. The descriptive statistics and the T-test on the means of the variables used in the following analysis can be found in the Appendix A.1.[10]

4 Results

4.1 Impact of experiencing poverty on adult outcomes

As a first step in the analysis we estimate by means of a probit model each individual's propensity score, i.e. his/her probability to be poor in childhood given the observed characteristics introduced in the previous section. In our sample the propensity score has a mean value of .11 a median of .08 and a standard deviation of .1.[11] As already explained in Section 2, the propensity score is a balancing score (Rosenbaum and Rubin, 1983), that is, within strata with the same value of $p(X)$ the probability that $T = \{0, 1\}$ (being poor or not) does not depend on the value of X. This balancing property, combined with the unconfoundedness assumption, implies that, for a given propensity score, exposure to a treatment status is random and therefore treated and control units should be on average "similar" conditional on observable characteristics. As a result, to be effective, propensity score based methods should balance characteristics across treatment groups. The extent to which this has been achieved can be explored by comparing balance in the pre-treatment covariates before and after adjusting for the estimated propensity score (PS). Figure 1 provides the standardized bias (in percentage) for unmatched and matched units,[12] showing a huge improvement in the balancing property when adjusting for the PS, with a bias always around 0.

Another important requirement for identification is given by the common support, which ensures that for each treated unit there are control units with the same observables. In the matched sample, the

[9]We use equivalized disposable income that is the total income of a household, after tax and other deductions, that is available for spending or saving, divided by the number of household members converted into equalized adults; household members are equalized using the so-called modified OECD equivalence scale.

[10]We are aware that individuals may suffer from retrospective recollection bias. In this specific case, we believe that the type of question asked is less affected by this problem compared, for example, with a direct question on the level of income in the household.

[11]For the results of this first step refer to table A2.1 in the Appendix A.2. Note that as a robustness check we have been esimating this probability also using interaction terms and our average treatment effects were not affected by this change.

[12]The reduction of bias due to matching is computed as: $BR = 100(1 - \frac{B_M}{B_0})$ where B_M is the standardized bias after matching $B_M = \frac{100(\bar{x}_{MC} - \bar{x}_{MT})}{\sqrt{\frac{S^2_{MC} + S^2_{MT}}{2}}}$ and B_0 is the standardized bias before matching $B_0 = \frac{100(\bar{x}_{0C} - \bar{x}_{0T})}{\sqrt{\frac{S^2_{0C} + S^2_{0T}}{2}}}$, where subscript M denotes after matching, 0 denotes before matching.

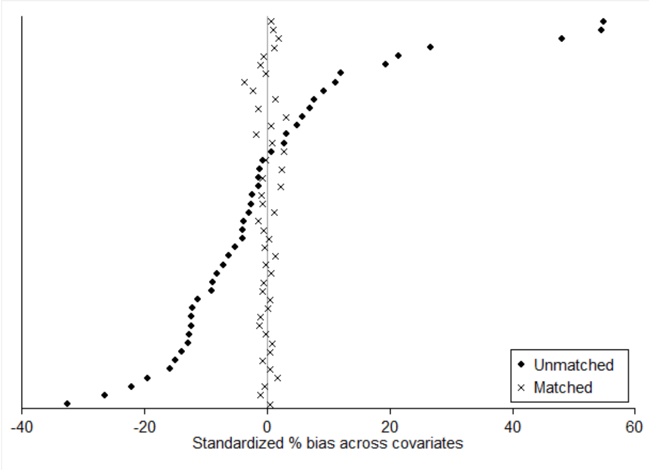


Figure 1: Standardized bias

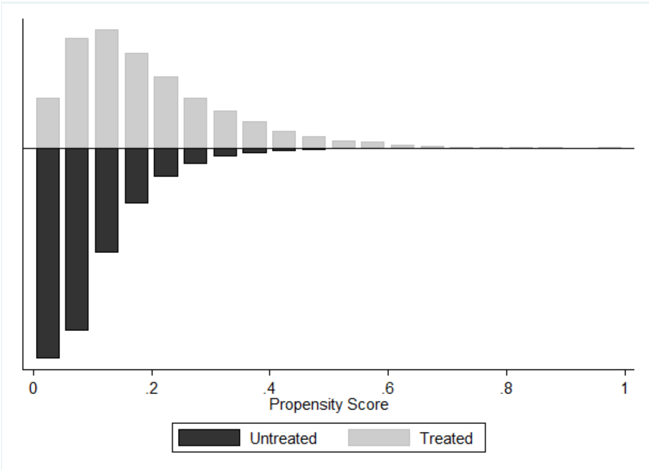


Figure 2: Common support

comparison of baseline covariates may be complemented by comparing the distribution of the estimated PS between treated and controls, as shown in figure 2. The common area of support serves as a first evaluation of whether covariates means included in the PS model are similar between the two groups (Ho et al., 2007). In our case, all the observations with non missing on the respective outcome variable belong to the common support.[13]

We implement the propensity score procedure applying a single nearest-neighbor matching to remove bias associated with differences in covariates[14] and estimate the effects of being poor on adulthood

[13]As shown in table A2.2 in Appendix A.2.

[14]To note that not only the mean value of the difference between the pscore of the treated and the pscore of his/her nearest neighbor is zero but also it is zero for the 99% of our data, reaching a maximum value at 0.03.

outcome, primarily equivalent income and probability of being poor later in life. [15]

In order to check the robustness of our results, we perform also a doubly-robust estimation of our treatment effect, which combines the entropy balancing method proposed by Hainmueller (2012) with a least squares (or probit) regression of the outcome on the treatment variable.

This balancing method constructs a weight for each control observation such that the sample moments of observed covariates are identical between the treatment and weighted control groups. We impose here balance on the second moments of the covariate distributions.[16] The marginal effect of the treatment is our doubly-robust estimate of the average treatment effect.[17]

Results are presented in table 1. The results on income show a substantial decrease in the equivalized income in adulthood due to exposure to poverty in childhood, of around 5%, and a significant increase of the probability of falling into poverty of 6 percentage points. With respect to our intermediate outcome of interest, our results show an average of 12 percentage points decrease in the probability of completing at least secondary education for the children experiencing financial difficulties while growing up. The average treatment effect estimate, relative to each outcome considered in our study, is overall very close between the two methods, both in the magnitude and in the significance.

Table 1: Average Treatment on the Treated

	Main Outputs		Intermediate Output
	Income	Poverty	Education
Propensity score matching	-0.05	0.07	-0.13
	(0.005)	(0.006)	(0.007)
Doubly-Robust Estimation	-0.05	0.06	-0.12
	(0.003)	(0.004)	(0.004)
N.	112746	112902	111833

The significant decrease in the probability of completing secondary education, double in magnitude of the impact on poverty risk, could be considered as a cause of the significant difference in income in adulthood that are maintained between these two groups of children. In the last part of the paper (section 5) we address this channel more in detail.

[15] In our analysis we use the `psmatch2` Stata package (Leuven and Sianesi, 2003).

[16] See table A2.3 in Appendix A.2 for the descriptive statistics of the sample pre and post reweighing.

[17] The results presented are the average marginal effect of the treatment variable, given by the OLS coefficient estimation for the continuous variable log of Income, and the average of the marginal effects calculated from the probit estimates for the probability of being poor in adulthood and to attain secondary education.

4.2 Sensitivity analysis

One of the central assumptions of our analysis is that being poor in childhood can be considered as random, given a set of precedent covariates X and thus the outcome of non-poor children can be used to estimate the counterfactual outcome of the poor children if they were not experiencing poverty in childhood. The plausibility of this assumption heavily relies on the quality and amount of information contained in X.[18] The validity of this assumption is not directly testable, since the data are completely uninformative about the distribution of the potential outcomes, but its credibility can be supported/rejected by additional sensitivity analysis. Our analysis would be in fact biased if we were to believe that even conditional on all the covariates we can observe (parental education and occupation, family situation, child own age, sex, year and country of birth and number of siblings, etc.), being poor in childhood would be linked to some unobserved parental genetic ability which would not only influence the probability of the parents of falling into poverty (being treated) but also the child's potential outcome as a result of the genetic transmission of ability. In this setting, it is assumed that the conditional independence assumption holds given X and the unobserved variable A: $Y_i(0) \perp T_i|X_i, A_i$ and knowing A would be sufficient to consistently estimate the ATT: $E[Y_{0_i}|T_i = 1, X_i, A_i] = E[Y_{0_i}|T_i = 0, X_i, A_i]$.

