I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10677

The Long-Run Impact of Childhood Poverty and the Mediating Role of Education

Luna Bellani
Michela Bia

MARCH 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10677

The Long-Run Impact of Childhood Poverty and the Mediating Role of Education

Luna Bellani
University of Konstanz and IZA

Michela Bia
LISER

MARCH 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Long-Run Impact of Childhood Poverty and the Mediating Role of Education*

This paper examines the role of education as causal channel through which growing up poor affects the economic outcomes in adulthood in the European Union. We apply a potential outcomes approach to quantify those impacts and we provide a sensitivity analysis on possible unobserved confounders, such as child ability. Our estimates indicate that being poor in childhood significantly decreases the level of income in adulthood and increases the average probability of being poor. Moreover, our results reveal a significant role of education in this intergenerational transmission. These results are particularly relevant for Mediterranean and Central and Easter European Countries.

JEL Classification: D31, I32, I24, J62

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

Corresponding author:
Luna Bellani
Department Economics
University of Konstanz
Universitätsstraße 10
78457 Konstanz
Germany
E-mail: luna.bellani@uni-konstanz.de

* The authors thank Alfonso Flores-Lagunes, Andreas Peichl, Guido Schwerdt, Philippe Van Kerm, Michael Vogt, Amelie Wuppermann and the participants to seminar at IZA and at LISER, to the 6th ECINEQ conference, to the workshop on Public Economics and Inequality in Berlin and the workshop on Uncovering Casual Mechanisms in Munich for comments on this or earlier version of the paper. This work has been supported by the second Network for the analysis of EU-SILC (Net-SILC2), funded by Eurostat. The European Commission bears no responsibility for the analyses and conclusions, which are solely those of the authors. In addition, 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). Previous versions of this paper have been circulated under the title "Intergenerational Poverty Transmission in Europe: the Role of Education". Usual disclaimers apply.

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 data from the Urban Institute 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 (Ratcliffe, 2015). In Europe the picture is similar. From the latest data provided by EUROSTAT, in 2014 in the EU-28 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, Romania and Bulgaria. Moreover, 50.5% of children whose parents' highest level of education was low were at risk of poverty compared to 8.0% of children whose parents' highest level of education was high (Eurostat, 2016). 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 for being low educated of 14.7 for adults having low educated parents. In 2011, in Bulgaria and Croatia, this association index was more than 40, while in Norway, Estonia, Denmark and Finland, it was less than 5. 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).

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 impacts 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 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 family structure, neighborhood influences, genetic transmissions, etc. Moreover, 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 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, the first method does not guarantee that estimates are unbiased, since there may remain some child-specific factors contributing to potential bias and the estimates are based on a selected type of family that could 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.

This paper is addressing the following questions: Does living in poverty as a child cause poverty in adulthood? If so, is education driving this causal effect?

We contribute to the above mentioned literature in three ways: i) we apply a potential outcomes framework to quantify the impact of experiencing financial difficulties while growing up; ii) we analyze the channel of this poverty transmission, introducing individual human capital accumulation as an intermediate variable and iii) we provide an extensive sensitivity analysis on unobserved confounders (e.g. parental ability) for both the direct and indirect impact.

Our analysis is based on the module on intergenerational transmission of 2011 of the EU-SILC data,[1]

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

where retrospective questions about parental characteristics (such as education, age, occupation) were asked. We find that, even considering possible unobserved confounders, such as child ability, being poor in childhood significantly decreases the level of income in adulthood and increases 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 Sections 4 and 5 we analyze the average and the distributional impact of growing up poor, respectively, while in Section 6 we focus on the mechanism behind this impact, analysing the role of education. Section 7 provides some heterogeneous effects by welfare state regime and Section 8 presents sensitivity analysis and robustness checks of our results. Finally, Section 9 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 context it is not possible for obvious ethical reasons to design such a randomized experiment on our population of interest (poor children), but even if we were able to randomly give unconditional cash transfer to poor parents today, we would not be able to analyze the long term impact of it on the children before another 15 to 20 years. [2] 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 on observables (unconfoundedness) (Rosenbaum and Rubin, 1983).[3]

Let us consider a set of N individuals, and denote each of them by subscript i: $i = 1, \dots, N$. Let

[2] See for example Haushofer and Shapiro (2016) for the short term impact of such transfers or Aizer et al. (2016) for an estimate of the long-run impact of cash transfers which compare accepted and rejected applicants to the Mothers' Pension program in the United States (1911-1935).

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

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 characteristics: $Y_i(0), Y_i(1) \perp T_i|X_i$. 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 up 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) (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 also analyze a mechanism behind this average effect. Human capital accumulation is essential to the individual's cognitive and non-cognitive abilities development and a key factor to the long-term reduction of poverty. Human development itself is very often the main target of a range of policies in both developed and developing countries. Hence, a better understanding of the interaction between cognitive skills development and the reduction of poverty will help to design more effective policies interventions. In order to do so we use a causal mediation analysis. The mediation analysis aims at quantifying the relative importance of a particular mechanism through which the effect of the treatment is mediated.