In order to asses if our estimated results are robust to the presence of such possible unobserved confouder, we implement the sensitivity analysis developed by Rosenbaum (2002), which relies on a sensitivity parameter - Γ - that measures the degree of departures from the assumption of random assignment of the treatment. The Γ parameter represents how much two individuals, with the same pre-treatment characteristics, may differ in the likelihood of receiving the treatment. In a randomized experiment, randomization ensures that $\Gamma = 1$. In observational studies, two individuals might be identical conditional on pre-treatment covariates, but one might be more likely to receive the treatment if they differ conditional on unobserved confounders (Rosenbaum, 2005) (e.g. if $\Gamma = 2$ one child might be twice as likely to be poor).

More formally, we have hidden bias if some subjects, j and i, have the same values on X, $X_j = X_i$, but different probabilities of receiving the treatment, $\pi_j \neq \pi_i$. Hence, if the probability of experiencing poverty for child i and child j is π_i, π_j, respectively, then the relative odds will be defined as: $\frac{\pi_i}{(1-\pi_i)}$ and $\frac{\pi_j}{(1-\pi_j)}$. The odds ratio of children with the same values on X is at most: $\frac{1}{\Gamma} \leq \frac{\frac{\pi_i}{(1-\pi_i)}}{\frac{\pi_j}{(1-\pi_j)}} \leq \Gamma$ for all i and j with $X_i = X_j$.

If A_i is the unobserved confounder for unit i, we can specify a logistic regression linking the odds to

[18] Refer to Section2 for a more detailed description of the assumptions made.

both the observed and unobserved covariates:

$$\log\left[\frac{\pi_i}{(1-\pi_i)}\right] = \beta X_i + \gamma A_i$$

with $0 \leq A_i \leq 1$. If child i and j have the same values on X, $X_i = X_j$, we can rewrite the model including the odds for units i and j in the following way:

$$\frac{\frac{\pi_i}{(1-\pi_i)}}{\frac{\pi_j}{(1-\pi_j)}} = \frac{e^{(\beta X_i + \gamma A_i)}}{e^{(\beta X_j + \gamma A_j)}} = e^{(\gamma(A_i - A_j))}.$$

That is, children differ in their odds of being poor by a factor of γ and the difference in the unobserved confounders.

The method developed by Rosenbaum (2002) includes different randomization tests according to the type of outcomes used in the analysis: Wilcoxon sign rank test (and the Hodges-Lehman point estimate) for continuous outcomes and the Mantel-Haenszel (MH) test for binary outcomes (see Becker and Caliendo (2007)).[19]

Results are presented in tables 2 and 3, for income and poverty, respectively.[20] Table 2 provides Rosenbaum's bounds for the p-values from Wilcoxon's signed rank test. We see that for an increase between 0.1 and 0.2 in Γ, the lower bound of the p-value increases to a level above the usual 0.05 threshold. This means that if the odds of one child being poor are more than 1.2 times higher because of different values on an unobserved covariate, e.g. parental ability, despite being identical on the matched covariates, our inference will not be significant anymore. If we then look at Rosenbaum's bounds for the additive effect due to treatment, i.e., the Hodges-Lehmann point estimate we see that the median difference in income if there is no hidden bias is -4.7%, for $\Gamma = 1.2$ this might be as high as -9.5% or as low as 0%.

Table 3 shows the results of the Mantel-Haenszel test for the binary outcome risk of poverty in adulthood. The two bounds in the table adjust the test statistic downward (upward) for the case of positive (negative) unobserved selection.

In our context if we were to believe that low able children, who are most likely to experience poverty as child because they are more likely to have low able parents, tend to have higher poverty rates in adulthood even without experiencing poverty growing up, a positive selection bias occurs, which in

[19]For binary outcomes, Aakvik (2001) suggests using the MH test statistic, that is, under the null-hypothesis of no treatment effect, the outcome distribution is hypergeometric. More formally, let $N_1 s$ and $N_0 s$ be the numbers of treated and control units in stratum s, where $N_s = N_0 s + N_1 s$. Then, let us define $Y_1 s$ and $Y_0 s$ the number of successful participants and non-participants, respectively, with $Y_s = Y_0 s + Y_1 s$. The Q_{MH} test statistics is given by: $Q_{MH} = \frac{|Y_1 - \sum_{s=1}^{S}(\frac{N_1 s Y_s}{N_s})| - 0.5}{\sqrt{\sum_{s=1}^{S}\frac{N_{1s} N_{0s} Y_s (N_s - Y_s)}{N_s^2 (N_s - 1)}}}$.

[20] In this part of our analysis we use the program `rbounds` provided by Gangl (2004), which allows to run sensitivity analyses for continuous outcomes, while to deal with binary outcomes we use the module built by Becker and Caliendo (2007) `mhbounds`.

Table 2: Rosenbaum bounds for Income

Γ	Wilcoxon's test significance level upper bound	lower bound	Hodges-Lehmann point estimate upper bound	lower bound
1	0	0	-.047	-.047
1.1	0	.000013	-.074	-.024
1.2	0	.465635	-.095	-.000
1.3	0	.99989	-.116	.020
1.4	0	1	-.135	.039
1.5	0	1	-.153	.057
1.6	0	1	-.169	.074
1.7	0	1	-.185	.089
1.8	0	1	-.2	.104
1.9	0	1	-.214	.118
2	0	1	-.227	.131

Γ =log odds of differential assignment due to unobserved factors

turns leads to an upward bias in the estimated treatment effects.

Table 3: Mantel-Haenszel bounds for Poverty

Γ	Test statistic over-estimation	under-estimation	Significance level over-estimation	under-estimation
1	12.1947	12.1947	0	0
1.1	9.23157	15.1726	0	0
1.2	6.53579	17.9064	3.2e-11	0
1.3	4.06164	20.4367	.000024	0
1.4	1.77401	22.7947	.038031	0
1.5	.322111	25.0046	.373684	0
1.6	2.3133	27.0861	.010353	0
1.7	4.18489	29.055	.000014	0
1.8	5.95135	30.9246	1.3e-09	0
1.9	7.62471	32.7057	1.2e-14	0
2	9.21505	34.4076	0	0

Γ = odds of differential assignment due to unobserved factors

The bounds under the assumption that we have over-estimated the treatment effect reveal that the confidence interval for the effect would include zero if an unobserved variable, e.g child ability, caused the probability of experiencing poverty in childhood to be 1.5 times higher for the treatment than the comparison groups, conditional on all the observable characteristics.

To analyze the extent of the possible overestimation, we also follow the approach suggested by Ichino et al. (2008) and assume that the unobserved ability variable A can be expressed as a binary variable taking value H=high, L=low. In addition, A is assumed to be i.i.d. distributed in the cells represented by the Cartesian product of the treatment and outcome values. The distribution of the binary confounding factor A can be fully characterized by the choice of four parameters: $p_{ij} \equiv Pr(A = 1|T = i, Y = j) = Pr(A = 1|T = i, Y = j, X)$ with $i, j \in 0, 1$, which give the probability that A = 1(high) in each of the four groups defined by the treatment status (poor as a child) and the outcome value (poverty in adulthood)[21]. Given arbitrary values of the parameters p_{ij}, a value of A is attributed to each individual, according to her/his belonging to one of the four groups defined by their poverty status in childhood and adulthood.

The simulated A is then treated as any other observed covariate and is included in the set of matching variables used to estimate the propensity score and to compute a simulated ATT estimate, derived as an average of the ATTs over the distribution of A.

We can thus control for the conditional association of A with Y_0 and T by measuring how each configuration of p_{ij} leads to an impact of A on Y_0 and T.

In order to do so, we estimate a logit model of $Pr(Y = 1|T = 0, A, X)$ at each iteration, reporting the average odds ratio of A as the "outcome effect" (Γ) and "selection effect" (Δ) of the simulated confounder:

$$\Gamma = \frac{\frac{Pr(Y=1|T=0,A=1,X)}{Pr(Y=0|T=0,A=1,X)}}{\frac{Pr(Y=1|T=0,A=0,X)}{Pr(Y=0|T=0,A=0,X)}},$$

i.e. the effect of parental ability on the outcome of non-poor children, controlling for the observable covariates (X),

$$\Delta = \frac{\frac{Pr(T=1|A=1,X)}{Pr(T=0|A=1,X)}}{\frac{Pr(T=1|A=0,X)}{Pr(T=0|A=0,X)}},$$

i.e. the effect of parental ability on the probability of experiencing poverty (T=1), controlling for the observable covariates (X).