Several ways to conceptualize the mediatory role of an intermediate variable in the treatment - outcome relationship have been proposed in the causal inference literature. These methods cover semi- and non-parametric estimation procedures (Imai et al., 2010b; Pearl, 2001b,a; Hafeman and Schwartz, 2009), matching based on the propensity score (Hill et al., 2003), weighting procedures (Peterson et al., 2006; VanderWeele, 2009; Hong, 2010), principal stratification approach (Frangakis and Rubin, 2002; Jo, 2008; Jo et al., 2011) and the g-computation based algorithm (Robins and Greenland, 1992). As emphasized by a recent work of Linden and Karlson (2013), many of these methodologies are conceptually interconnected, or serve as basis for the extension of other techniques. For example, the propensity score (Rosenbaum and Rubin, 1983) performs as a way of coping with the Sequential Ignorability assumption, in either a separate procedure (Hill et al., 2003) or as a basis for weighting and principal stratification (Peterson et al., 2006; VanderWeele, 2009; Hong, 2010; Jo et al., 2011). One of the most popular framework for identifying and estimating causal mechanisms is the above mentioned principal stratification approach (PRS). [4] PRS defines causal effects by comparing individuals with the same potential values of the post-treatment variable under each of 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 mechanisms of a binary treatment variable under the unconfoundedness assumption, basing his estimation strategy on inverse weighting. This significant increase in the number of methods recently introduced to conduct mediation analysis leads to the question of how they should be compared and selected. Huber et al. (2016) help address this question using a simulation study based on Swiss jobseekers data in order to investigate the finite sample properties of different classes of parametric and semi-parametric estimators under sequential conditional independence assumptions. Linden and Karlson (2013) apply instead a variety of methods to the JOBS II dataset used in Imai et al. (2010a,b) and to simulated data in a Monte Carlo study. In the first analysis, the so-called g-computation technique dominates, but differences between estimators are often minor in the various scenarios. The second study also suggests that some methods perform better than others, but no significant difference was found among best methods in terms of performance. In particular, in this latter case, the approach by Imai et al. (2010a,b) is among the best performers overall, and, given its flexibility in terms of models for outcomes and mediating variables and the possibility of conducting sensitivity analysis with respect to key identification assumptions, it can be easily implemented in most scenarios often met in applied research. Therefore, 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 set of information, as the one included in our EU-SILC data, we follow the

[4]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 and inverse weighting.

procedure described in Imai et al. (2010a), that allows us 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 analyzes, 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-treatment characteristics. Assumption 2 states that the mediator variable is now ignorable, given the observed treatment level and the observed characteristics, that is, among those individuals with the same poverty status and the same pre-treatment observable characteristics, the level of education can be considered as if it were randomized. The average causal mediation effects can be then consistently estimated by nonparametric identification, that is, under the SI assumption, the distribution of any counterfactual outcome is identified without being based on any specific model. [5] In particular, Imai et al. (2010a) propose an algorithm to estimate the average causal mediation effects based on the quasi-Bayesian Monte Carlo approximation proposed by King et al. (2000), in which the posterior distribution of the quantities of interest is approximated by their sampling distribution. This procedure can be used for any parametric specification and, as explained in Imai et al. (2010a), it is based on the following steps:1) Fit the models for the observed outcome and mediator variables. 2) Simulate the model parameters from their sampling distribution. 3) Repeat the following three steps: i) simulate the potential values of the mediator, ii) simulate the potential outcomes given the simulated values of the mediator, iii) compute the causal mediation effects. 4) Compute summary statistics such as point estimates and confidence intervals.

Our key quantity of interest is the change in the outcome (children outcomes later in life) corresponding

[5] Refer to the Theorem on *Nonparametric Identification* in Imai et al. (2010a).

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). 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: $\bar{\gamma}_t = E[Y_i(1, M_i(t)) - Y_i(0, M_i(t))]$.

In our study the mediating variable is binary (achieving at least secondary education) and a probit model is used:

$$M_i = \mathbf{1}\left\{M_i^* > 0\right\} \tag{3}$$

where

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

When the outcome variable is continuous (equivalized disposable income) a linear regression model is implemented:

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

while when our outcome is binary (adult poverty) a probit model is implemented.[6]

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_{i3}, 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, with mean zero and covariance $\rho\sigma_3^2$, where ρ is the correlation between the two error terms. [7]

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). Moreover, since assuming no interaction between the treatment and the mediator variable is often unrealistic, we also include the interaction term in equation 4. Such an interaction might arise if, for example, the effect of the educational level depends on whether the individual grew up in poverty or not.

Following this approach, we provide in section 6 information on the extent to which a causal effect of growing up in poverty on later outcomes is due to the causal effect of growing up poor on human capital

[6] $Y_i = \mathbf{1}\left\{Y_i^* > 0\right\}$ where $Y_i^* = \alpha_3 + \beta_3 T_i + \kappa T_i M_i + \gamma M_i + \xi' X_i + \epsilon_{i3}$.

[7] 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.

accumulation. Moreover, we provide sensitivity analyzes to quantify the extent to which our empirical findings are robust to the existence of unobserved confounders.

3 Data

The analysis is based on data from 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.[8] For the specific purpose of this paper we use the module on intergenerational transmission of 2011,[9] 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. We define an individual as experiencing poverty while growing up if she/he reported the financial situation of the household as very bad or bad in her/his early adolescence.[10]

We are aware that individuals may suffer from retrospective recollection bias.[11] A recall bias would be present if there are selective preconceptions between groups of individuals (rich or poor, more or less educated) about the association between having experienced financial problems during childhood and the outcome of interest (income and education). In our context we argue that the type of questions asked are less affected by this problem compared, for example, with a direct question on the level of income in the household during the same period. In support of our argument on the validity of this measure, previous studies in the fields of medicine and psychology have found that the retrospective recall in adulthood of serious negative experiences in childhood is sufficiently valid, and it is more likely to have significant under-reporting than "false-positive" (see Hardt and Rutter (2004) for a review). In particular, closer to our setting, results from a test for concordance between siblings performed by Robins et al. (1985) show that young children are likely to have a general concept of whether the family was rich or poor, and, more importantly for the validity of treatment-control group analysis, they found no difference in the sibling agreements between group of patients with alcoholism or depression, and a control group free of psychiatric disorder, suggesting that interviews requiring recall of childhood environment may be reasonably valid. [12]

[8] Refer to chapter 2 in Atkinson et al. (2017) for a detailed description of this database.

[9] 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 a preliminary analysis of the 2005 module refer to Bellani and Bia (2017).

[10] As a robustness check we use a different definition of childhood poverty based on the ability of making ends meet. The results are presented in Section 8.

[11] Recall bias is said to occur when accuracy of recall is different by outcome.

[12] Refer also to Akerlof and Yellen (1985) and Jürges (2007) for examples of the use of retrospective questions in the context of occupational mobility and their accuracy.

Table 1: Descriptive Statistics

	mean	sd	min	max
Child characteristics				
quarter of birth	2.46	1.11	1	4
year of birth	1965.8	5.97	1956	1976
sex	1.53	0.50	1	2
country_of_birth==EU	0.034	0.18	0	1
country_of_birth==OTH	0.049	0.22	0	1
Parents characteristic				
year of birth of father	1935.9	8.57	1907	1957
year of birth of mother	1938.9	8.26	1913	1959
father not born in country of residence	0.11	0.31	0	1
mother not born in country of residence	0.11	0.31	0	1
father primary education	0.62	0.48	0	1
mother primary education	0.69	0.46	0	1
father self-Employed	0.18	0.39	0	1
father Employee	0.80	0.40	0	1
mother self-Employed	0.10	0.31	0	1
mother Employee	0.49	0.50	0	1
Household characteristic				
Single parent	0.036	0.19	0	1
Tenancy_status== Owner	0.73	0.45	0	1
n. of adult in hh	2.51	0.97	1	7
n. of children in hh	2.31	1.22	0	7
n. of person in work	1.85	0.88	0	6
Treatment				
Poor as child	0.10	0.30	0	1
Outcomes				
At least secondary education	0.79	0.40	0	1
Income	4.02	0.41	0.40	5.03
Poor as adult	0.14	0.34	0	1
Observations	108747			

We use as predictors of the probability of experiencing poverty in growing up the individuals' country of residence and of birth, the gender, the year and quarter of birth, single parenthood, year and country of birth, highest level of education, main occupation of the father and of the mother, the tenancy status, the number of adults, children and persons in the labor force in the household.

The outcomes in adulthood that we are interested in are the log of the equivalised income,[13] and the risk of poverty (defined as having an equivalised 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 we analyze the probability of having completed at least secondary education. We further restrict our sample to the individuals 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.[14]

As presented in tables 1 and 2, we are left with a sample of more than 100,000 individuals, belonging

[13] We use equivalised 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.

[14] See for example Chetty et al. (2014) who show that life cycle bias in measure of intergenerational mobility are negligible when the child income is measured after age 32 .

Table 2: T-test on the control variables means by treatment and mediator status

	Poor			Secondary Education		
	No	Yes	Diff	No	Yes	Diff
quarter of birth	2.46	2.45	0.01	2.45	2.46	-0.01
year of birth	1965.92	1964.48	1.44***	1964.95	1965.98	-1.04***
sex	1.53	1.52	0.01	1.53	1.53	-0.00
n. of adult in hh	2.48	2.70	-0.21***	2.74	2.44	0.29***
n. of children in hh	2.25	2.86	-0.60***	2.65	2.23	0.42***
n. of person in work	1.87	1.77	0.10***	1.87	1.85	0.02***
year of birth of father	1936.08	1934.31	1.77***	1934.47	1936.27	-1.81***
year of birth of mother	1939.14	1937.24	1.90***	1937.48	1939.33	-1.85***
Single parent	0.03	0.08	-0.05***	0.03	0.04	-0.01***
country_of_birth==EU	0.03	0.04	-0.01**	0.04	0.03	0.01***
country_of_birth==LOC	0.92	0.89	0.03***	0.90	0.92	-0.02***
country_of_birth==OTH	0.05	0.07	-0.02***	0.06	0.05	0.01***
father not born in country of residence	0.10	0.13	-0.03***	0.11	0.11	0.00
mother not born in country of residence	0.10	0.13	-0.03***	0.11	0.10	0.00
father primary education	0.60	0.83	-0.23***	0.90	0.55	0.35***
mother primary education	0.67	0.88	-0.22***	0.94	0.62	0.31***
Tenancy_status== Owner	0.74	0.61	0.12***	0.69	0.74	-0.05***
father Employee	0.81	0.73	0.08***	0.73	0.82	-0.09***
father self-Employed	0.18	0.22	-0.04***	0.24	0.16	0.08***
mother Employee	0.50	0.33	0.17***	0.28	0.54	-0.27***
mother self-Employed	0.10	0.14	-0.04***	0.13	0.10	0.03***

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

to the 28 European Union's countries,[15] of which 10% reported to have experienced poverty growing up.[16] Among those we have on average more first generation migrants (11% against 8%), more children of single parent (8% versus 3%) and more children with mother out of the labor force (more than 50% against 40%). The most sizable significant difference among these two groups of children can be found with respect to parental education, the poor being around 23 percentage points more likely to have at maximum a primary education degree. This difference is even more pronounced once we look at individuals who have less than secondary education, of which almost 95%(90%) have a mother (father) with only a primary education degree, 30 (35) percentage points more than in the more educated group. Also remarkable is the higher difference in maternal employment, the individuals with less than secondary education are around 24 percentage points more likely to have a mother out of the labor force (almost 60%). Interestingly there is not significant difference in the parental migration status.

To give a better sense of how representative is our sample with respect to the total population, we can look at the average risk of poverty in 2011 in the EU-27, as reported by EUROSTAT. This was of 24.2% of the total population. If we then focus on people between 25 and 54 years old, closer to our sample of interest, this percentage is 22.9 of the total population. Knowing that around 42% of the EU-27 population is in this range of age, while 29% is in the range we are interested in, we can easily see that if

[15] Refer to table A1.1 for a list of the countries included in our study.

[16] In table A1.2 we provide the results of a test on the average differences in the relevant outputs between our sample and the sample of individuals in the same range of age with missing value of our treatment variable. We can notice that there is no difference in their poverty risk, but the individual in our sample are on average richer and slightly less educated.

poverty risk was uniform in this age group, around 15% of our adult population should be poor to be in line with the official data. Our sample shows a slightly lower poverty risk, 14%, and it could be credibly argued that this is due to the higher poverty risk in the age group between 25 and 29 years (25.5% of the overall population), who is excluded by construction from our sample.

4 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 unmatched sample the propensity score has a mean value of .11, a median of .08 and a standard deviation of .1.[17] 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 covariates before and after adjusting for the estimated propensity score (PS). Figure 1 (a) provides the standardized bias (in percentage) for unmatched and matched units,[18] 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 comparison of baseline covariates may be complemented by comparing the distribution of the estimated PS between treated and controls, as shown in figure 1(b).