We perform two simulation exercises.[22]In the first one, the p_{ij} are set so as to let our simulated parental ability A mimic the behavior of parental education variables, as their strong, although not perfect, positive correlation is well known in the literature (see among others Black et al. (2009); Anger and Heineck (2010); Björklund et al. (2010)).

In the second one, a set of different p_{ij} is built in order to capture the characteristics of this potential confounder that would drive the ATT estimates to zero (*Killer* confounder).

In tables 4 and 5 the results of these sensitivity checks are presented. Both the significance and the size of the estimated effects are robust to the introduction of the calibrated unobserved ability. Moreover

[21]Note that, in order to perform the simulation, two assumptions are made: i) binary confounder ii) conditional independence of A given X.

[22]For this sensitivity analysis we use the `sensatt` program developed by Nannicini (2007).

these results show that both the outcome and the selection effect need to be quite strong (4.8 and 4.9, respectively) in order to "kill" the ATT, i.e., to explain entirely the positive baseline estimate of the ATT. Finally, we can see that the effect of our calibrated parental ability on the probability of being poor as adults of non-poor children is of 0.2 points higher than the effect of parental ability on the probability of experiencing poverty in childhood.

Table 4: ATT estimation

Baseline	Father's Educ-Calibrated	Mother's Educ-Calibrated	Killer
0.06	0.04	0.04	0.00
(0.004)	(0.007)	(0.007)	(0.008)

Table 5: Sensitivity Analysis: p_{ij} values and odds ratio

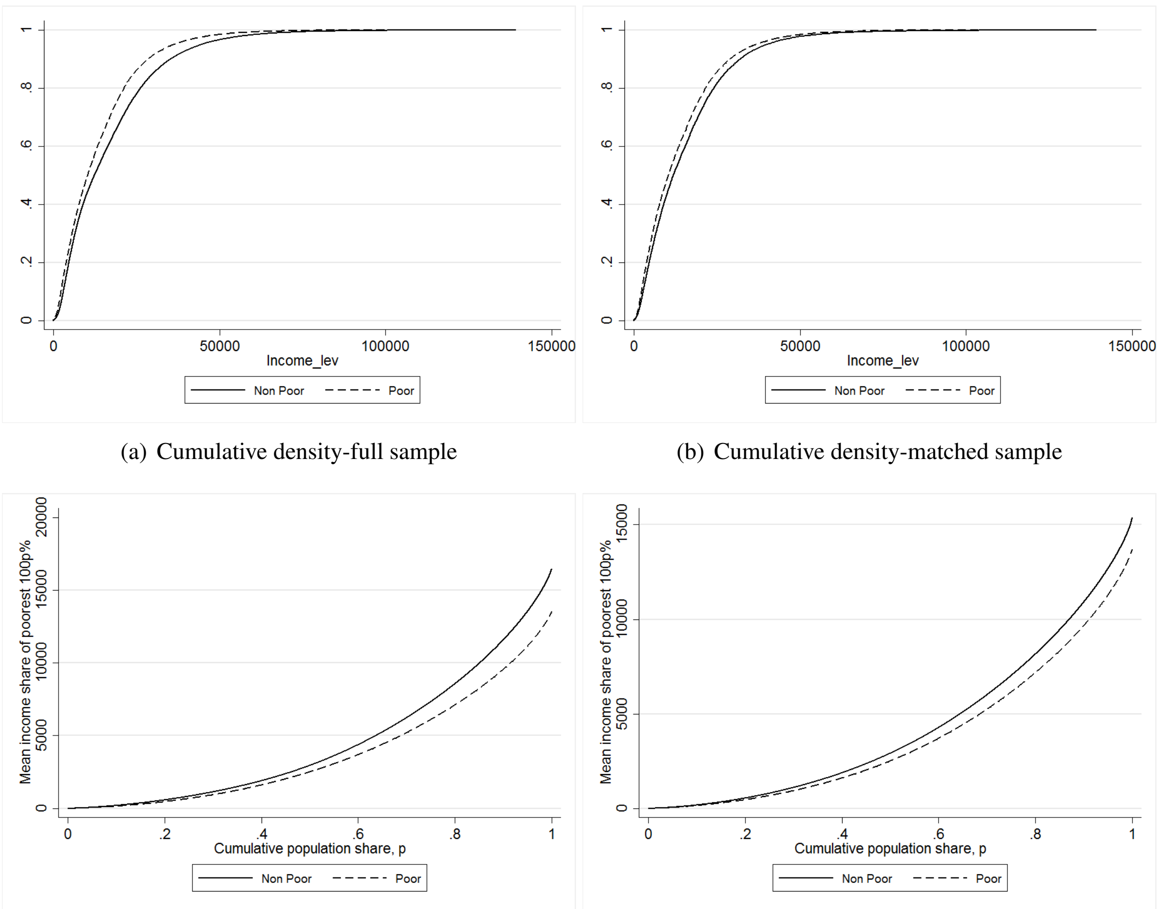
	Father's Educ-Calibrated	Mother's Educ-Calibrated	Killer
p_{11}	0.1	0.1	0.4
p_{10}	0.2	0.1	0.3
p_{01}	0.2	0.2	0.2
p_{00}	0.4	0.3	0.1
Outcome Effect	0.5	0.5	4.8
Selection Effect	0.3	0.2	4.9

To summarize, over the 27 European countries considered in our study, growing up poor as a child leads to an increase of 4 percentage points in the risk of being poor later in life, in our most conservative estimates, and to a decrease of 5% in the adult equivalent income. The impacts of experiencing poverty in childhood on the two adult outcomes would remain significant even if parental unobservable ability alone, conditional on all the other observable parental characteristics, were to increase the probability of experiencing financial problems of about 15% and 30%, respectively. This suggests that a significantly greater unobserved difference in the covariates should occur in the second case for our inference at the income level to be changed.

Finally, when following the approach of Ichino et al. (2008), our findings are robust to the introduction of both "calibrated" and "killer" confounders, needing the odds of experiencing poverty in childhood to be 4.9 times higher for high ability parents than for low ability ones and the odds of being at risk of poverty in adulthood 4.8 times higher for non-poor children of high ability parents than the ones of low ability parents, controlling for the observable covariates, to make our results invalid.

4.3 Distributional effect

Following the results presented in column 1 of table 1 we know that experiencing poverty while growing up leads to a decrease of income in adulthood of on average 5%. In this subsection we are interested in exploring this impact more in detail. In order to do so we present and compare here the cumulative densities of the distributions of income in adulthood of the children belonging to different groups. At first, we consider the whole sample and compare the distribution of individuals' income in both the poor and non-poor group of children without controlling for their probability of experiencing poverty in childhood. Looking at part (a) and (c) in figure 3 we can notice that not surprisingly the distribution of the non-poor children first order stochastic dominates the other, implying thus higher social welfare in an hypothetical society in which no one experience poverty as a child than one in which childhood poverty is common.