As a second step, we apply a single nearest-neighbor matching to remove bias associated with differences in covariates[19] and estimate the effects of being poor on adulthood outcome, primarily equivalent income and being at risk of poverty. [20]

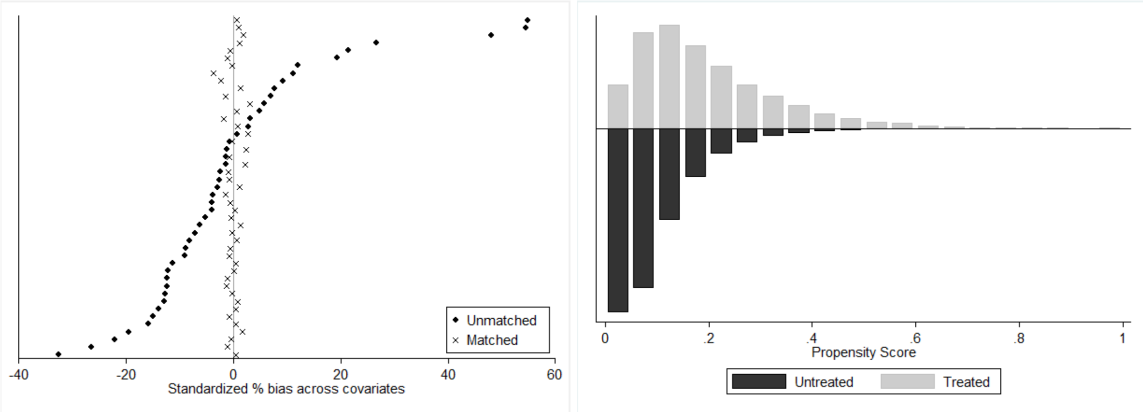
In order to check the robustness of our results, we perform also a doubly-robust estimation of our treat-

[17]For the results of this first step refer to column (1) of table A1.4 in the Appendix A.1. Note that as a robustness check we have been estimating this probability also using interaction terms, the results are presented in Section 8.

[18]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.

[19]To note that not only the mean value of the difference between the propensity score of the treated individual and the one 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.

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



(a) (b)

Figure 1: Standardized bias and common support

ment 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.[21] The marginal effect of the treatment is our doubly-robust estimate of the average treatment effect.[22]

Results are presented in table 3. The results on income show a substantial decrease in the equivalized income in adulthood due to exposure to poverty in childhood, of around 5%, equal to an average loss of around 764 €which is half of the average gross monthly earning of people with less than secondary degree in 2010, in the EU-28. [23] We also find a significant increase in the probability of falling into poverty of 5 percentage points, which means that on average poor children are 1.5 times more at risk of poverty in adulthood then the non poor. Regarding our intermediate outcome of interest, our results show an average of 11 percentage points decrease in the probability of completing at least secondary education, which means that on average the children who did not experience poverty are 1.8 times more likely to have at least a secondary education degree. The estimates are overall very close between the two methods, both in the magnitude and in the significance.

[21]See table A1.5 in Appendix A.1 for the descriptive statistics of the sample pre and post reweighing.

[22]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.

[23]Source: `http://ec.europa.eu/eurostat/web/labor-market/earnings/database`.

Table 3: Average Treatment on the Treated

	Main Outputs		Intermediate Output
	Income	Poverty	Education
Propensity score matching	-0.06	0.06	-0.12
	[-0.071, -0.043]	[0.043 ,0.071]	[-0.134, -0.104]
Doubly-Robust Estimation	-0.05	0.05	-0.11
	[-0.054, -0.044]	[0.044 ,0.060]	[-0.122, -0.104]
N	108355	108355	107566

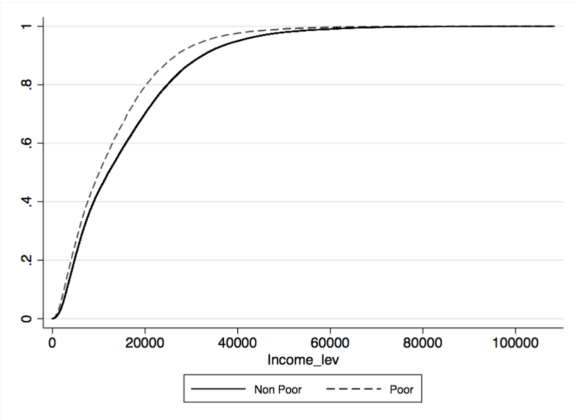
Note: 95% Confidence Intervals in brackets.

5 Distributional effect of childhood poverty and education

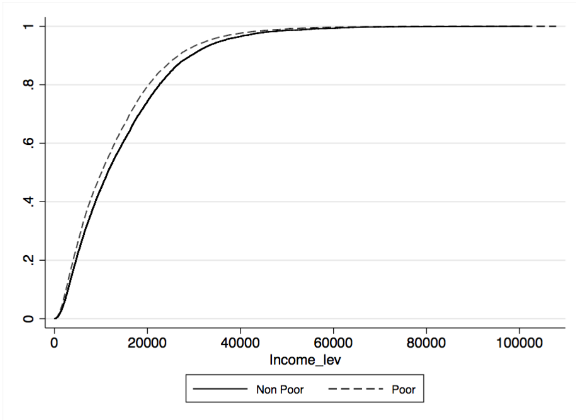
After having shown that childhood poverty has indeed a relevant detrimental impact on economic outcomes in adulthood, a fundamental question left to answer to be able to make more effective policy recommendation regards the channels through which being raised in a poor family affects the individual's economic and social status as an adult. In order to do so, as a first step in this direction, we begin by exploring this impact more in detail presenting and comparing 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) in figure 2 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 experiences poverty as a child than one in which childhood poverty is common. 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, 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) in figure 2).

The Kolmogorov-Smirnov test shows in fact that there are significant differences in the income distribution for these two groups. The largest difference between the distribution functions is of almost 10% in the unmatched sample which reduces to a 6% in the matched sample, remaining always highly significant. Focusing on the risk of poverty, we can see that not only the incidence of poverty is higher, around 21% against 13% (15% in the matched sample) for the non poor children, but also its intensity. Finally, it is worth to notice that the impact of childhood poverty seems to decrease the income achievable by

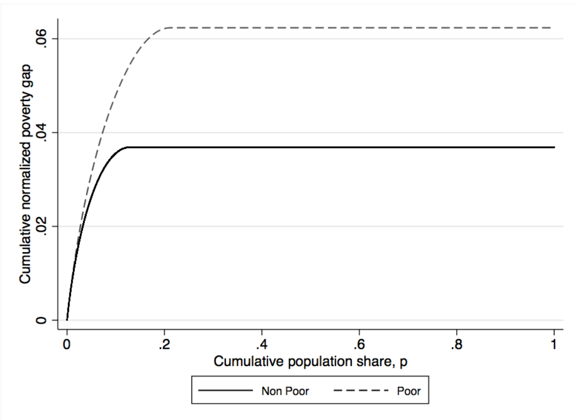
[24] In Figure A1.1 in Appendix A.1 the bottom and the top of the distribution is plotted for both cases to show that the cumulative densities are not crossing.



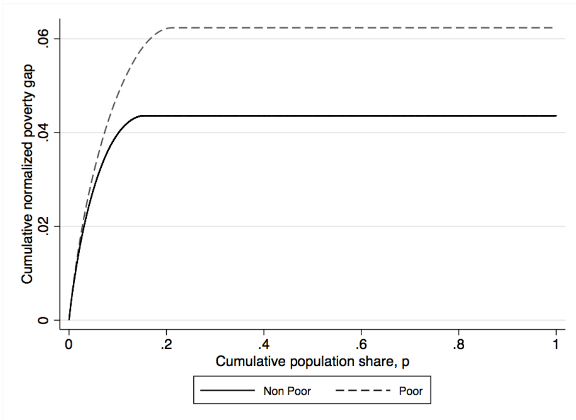
(a) Cumulative Distribution Function-full sample



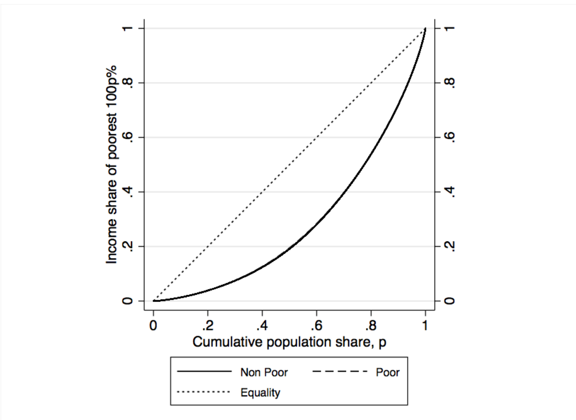
(b) Cumulative Distribution Function-matched sample



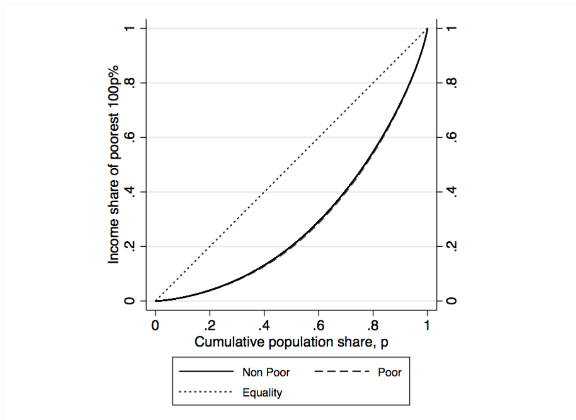
(c) Three I's of Poverty (TIP) curves-Full sample