(a) Cumulative density-full sample

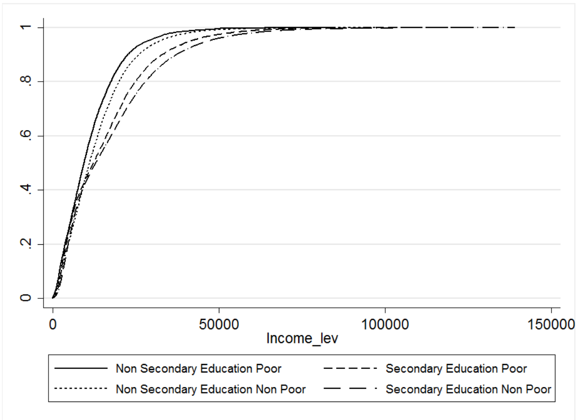
(b) Cumulative density-matched sample

(c) Generalized Lorenz Curves-full sample

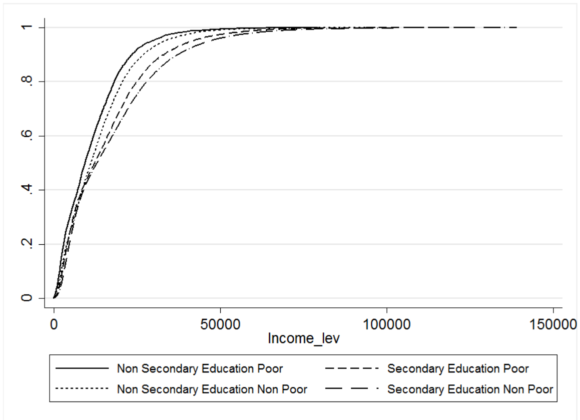
(d) Generalized Lorenz Curves-matched sample

Figure 3: Distributional Graphs

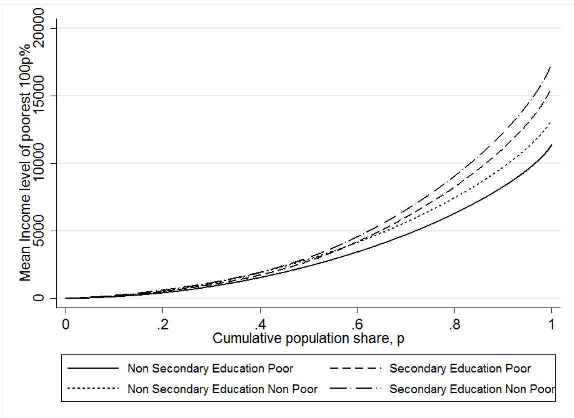
More interestingly, when we look at the matched sample, i.e., the sample where the characteristics of the children are matched such as to not significantly differ between the poor and non poor [23], this result still holds, suggesting that even when we control for the observable characteristics which are associated with experiencing poverty in the first place, the impact of the parental financial difficulties, does predict lower welfare achievement in the next generation[24] (see part (b) and (d) in figure 3).



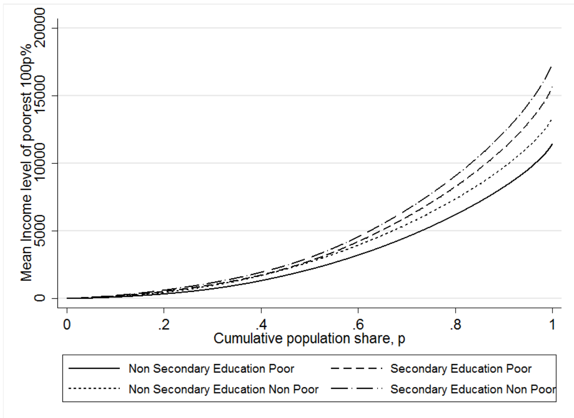
(a) Cumulative density-full sample



(b) Cumulative density-matched sample



(c) Generalized Lorenz Curves-full sample



(d) Generalized Lorenz Curves-matched sample

Figure 4: Distributional Graphs-Education

To conclude this part, before moving to the analysis of the role of education, we show in figure 4 the distribution of incomes of our sample divided in four groups, characterized not only by growing up poor as a child, but also by their own subsequent educational choice. As we can notice from both the full and the matched sample, acquiring at least secondary education plays as an important discriminate in being able to reach higher levels of income along the whole distribution. It seems in fact to have higher impact that not being poor in childhood along the whole distribution when we control for other differences in observable characteristics between the group of poor and non poor children (see part (b) and (d) of

[23]See table A2.4 in Appendix A.2 for a T-test on the means in this matched sample

[24]In Figure A2.1 in Appendix A.2 the bottom and the top of the distribution is plotted for both cases to show that the cumulative densities are not crossing.

figure 4) and for the upper 40% of the distribution when we do not control(part (a) and (c) of figure 4). Moreover we can notice that even acquiring at least a secondary level of education is not completely offsetting the lower income reached along the whole distribution for the one who experienced poverty in childhood.

5 The role of education

After having shown that childhood poverty has indeed a relevant detrimental impact on economic outcomes in adulthood, a fundamental question left to answer to be able to make more effective policy recommendation regards the channels through which being raised in a poor family affects the individual's economic and social status as an adult. In order to do so, as a first step in this direction, this section provides evidence of the important role that human capital has in this process.

As introduced at the end of section 2, in this last part of our analysis we implement a causal mediation analysis to our research question in order to uncover the role of human capital accumulation in the intergenerational poverty transmission. In particular, we study whether being poor as a child led to substantial lower levels of income later in life by decreasing the likelihood of graduating from High School. Our mediating variable is therefore secondary education, while the outcome of interest is the income level in adulthood. First, based on the fitted mediator model (eq.3, replaced by a probit model), we generate predicted secondary education attainment levels for the children who experienced and not experienced poverty. Next, we use the outcome model (eq.4) to impute potential outcomes. The ACME, is computed as the average difference in disposable income under the treatment across the levels of high school graduation with and without having experienced poverty. Finally, we repeat the two simulation steps 500 times in order to estimate the standard errors. We first estimate the ACME and the ADE assuming a bivariate normal distribution of the error terms, with mean 0 and covariance $\rho\sigma_3$, where ρ is the correlation between the two error terms.

Table 6 shows the estimated ACME, ADE and average total effect. The total effect, as we knew from the analysis in the previous sections, is equal to -5%. The indirect effect (ACME), i.e. the portion of the total effect, conveyed through the educational level, is estimated to be around -2%, suggesting that a significant portion of the average total effect is attributable to a decrease in the probability of graduating from High School. Hence, growing up poor as a child "induces" a decreased level of education which accounts for 34% of the total effect on adult income.

While education is a major contributor, the part of the transmission mechanisms that remains unexplained is substantial. As a first possible explanation for it we should remind the reader that due to data availability our measure of education is quite imprecise, not allowing to distinguish between different quality of this same education. Moreover, the extent to which poverty status is transmitted from parents to their children also depends on the combined effect of the investment in education and the rate of return on these investments. The extent to which education is publicly financed and rewarded in the labor market also matters and it is in turn affected by the way both the society and the market operate in the environment where the children are raised.

To conclude, focusing on possible others unobservable factor which could explain the remaining part of the this impact, it is worth highlighting that parental poverty is likely to be related to lower levels of good health, nutrition and housing, all of which affect child development and thus future incomes. Furthermore, the home and social environment is where beliefs, attitudes and values are shaped and these are likely to have effects on children future attitudes to work, health and family formation.

Table 6: Mediation Causal Analysis

	Direct Indirect Effects	95% Confidence Intervals	
ACME	-.017	-.019	-.016
Direct Effect	-.033	-.039	-.027
Total Effect	-.051	-.057	-.044
% Tot.Eff. Mediated	34	30	39

5.1 Sensitivity Analysis

As previously mentioned in section 2, both the assumptions 1 and 2 rely on the the quality and richness of the data. Despite the rich set of pretreatment characteristics, if there were unobserved confounders that affect both the educational level and the income, the SI assumption will no longer be satisfied, and the ACME and ADE will not be identified. As an example, pre-existing cognitive or non-cognitive problems might reduce the likelihood of graduating from secondary school, as well as the likelihood of higher income levels later in life. In order to deal with this hypothetical violation of the SI assumption, we assess the role of unobserved confounders via a sensitivity analysis.

We first apply the analysis based on the estimated ρ parameter via the quasi-Bayesian technique and report the indirect effects as a function of ρ. When ρ is 0, assumption 2 is satisfied, thus, there is no correlation between the error terms of the mediator and outcome models. Conversely, values of ρ different from 0 lead to violations of the SI (Keele et al., 2015).

In our study, ACME equals 0 for $\rho = 0.3$ (see Figure 5), that means that if there were a modest violation of the SI, the true mediated effect could be 0.