(d) Three I's of Poverty (TIP) curves-matched sample

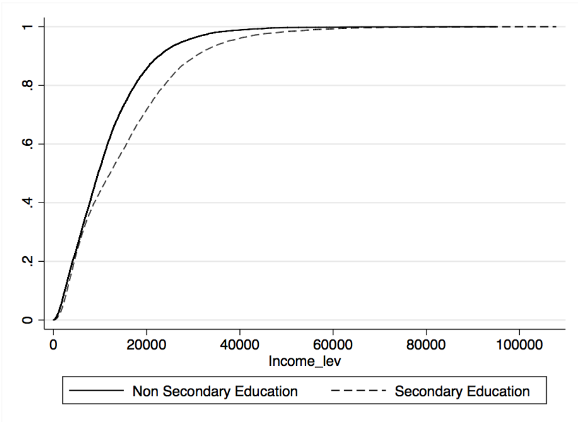


(e) Lorenz Curves-full sample

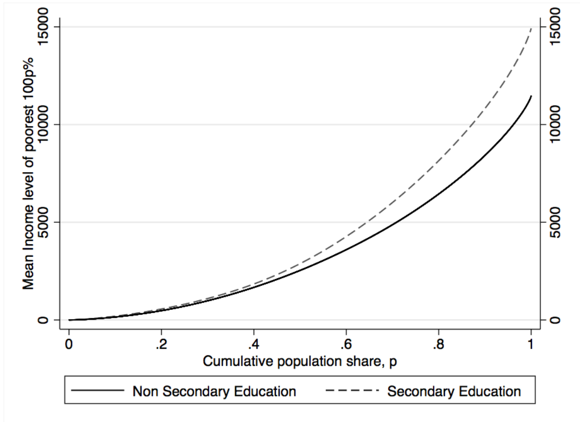


(f) Lorenz Curves-matched sample

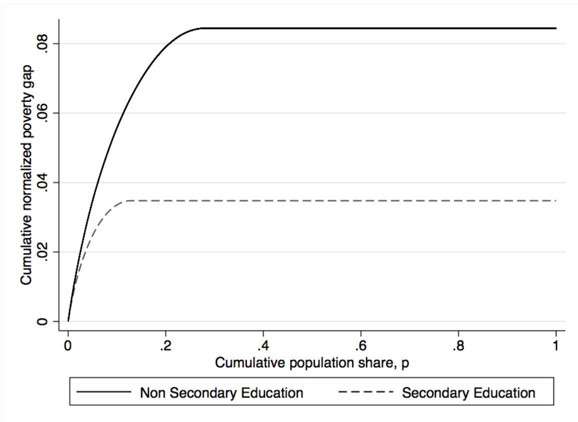
Figure 2: Distributional Graphs



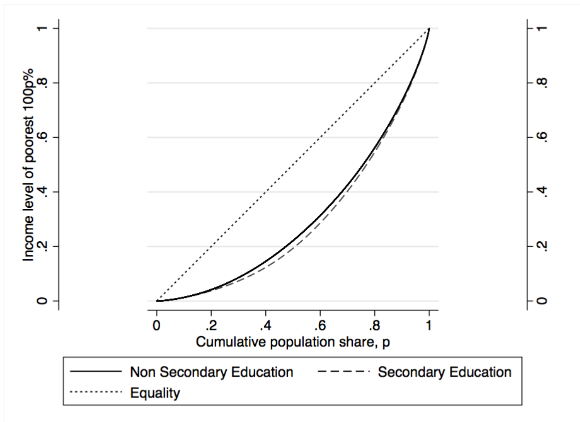
(a) Cumulative Distribution Function-matched sample



(b) Generalized Lorenz Curves-matched sample



(c) Three I's of Poverty (TIP) curves-matched sample

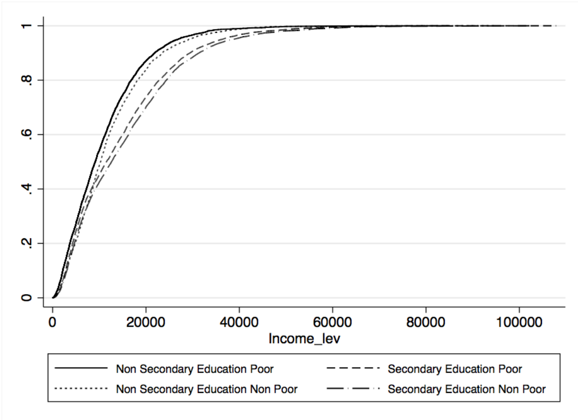


(d) Lorenz Curves-matched sample

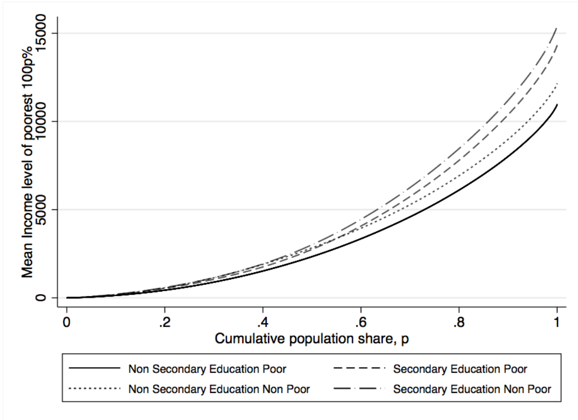
Figure 3: Distributional Graphs-Education

the two groups of children but not to impact how this is distributed within each group (see part (e) and (f) of figure 2).

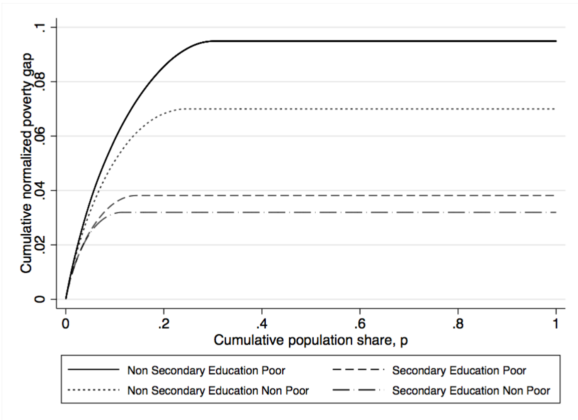
Performing the same type of exercise, but focusing on the impact of secondary education, we find that the individuals having only less than secondary education have consistently lower income at any percentile of the distribution reaching a maximum of 15% difference with respect to the more educated ones, a difference which is also in this case highly significant, as we can see from part (a) and (b) of figure 3. When we look at the risk of poverty by educational level these differences are more pronounced, in fact not only the incidence of poverty is higher, around 27% with respect to 13%, but the difference in its intensity are of particular significance as the average income of the poor drops at only 54% of the median income. Inequality, not so surprisingly, is slightly higher for the more educated group (see part (c) and (d) of figure 3, for poverty and inequality, respectively). To conclude this part, before moving to the mediation analysis on 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, but also by their own subsequent educational choice. As we were expecting, the income distribution of the poor children



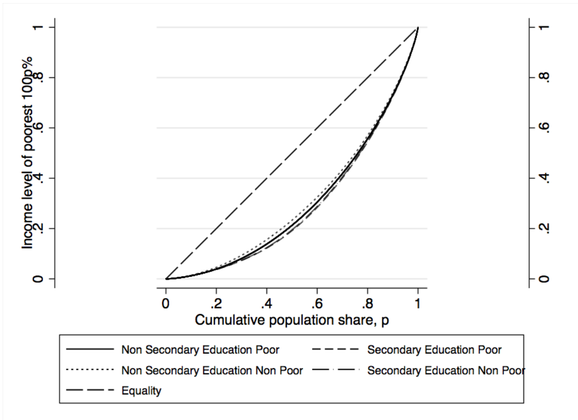
(a) Cumulative Distribution Function-matched sample



(b) Generalized Lorenz Curves-matched sample



(c) Three I's of Poverty (TIP) curves-matched sample



(d) Lorenz Curves-matched sample

Figure 4: Distributional Graphs-Poverty & Education

without a secondary degree is dominated by all the others, but what is less expected, education does not seem to be able to offset the impact of having experienced poverty in childhood for the bottom 40% of the income distribution, while it does so, although never completely, along the rest of the distribution. If we then look at poverty risk, we can notice that acquiring a secondary degree reduces the incidence and intensity of poverty quite drastically for the poor children, of whom only 13% are poor, compared with 11% of the ones who did not experience poverty, while they are consistently less likely to fall into poverty than the children who did not experience poverty, but who did not obtain higher education, for whom the incidence of poverty is instead of 23%.

6 The mediating role of education

Our analysis has shown so far that i) the individuals who experience poverty growing up are significantly less likely to achieve secondary education and ii) the individuals who have less than a secondary degree have consistently lower income. Therefore, it seems worth focusing at first on the mediating role of human capital accumulation. In order to do so, as introduced at the end of section 2, in this last part of our analysis we implement a causal mediation analysis to uncover the role of education 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 at least graduating from High School. Our mediating variable is therefore secondary education, while the outcomes of interest are the income level and the poverty risk in adulthood. First, based on the fitted mediator model (eq.3), we generate predicted secondary education attainment levels for the children who experienced and who did not experience poverty. Next, we use the outcome model (eq.4) to impute potential outcomes. [25] The average causal mediation effect (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 1000 times in order to estimate the standard errors.

Table 4: Mediation Causal Analysis

	Income	Poverty
Total Effect	-0.05	0.06
	[-0.058, -0.045]	[0.046, 0.066]
Average		
ACME	-0.02	0.02
	[-0.018, -0.015]	[0.016, 0.021]
Direct Effect	-0.03	0.04
	[-0.041, -0.029]	[0.027, 0.047]
Tot.Eff. Mediated	0.32	0.33
	[0.285, 0.365]	[0.276, 0.400]
Poor		
ACME	-0.02	0.02
	[-0.020, -0.015]	[0.017, 0.022]
Direct Effect	-0.04	0.04
	[-0.043, -0.030]	[0.028, 0.049]
Tot.Eff. Mediated	0.34	0.35
	[0.302, 0.386]	[0.296, 0.428]
Non Poor		
ACME	-0.02	0.02
	[-0.017, -0.014]	[0.014, 0.020]
Direct Effect	-0.03	0.04
	[-0.040, -0.028]	[0.027, 0.046]
Tot.Eff. Mediated	0.30	0.30
	[0.269, 0.344]	[0.256, 0.371]

Note: 95% Confidence Intervals in brackets.

Table 4 shows the estimated ACME, average direct effect and average total effect. The total effect is equal to -5% of equivalized income and $+5p.p.$ risk of being poor. The indirect effect (ACME), i.e. the portion of the total effect, conveyed through the educational level, is estimated to be around -2% and $2p.p.$, suggesting that a significant portion of the average total effect is attributable to a decrease in the

[25]For the binary outcome risk of poverty we use a probit model.

probability of graduating from High School. Hence, growing up poor as a child induces a lower level of education which accounts for more then 30% of the total effect on adult income and on poverty risk. The impact of education does not show any significant heterogeneous impact by poverty status on both outcomes.

While education is a major contributor, the part of the transmission mechanisms that remains unexplained is substantial (around 3% of income and $4p.p.$ in the risk of poverty). As a first possible explanation for it, we should remind the reader that, due to data availability, our measure of education only give us information on the degree completed, but does not allow us to distinguish between different quality of those degrees. 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.

7 Heterogeneous effect by Welfare Regime

In order to analyze possible heterogeneous effects of education due to different welfare regimes, in this subsection we provide the results for both the direct and the indirect impact of poverty by sub-sample of countries, defined following the well-known and established classification of Esping-Andersen (1990) with the addition of the Central and Eastern European countries (Hemerijck, 2012).[26]

Table 5: Average Treatment Effect on the Treated by Welfare State Regime

	Main Outputs		Intermediate Output
	Income	Poverty	Education
Continental	-0.03	0.01	-0.07
	[-0.040, -0.010]	[-0.011, 0.032]	[-0.097, -0.039]
N.	24733	24733	24648
Mediterranean	-0.06	0.07	-0.16
	[-0.082, -0.048]	[0.037 ,0.094]	[-0.190, -0.126]
N.	43618	43618	43343
Social Democratic	-0.04	0.03	-0.03
	[-0.088, 0.009]	[-0.029 ,0.096]	[-0.104, 0.036]
N.	4282	4282	4252
Central and Eastern Europe	-0.03	0.05	-0.10
	[-0.044, -0.014]	[0.031 ,0.072]	[-0.120, -0.076]
N.	43618	43618	43343
Liberal	-0.01	0.00	-0.05
	[-0.054, 0.036]	[-0.070 ,0.076]	[-0.121, 0.013]
N.	4623	4623	4509

Note: 95% Confidence Intervals in brackets.

[26]See table A1.6 in Appendix A.1 for the categorization of the countries.

As presented in table 5, Scandinavian and Anglo-Saxon countries do not show significant causal impacts of childhood poverty on later life outcomes, while these impacts are most pronounced for Mediterranean countries, which show a higher income loss (6% vs 3%) and lower probability of attaining at least secondary education (16p.p. vs 7p.p. and 10p.p., respectively) than both Continental and Central and Easter European countries. It is quite well known that Scandinavian countries have a robust dual earner model with universal income security and strong incentive for high female employment participation. Therefore, the results about this group is not so surprising, suggesting that the average difference in income, poverty risk and education level existing between poor and non poor children are substantially driven by the factors who are responsible for childhood poverty in the first place. [27] The results on the Liberal countries, UK and Ireland, seem at a first glance at odds with previous results in the literature (in particular regarding the UK) but we ought to remember that previous contributions looked either at general male earnings mobility (Blanden et al., 2007) or at a particular childhood circumstance, as single motherhood or joblessness (Ermisch et al., 2004), both of which in this contribution are controlled for. Blanden and Gibbons (2006) look explicitly at poverty transmission but only at associations, they find that 19% of men who experienced poverty as teenagers in the 1970s were in poverty in their thirties, compared with only 10% of those who were not poor in their teens, with the size of the transmission similar at ages 33 and 42. However, when they compare people who had similar teenage family background in other ways apart from family poverty, they find that the direct effect of teenage poverty becomes small and statistically weak showing that the persistence of poverty from teenage through to middle age can be fully explained by the other aspects of teenage disadvantage, in particular whether one had low-educated or non-employed parents.