As the sensitivity parameter itself is rather difficult to interpret directly, we show here an alternative approach which is to express the degree of sensitivity as a function of R^2, that is the usual coefficient

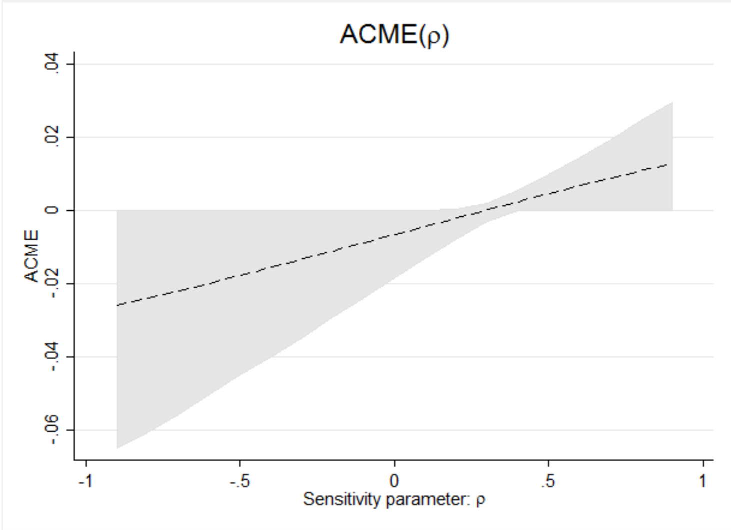


Figure 5: Sensitivity Analysis

of determination. In presence of an omitted confounder U_i, the error term will be a function of U_i and will be equal to $\epsilon_{ij} = \lambda_j U_i + \epsilon'_{ij}$, with $j = 2, 3$ for the mediator and outcome model, respectively, and λ_j representing the unknown coefficient for each equation. The sensitivity analysis is based on the proportion of original variance that is explained by the omitted confounder in the mediator and outcome model, equal to $\tilde{R}^2_M \equiv [Var(\epsilon_{i2}) - Var(\epsilon'_{i2})]/Var(M_i)$ and $\tilde{R}^2_Y \equiv [Var(\epsilon_{i3}) - Var(\epsilon'_{i3})]/Var(Y_i)$. In this setting, ρ is a function of the unexplained variances proportions in the mediator and outcome models.[25] The relationship between the ACME and the R^2 can be expressed as the product of the mediating and outcome variables' R^2, with $\rho = sgn(\lambda_2\lambda_3)R^*_M R^*_Y$ for the unexplained variances, and $\rho = sgn(\lambda_2\lambda_3)\tilde{R}_M\tilde{R}_Y/\sqrt{(1-R^2_M)(1-R^2_Y)}$ for the original variances (Hicks and Tingley, 2011). [26] To represent the results of this sensitivity analysis we show how much of the observed variations in the mediating ($\tilde{R}^2_M$) and outcome ($\tilde{R}^2_Y$) variables are explained by a potential unobserved confounder. In Figure 6 these proportions are reported on the horizontal and vertical axes, respectively. The dark line represents the combination of explained variations for which the ACME is equal 0. In particular, the true ACME would change sign when the product of the proportions is greater than 0.2. For example, we might think that preexisting cognitive and non-cognitive problems will turn into a decrease of both graduation rates and income level in adulthood. In this case, the true ACME would be 0 if these problems explained around 14% of the variances for both of these variables. At higher values of both $\tilde{R}^2_Y$ and $\tilde{R}^2_M$, the estimated causal mediation effect would be positive. For example, the unobserved

[25] $R^{2*}_M \equiv 1 - Var(\epsilon'_{i2})/Var(\epsilon_{i2})$, $R^{2*}_Y \equiv 1 - Var(\epsilon'_{i3})/Var(\epsilon_{i3})$

[26] When the mediating or outcome variable is binary, the pseudo-R^2 developed by McKelvey and Zavoina (1975) is implemented.

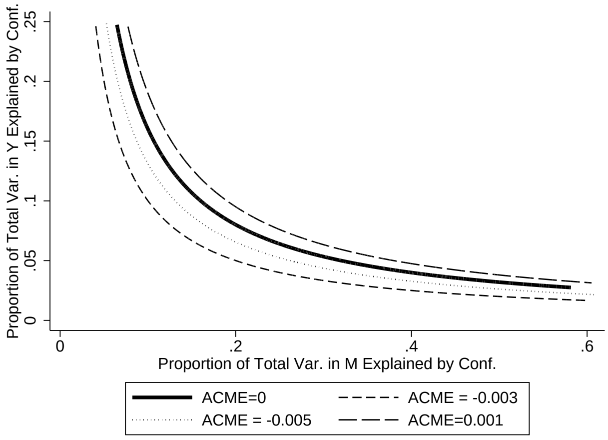


Figure 6: Sensitivity Analysis

factors should explain around 62% of the original variance in the secondary education and around 33% of the original variance in the income level later in life for an ACME equal to just 0.01.

To sum up, as reported in table 7 both $\tilde{R}^2_Y$ and $\tilde{R}^2_M$ must be substantially higher for the original conclusion to be changed, showing that negative mediation effects for the equivalized adult income is quite robust to deviations from the standard SI assumption of no unobserved confounding factors.

Table 7: Sensitivity results for the ACME

ρ at which ACME = 0:	0.3
$R^{2*}_M R^{2*}_Y$ at which ACME = 0:	0.09
$\tilde{R^2_M}\tilde{R^2_Y}$ at which ACME = 0:	0.02

6 Concluding remarks

This paper examines the causal channels through which growing up poor affects the individual's economic outcomes as an adult. We employ a propensity score matching method under the assumption that, conditional on observable characteristics, growing up in poverty is independent of the income level and the probability of being poor later in life. We also perform a doubly-robust estimation of our treatment effect, implementing the entropy balancing method with a least squares (or probit) regression of our outcomes of interest on experiencing financial problems in childhood. The richness of our data and a series of sensitivity analyses performed to check for unobserved confounders, e.g. genetically transmitted ability, augment the credibility of our identifying assumptions.

Our analysis is based on the 2011 module on intergenerational transmission of EU-SILC data, where retrospective questions on parental characteristics were asked to each household member older than 24 and younger than 66. For homogeneity purposes over the life cycle period, we restrict our sample to working age between 35 and 55.

Our results show that, on average, over the 27 European countries considered, growing up in financial distress leads to an increase of 4 percentage points in the risk of being poor, in our most conservative estimates, and to a decrease of 5% in the adult equivalent income, which would remain significant even if parental unobservable ability alone, conditional on all the others observable parental characteristics (e.g. age, occupation, education, migration status), were to increase their probability of experiencing financial problems of 15% (in our most restrictive case).

We also investigate the average effect on income later in life more in detail, looking at the cumulative densities of the distributions of children income in adulthood belonging to different groups. We find that, even after controlling for the probability of growing up poor as a child, the distribution of the non-poor children first order stochastic dominates the one of the poor children, implying higher social welfare in an hypothetical society where no one experiences poverty as a child against one where childhood poverty is common. This finding remains valid even after achieving (at least) a secondary level of education.

Moreover, we find that experiencing poverty during childhood will more likely translate into an exclusion from secondary education (of 12 percentage points on average) and that education plays indeed a substantial role in terms of intermediate variable to define the causal effect estimation of childhood poverty on labor market outcomes in adulthood, accounting for almost 35% of the total effect.

Some policy implications can be derived by our study. On the one hand, our results are in line with some previous studies which suggest that progressive government spending on education can increase intergenerational mobility, offsetting parental sub-optimal investment in education (Solon, 2004; Mayer and Lopoo, 2008). On the other hand, our analysis reinforces the need for policies devoted to eliminate the source at the basis of this increased risk, reducing childhood poverty.

Our research also shows the need for further studies devoted to the analysis of the other factors driving the impact of childhood poverty on adult incomes. As suggested in our discussion parental poverty is likely to be related to lower levels of health and housing, which affect not only children cognitive but also their non-cognitive development as beliefs, attitudes and values are also shaped in those years. Thus, helping parents work can be more effective than giving them cash transfers, as this may contribute to change attitudes or behaviors. Reducing the stress and anxiety of children, targeting intensive health, nutrition and care supports on particularly deprived households or areas is highly desirable.

References

Aakvik, A. (2001). Bounding a Matching Estimator: The Case of a Norwegian Training Program. *Oxford Bulletin of Economics and Statistics*, 63(1):115–43.

Acemoglu, D. and Pischke, J.-S. (2001). Changes in the wage structure, family income and childrens education. *European Economic Review*, 45:890–904.

Anger, S. and Heineck, G. (2010). Do smart parents raise smart children? the intergeneretional transmission of cognitive abilities. *Journal of Population Economics*, 23:1255–1282.

Atkinson, T., Guio, A.-C., and Marlier, E., editors (Forthcoming). *Monitoring Social Europe*. Eurostat Statistical.

Becker, S. O. and Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *Stata Journal*, 7(1):71–83.

Bellani, L. and Bia, M. (Forthcoming). The impact of growing up poor in europe. In Atkinson, T., Guio, A.-C., and Marlier, E., editors, *Monitoring Social Europe*. Eurostat Statistical.

Björklund, A., Eriksson, K. H., and Jäntti, M. (2010). Iq and family background: Are associations strong or weak? *The B.E. Journal of Economic Analysis & Policy*, 10.

Black, S. E., Devereux, P. J., and Salvances, K. G. (2009). Like father, like son? a note on the intergenerational transmission of iq scores. *Economic Letters*, 105:138–140.

Blanden, J., Gregg, P., and Macmillan, L. (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal*, 117:C43C60.

Ermisch, J., Francesconi, M., and Pevalin, D. J. (2004). Parental partnership and joblessness in childhood and their influence on young peoples outcomes. *Journal of the Royal Statistical Society*, 167:69101.

Flores, C. and Flores-Lagunes, A. (2009). Identification and estimation of causal mechanisms and net effects of a treatment under unconfoundedness. DP 4237, IZA.

Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58(1):21–29.

Gangl, M. (2004). RBOUNDS: Stata module to perform Rosenbaum sensitivity analysis for average treatment effects on the treated. Statistical Software Components, Boston College Department of Economics.

Grundiza, S. and Lopez Vilaplana, C. (2013). Is the likelihood of poverty inherited? Statistics in focus 27/2013, Eurostat.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20:25–46.

Hicks, R. and Tingley, D. (2011). Causal mediation analysis. *The Stata Journal*, 11:605–619.

Ho, D., , K., King, G., and Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15.

Huber, M. (2014). Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, 29(6):920–943.

Ichino, A., Mealli, F., and Nannicini, T. (2008). From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, 23(3):305–327.

Imai, K. L., Keele, L., and Yamamoto, T. (2010). Identification, inference and, sensitivity analysis for causal mediation effects. *Statistical Science*, 25:309–334.

Imbens, G. (2004). Non parametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics*, 86:4–29.

Joffe, M., Small, D., and Hsu, C.-Y. (2007). Defining and estimating intervention effects for groups that will develop an auxiliary outcome. *Statistical Science*, 22:74–97.

Keele, L., Tingley, D., and Yamamoto, T. (2015). Identifying mechanism behind policy interventions via causal mediation analysis. *Journal of Policy Analysis and Management*, 34:937–963.

Leuven, E. and Sianesi, B. (2003). Psmatch2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components, Boston College Department of Economics.

Lopez Vilaplana, C. (2013). Children were the age group at the highest risk of poverty or social exclusion in 2011. Statistics in focus 4/2013, Eurostat.

Mayer, S. and Lopoo, L. (2008). Government spending and intergenerational mobility. *Journal of Public Economics*, 92.

Mayer, S. E. (1997). *What Money Cant Buy. Family Income and Childrens Life Chances*. Harvard University Press, Cambridge MA.

McKelvey, R. and Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 4(3):103–120.

Mealli, F. and Rubin, D. B. (2003). Assumptions allowing the estimation of direct causal effects. *Journal of Econometrics*, 112(1):79–87.

Nannicini, T. (2007). Simulation-based sensitivity analysis for matching estimators. *The Stata Journal*, 7(3):334–350.

Piopiunik (2014). Intergenerational transmission of education and mediating channels: evidence from a compulsory schooling reform in germany. *The Scandinavian Journal of Economics*, 116:878–907.

Ratcliffe, C. (2015). Child poverty and adult success. Brief, Urban Institute.

Robins, J. M. and Greenland, S. (1992). Indentifiability and exchangeability for direct and indirect effects. *Epidemiology*, 3:143–155.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer-Verlag New York.

Rosenbaum, P. R. (2005). Sensitivity analysis in observational studies. In Everitt, B. S. and Howell, D. C., editors, *Encyclopedia of Statistics in Behavioral Science*. John Wiley & Sons, Ltd, Chichester,.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Education Psychology*, 66.

Rubin, D. (1978). Estimating causal effects of treatments in randomized and nonrandomized studies. *Annals of Statistics*, 6:3458.

Rubin, D. B. (2004). Direct and indirect causal effects via potential outcomes. *Scandinavian Journal of Statistics*, 31(2):161–170.

Shea, J. (2000). Does parents money matter? *Journal of Public Economics*, 77:155–184.

Solon, G. (2004). A model of intergenerational mobility variation over time and place. In Corak, M., editor, *Generational Income Mobility in North America and Europe*, pages 38–47. Cambridge University Press,Cambridge, England.

A Appendix

A.1 Descriptive statistics

Table A1.1: Descriptive Statistics

	mean	sd	min	max
quarter of birth	2.45	1.11	1	4
year of birth	1965.7	5.98	1956	1976
sex	1.53	0.50	1	2
n. of adult in hh	2.56	1.16	0	67
n. of children in hh	2.41	1.43	0	41
n. of person in work	1.91	1.09	0	30
year of birth of father	1935.6	9.24	1890	1981
year of birth of mother	1938.7	8.78	1841	1988
country==AT	0.036	0.19	0	1
country==BE	0.030	0.17	0	1
country==BG	0.035	0.18	0	1
country==CH	0.036	0.19	0	1
country==CY	0.027	0.16	0	1
country==CZ	0.034	0.18	0	1
country==DK	0.014	0.12	0	1
country==EE	0.024	0.15	0	1
country==EL	0.026	0.16	0	1
country==ES	0.085	0.28	0	1
country==FI	0.016	0.12	0	1
country==FR	0.055	0.23	0	1
country==HU	0.069	0.25	0	1
country==IS	0.0084	0.091	0	1
country==IT	0.12	0.32	0	1
country==LT	0.027	0.16	0	1
country==LU	0.037	0.19	0	1
country==LV	0.028	0.17	0	1
country==MT	0.025	0.15	0	1
country==NL	0.030	0.17	0	1
country==NO	0.014	0.12	0	1
country==PL	0.076	0.26	0	1
country==PT	0.032	0.17	0	1
country==RO	0.037	0.19	0	1
country==SE	0.011	0.10	0	1
country==SK	0.037	0.19	0	1
country==UK	0.034	0.18	0	1
Lived with parent(s)	0.99	0.090	0	1
country_of_birth==EU	0.041	0.20	0	1
country_of_birth==OTH	0.054	0.23	0	1
father_country_of_birth==country of residence	0.88	0.32	0	1
mother_country_of_birth==country of residence	0.88	0.32	0	1
father primary education	0.65	0.48	0	1
mother primary education	0.71	0.45	0	1
Tenancy_status== Owner	0.72	0.45	0	1
father Employed	0.98	0.14	0	1
mother Employed	0.59	0.49	0	1
Poor as child	0.11	0.31	0	1
Income	4.03	0.43	0.40	5.14
Poor as adult	0.15	0.35	0	1
At least secondary education	0.77	0.42	0	1
Observations	114278			

Table A1.2: T-test on the pre-treatment variables means by treatment status

	Non poor	Poor	Diff
quarter of birth	2.46	2.45	0.01
year of birth	1965.88	1964.59	1.28***
sex	1.53	1.52	0.00
n. of adult in hh	2.53	2.82	-0.30***
n. of children in hh	2.32	3.13	-0.81***
n. of person in work	1.91	1.91	-0.01
year of birth of father	1935.81	1933.97	1.84***
year of birth of mother	1938.93	1937.07	1.86***
country==AT	0.03	0.05	-0.02***
country==BE	0.03	0.02	0.01***
country==BG	0.04	0.01	0.02***
country==CH	0.04	0.03	0.01***
country==CY	0.02	0.06	-0.04***
country==CZ	0.04	0.02	0.01***
country==DK	0.01	0.01	0.01***
country==EE	0.03	0.01	0.02***
country==EL	0.03	0.04	-0.01***
country==ES	0.08	0.10	-0.02***
country==FI	0.02	0.01	0.01***
country==FR	0.05	0.05	0.00
country==HU	0.07	0.07	0.00
country==IS	0.01	0.01	0.00**
country==IT	0.12	0.12	0.00
country==LT	0.03	0.02	0.00**
country==LU	0.04	0.04	-0.01**
country==LV	0.03	0.01	0.02***
country==MT	0.02	0.03	-0.01***
country==NL	0.03	0.01	0.02***
country==NO	0.02	0.00	0.01***
country==PL	0.07	0.07	0.01*
country==PT	0.03	0.08	-0.06***
country==RO	0.03	0.06	-0.02***
country==SE	0.01	0.01	0.00***
country==SK	0.04	0.02	0.02***
country==UK	0.03	0.03	0.01***
Lived with parent(s)	0.99	0.99	0.01***
country_of_birth==EU	0.04	0.05	-0.01***
country_of_birth==OTH	0.05	0.08	-0.03***
father_country_of_birth==country of residence	0.89	0.84	0.04***
mother_country_of_birth==country of residence	0.89	0.85	0.04***
father primary education	0.62	0.85	-0.23***
mother primary education	0.69	0.90	-0.21***
Tenancy_status== Owner	0.74	0.59	0.15***
father Employed	0.98	0.95	0.04***
mother Employed	0.60	0.47	0.13***

Table A1.3: T-test on the post-treatment variables means by treatment status

	Non poor	Poor	Diff
Income	4.04	3.96	0.09***
Poor as adult	0.13	0.24	-0.10***
At least secondary education	0.80	0.55	0.25***

A.2 Additional Tables and Figures

Table A2.1: Probit

	Poor as child	
Poor as child		
quarter of birth	-0.000555	(-0.11)
year of birth	-0.00828***	(-8.93)
sex	-0.00766	(-0.71)
n. of adult in hh	0.0647***	(13.98)
n. of children in hh	0.140***	(41.99)
n. of person in work	-0.0352***	(-6.84)
year of birth of father	-0.000129***	(-6.32)
year of birth of mother	-0.00000220	(-0.10)
country==AT	0.439***	(10.01)
country==BE	-0.0231	(-0.47)
country==BG	-0.0959	(-1.83)
country==CH	0.139**	(2.95)
country==CY	0.596***	(13.32)
country==CZ	-0.0503	(-1.04)
country==DK	0.0755	(1.15)
country==EE	-0.0930	(-1.62)
country==EL	0.393***	(8.52)
country==ES	0.112**	(2.76)
country==FI	-0.0146	(-0.23)
country==FR	0.0675	(1.56)
country==HU	0.234***	(5.61)
country==IS	0.222**	(3.13)
country==IT	-0.00696	(-0.18)
country==LT	0.0604	(1.22)
country==LU	0.0481	(1.02)
country==LV	-0.262***	(-4.81)
country==MT	0.0482	(1.06)
country==NL	-0.338***	(-6.62)
country==NO	-0.0264	(-0.35)
country==PL	0.170***	(4.12)
country==PT	0.620***	(14.39)
country==RO	0.383***	(8.79)
country==SE	0.250***	(3.73)
country==SK	0.103*	(2.14)
Lived with parent(s)	-0.268***	(-4.51)
country_of_birth==EU	0.0725*	(2.04)
country_of_birth==OTH	0.0288	(0.88)
father_country_of_birth==country of residence	-0.112***	(-3.57)
mother_country_of_birth==country of residence	-0.0336	(-1.05)
father primary education	0.353***	(21.71)
mother primary education	0.313***	(17.26)
Tenancy_status== Owner	-0.391***	(-33.57)
father Employed	-0.599***	(-19.94)
mother Employed	-0.0544***	(-4.20)
Constant	15.43***	(8.45)
Observations	112902	

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2.2: Common Support

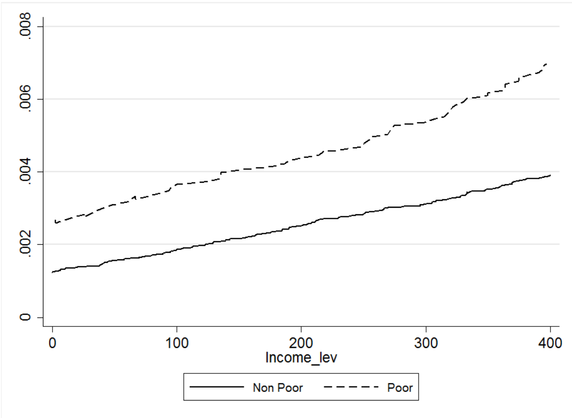
	Income		Poverty		Education	
Treatment assignment	On Support	Total	On Support	Total	On Support	Total
Untreated	100,616	100,616	100,741	100,741	99,996	99,996
Treated	12,130	12,130	12,161	12,161	11,837	11,837
Total	112,746	112,746	112,902	112,902	111,833	111,833

Table A2.3: Balancing Table

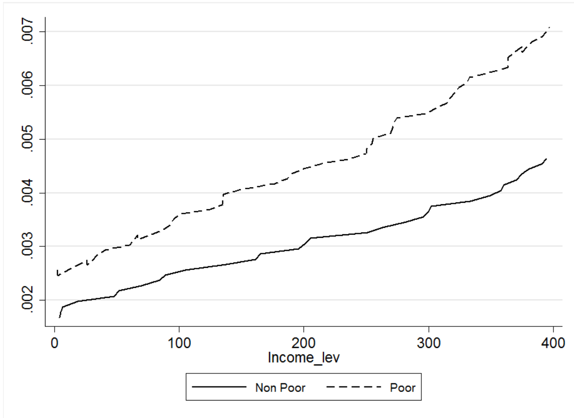
	Mean			Variance			Skewness			Std difference	
	treated	control	weighted cont	treated	control	weighted cont	treated	control	weighted cont	treat&control	treat& weighted cont
quarter of birth	2.2	2.1	2.2	1.9	2.1	1.9	-.51	-.55	-.52	.052	2.8e-05
year of birth	1965	1966	1965	35	36	35	.28	.019	.28	-.22	.0032
sex	1.5	1.5	1.5	.25	.25	.25	-.095	-.11	-.095	-.0064	3.1e-05
n. of adult in hh	2.8	2.5	2.8	2.5	1.2	2.5	7	3.5	4.7	.18	3.1e-05
n. of children in hh	3.1	2.3	3.1	3.8	1.8	3.9	1.6	2.1	2	.41	3.0e-05
n. of person in work	1.9	1.9	1.9	2.1	1.1	2.1	3.7	3.2	5	.0074	-7.3e-05
year of birth of father	1799	1867	1799	2.4e+05	1.3e+05	2.4e+05	-3.4	-5	-3.4	-.14	3.5e-07
year of birth of mother	1828	1883	1828	2.0e+05	1.1e+05	2.0e+05	-3.8	-5.6	-3.8	-.12	7.2e-06
country==AT	.053	.034	.053	.05	.033	.05	4	5.2	4	.085	-1.1e-05
country==BE	.024	.03	.024	.023	.03	.023	6.2	5.5	6.2	-.042	-8.7e-06
country==BG	.014	.038	.014	.014	.036	.014	8.2	4.8	8.2	-.2	-1.9e-04
country==CH	.03	.037	.03	.029	.035	.029	5.5	4.9	5.5	-.041	-3.6e-05
country==CY	.061	.023	.061	.057	.022	.057	3.7	6.4	3.7	.16	3.7e-05
country==CZ	.021	.036	.021	.021	.035	.021	6.6	5	6.6	-.1	-9.7e-06
country==DK	.0072	.015	.0072	.0072	.014	.0072	12	8.1	12	-.087	-4.4e-05
country==EE	.01	.026	.01	.0099	.026	.0099	9.8	5.9	9.8	-.17	-1.7e-04
country==EL	.039	.025	.039	.037	.025	.037	4.8	6	4.8	.07	3.2e-05
country==ES	.1	.084	.1	.09	.077	.09	2.7	3	2.7	.055	5.9e-05
country==FI	.0077	.017	.0077	.0077	.016	.0077	11	7.6	11	-.1	-4.6e-05
country==FR	.05	.053	.05	.048	.051	.048	4.1	4	4.1	-.014	1.6e-05
country==HU	.066	.07	.066	.062	.065	.062	3.5	3.4	3.5	-.014	1.2e-05
country==IS	.0064	.0087	.0064	.0064	.0086	.0064	12	11	12	-.028	-2.7e-06
country==IT	.12	.12	.12	.1	.11	.1	2.4	2.3	2.4	-.0079	6.4e-05
country==LT	.023	.027	.023	.022	.027	.022	6.4	5.8	6.4	-.031	-2.6e-06
country==LU	.042	.037	.042	.04	.035	.04	4.6	4.9	4.6	.027	3.2e-06
country==LV	.012	.031	.012	.012	.03	.012	8.9	5.4	8.9	-.17	-1.3e-04
country==MT	.029	.024	.029	.028	.023	.028	5.6	6.2	5.6	.031	2.8e-05
country==NL	.013	.032	.013	.013	.031	.013	8.4	5.3	8.4	-.16	-1.1e-04
country==NO	.0042	.015	.0042	.0042	.015	.0042	15	8	15	-.17	-3.8e-04
country==PL	.067	.073	.067	.062	.067	.062	3.5	3.3	3.5	-.024	-2.3e-05
country==PT	.086	.025	.086	.078	.025	.078	3	6	3	.22	7.5e-05
country==RO	.058	.035	.058	.055	.034	.055	3.8	5.1	3.8	.1	5.3e-05
country==SE	.023	.039	.023	.023	.037	.023	6.3	4.8	6.3	-.1	-6.2e-05
country==SK	.025	.034	.025	.024	.033	.024	6.1	5.2	6.1	-.057	5.3e-07
Lived with parent(s)	.99	.99	.99	.01	.0052	.01	-9.6	-14	-9.6	-.053	-5.8e-06
country_of_birth==EU	.054	.04	.054	.051	.038	.051	3.9	4.7	3.9	.065	-4.5e-06
country_of_birth==OTH	.08	.05	.08	.073	.047	.073	3.1	4.1	3.1	.11	-3.5e-06
father_country_of_birth==country of residence	.85	.89	.85	.13	.099	.13	-1.9	-2.5	-1.9	-.12	4.5e-05
mother_country_of_birth==country of residence	.85	.89	.85	.13	.099	.13	-1.9	-2.5	-1.9	-.11	4.6e-05
father primary education	.85	.62	.85	.13	.24	.13	-2	-.5	-2	.65	5.3e-04
mother primary education	.9	.69	.9	.09	.22	.09	-2.7	-.8	-2.7	.72	7.4e-04
Tenancy_status== Owner	.59	.74	.59	.24	.19	.24	-.36	-1.1	-.36	-.31	-1.7e-05
father Employed	.95	.98	.95	.049	.016	.049	-4.1	-7.6	-4.1	-.16	-1.6e-05
mother Employed	.47	.6	.47	.25	.24	.25	.11	-.42	.11	-.26	-1.4e-04

Table A2.4: T-test on the pre-treatment variables means by treatment status

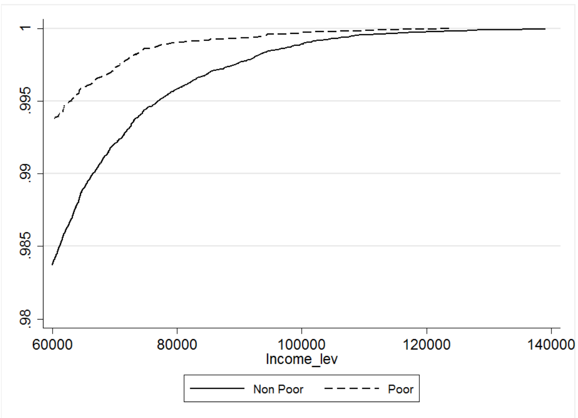
	Non poor	Poor	Diff
quarter of birth	2.22	2.21	0.01
year of birth	1964.78	1964.69	0.09
sex	1.53	1.52	0.01
n. of adult in hh	2.74	2.77	-0.03
n. of children in hh	2.94	2.93	0.02
n. of person in work	1.88	1.91	-0.03
year of birth of father	1809.65	1806.38	3.26
year of birth of mother	1835.12	1836.08	-0.95
country==AT	0.05	0.05	0.00
country==BE	0.02	0.02	0.00
country==BG	0.01	0.02	-0.00
country==CH	0.03	0.03	0.00
country==CY	0.06	0.05	0.00
country==CZ	0.03	0.02	0.00
country==DK	0.01	0.01	-0.00
country==EE	0.01	0.01	0.00
country==EL	0.04	0.04	0.00
country==ES	0.10	0.10	-0.00
country==FI	0.01	0.01	-0.00
country==FR	0.05	0.05	0.00
country==HU	0.06	0.07	-0.00
country==IS	0.01	0.01	0.00
country==IT	0.11	0.12	-0.01*
country==LT	0.02	0.02	-0.00
country==LU	0.04	0.04	0.00
country==LV	0.01	0.01	0.00
country==MT	0.03	0.03	-0.00
country==NL	0.02	0.01	0.00
country==NO	0.00	0.00	-0.00
country==PL	0.07	0.07	-0.00
country==PT	0.07	0.07	0.00
country==RO	0.06	0.06	-0.00
country==SE	0.01	0.01	-0.00
country==SK	0.02	0.03	-0.00
country==UK	0.03	0.03	0.00
Lived with parent(s)	0.99	0.99	-0.00
country_of_birth==EU	0.05	0.05	-0.00
country_of_birth==OTH	0.08	0.07	0.00
father_country_of_birth==country of residence	0.85	0.85	-0.00
mother_country_of_birth==country of residence	0.85	0.85	-0.00
father primary education	0.83	0.84	-0.01
mother primary education	0.88	0.89	-0.01
Tenancy_status== Owner	0.61	0.62	-0.01
father Employed	0.95	0.96	-0.00
mother Employed	0.49	0.48	0.00



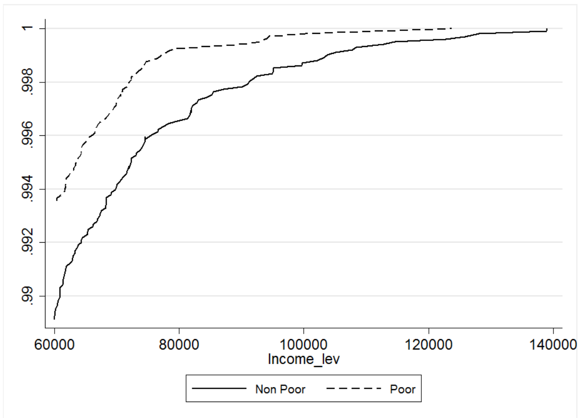
(a) Without Matching



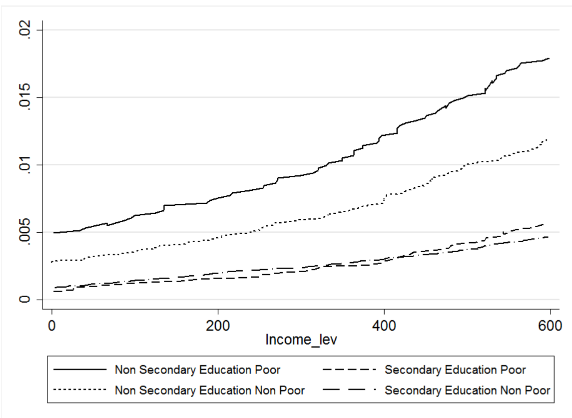
(b) Matched



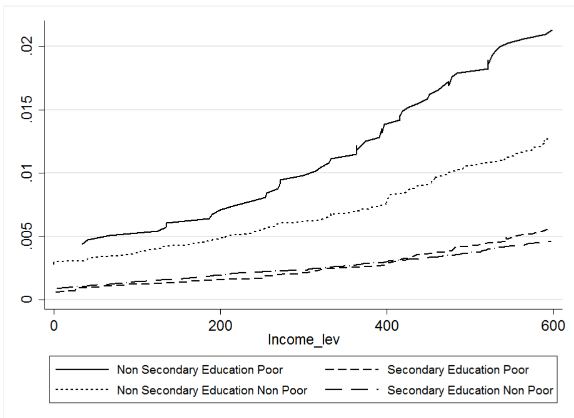
(c) Without Matching



(d) Matched



(e) Without Matching



(f) Matched

Figure A2.1: Cumulative density at the bottom and top of the distribution.

UNIVERSITY OF KONSTANZ
Department of Economics

Universitätsstraße 10
78464 Konstanz
Germany

Phone: +49 (0) 7531-88-3713
Fax: +49 (0) 7531-88-3130

www.wiwi.uni-konstanz.de/econdoc/working-paper-series/