When we analyze more in detail the mechanisms behind the significant impacts in the Continental, Mediterranean and the Central and Easter European countries, we can notice that in Continental Europe, although poor children are on average 7 percentage points less likely to obtain at least a secondary degree, this does not seem to have a significant direct impact on the lower average income of this group. Moreover, although Mediterranean countries exhibit on average a higher impact of childhood poverty, the role of secondary education in this process do not seem to differ between those and the CEE countries, which in fact display almost the same total average effect and the same composition between direct and indirect effect. Results are presented in tables 6 and 7.

[27] See table A1.7 for these average differences by welfare state regime.

Table 6: Mediation Causal Analysis for Income by Welfare State Regime

	Continental	Mediterranean	Central & Eastern Europe
Total Effect	-0.02	-0.06	-0.05
	[-0.031, -0.004]	[-0.070, -0.043]	[-0.057, -0.037]
Average			
ACME	-0.01	-0.02	-0.02
	[-0.010, -0.004]	[-0.029, -0.021]	[-0.021, -0.014]
Direct Effect	-0.01	-0.03	-0.03
	[-0.023, 0.004]	[-0.045, -0.018]	[-0.040, -0.020]
Tot.Eff. Mediated	0.39	0.44	0.37
	[0.225, 1.203]	[0.225, 0.572]	[0.303, 0.473]
Poor			
ACME	-0.01	-0.02	-0.02
	[-0.009, -0.004]	[-0.029, -0.020]	[-0.021, -0.014]
Direct Effect	-0.01	-0.03	-0.03
	[-0.023, 0.004]	[-0.045, -0.019]	[-0.041, -0.019]
Tot.Eff. Mediated	0.33	0.43	0.38
	[0.191, 1.021]	[0.353, 0.564]	[0.307, 0.479]
Non Poor			
ACME	-0.01	-0.03	-0.02
	[-0.011, -0.005]	[-0.029,]	[-0.020, -0.014]
Direct Effect	-0.01	-0.03	-0.03
	[-0.024, 0.003]	[-0.046, -0.017]	[-0.039, -0.020]
Tot.Eff. Mediated	0.45	0.45	0.37
	[0.259, 1.386]	[0.362, 0.579]	[0.299, 0.466]

Note: 95% Confidence Intervals in brackets.

Table 7: Mediation Causal Analysis for Poverty by Welfare State Regime

	Continental	Meditteranean	Central & Eastern Europe
Total Effect	0.02	0.05	0.06
	[0.001, 0.045]	[0.037, 0.070]	[0.047, 0.081]
Average			
ACME	0.01	0.02	0.02
	[0.006, 0.014]	[0.017, 0.026]	[0.018, 0.030]
Direct Effect	0.01	0.03	0.04
	[-0.011, 0.037]	[0.014, 0.051]	[0.022, 0.057]
Tot.Eff. Mediated	0.43	0.40	0.36
	[0.197, 2.108]	[0.301,]	[0.284, 0.493]
Poor			
ACME	0.01	0.02	0.02
	[0.005, 0.012]	[0.019, 0.029]	[0.019, 0.030]
Direct Effect	0.01	0.03	0.04
	[-0.014, 0.035]	[0.016, 0.053]	[0.022, 0.059]
Tot.Eff. Mediated	0.36	0.44	0.38
	[0.163, 1.751]	[0.331, 0.633]	[0.295, 0.512]
Non Poor			
ACME	0.01	0.02	0.02
	[0.007, 0.017]	[0.014, 0.024]	[0.017, 0.029]
Direct Effect	0.01	0.03	0.04
	[-0.007, 0.038]	[0.013, 0.050]	[0.023, 0.056]
Tot.Eff. Mediated	0.51	0.36	0.35
	[0.230, 2.464]	[0.270, 0.517]	[0.273, 0.474]

Note: 95% Confidence Intervals in brackets.

8 Robustness and Sensitivity Analysis

8.1 Sensitivity analysis on the ATT

One of the central assumptions of our analysis is that being poor in childhood can be considered as good as random, conditional on the set of covariates X. This implies that 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.[28] 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 biased if we were to believe that even conditional on all the covariates we can observe (parental education and occupation, 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: $E[Y_{0_i}|T_i = 1, X_i, A_i] = E[Y_{0_i}|T_i = 0, X_i, A_i]$.

As a first test on the robustness of our results we follow the recent contribution by Oster (2016) and we evaluate the robustness to an omitted variable bias under the assumption that the relationship between treatment and unobservable can be recovered from the relationship between treatment and observables. Oster (2016) builds on the work done by Altonji et al. (2005) and extends it by explicitly connecting the bias to coefficient stability, showing that it is necessary to take into account both coefficients and R^2 movements in evaluating robustness to omitted variable bias. She develops a consistent, closed-form, estimator for omitted variable bias which allows to calculate a consistent estimate of the bias-adjusted treatment effects under the assumptions that, as in Altonji et al. (2005), the R^2 from an hypothetical regression of the outcome on treatment and both observed and unobserved controls is equal to 1 and that there is proportional selection on observed and unobserved variables. Moreover, with this method, is possible to calculate the degree of selection on unobservables relative to observables which would be necessary to drive the effect to zero. Formally, an approximation of the bias-adjusted treatment effect

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

β^* can be calculated as follows:

$$\beta^* \approx \tilde{\beta} - \delta[\dot{\beta} - \tilde{\beta}]\frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}}$$

where $\dot{\beta}$ and $\dot{R}$ are the coefficients and the R^2, respectively, of the regression of adult income on the childhood poverty dummy only, while $\tilde{\beta}$ and $\tilde{R}$ are the coefficient and the R^2, respectively, of the regression of adult income on the childhood poverty dummy and all the observables. R_{max} is the R^2 of a hypothetical regression of our outcome on the treatment, the observables and the unobservables, and δ measures the degree of selection on unobservables relative to observables.

Under the above-mentioned assumptions, we estimate the coefficient of proportionality which moves the effect of being poor as a child on adult income towards 0 and find that the unobservables would need to be 20 times as important as the observables to produce a null treatment effect. Moreover, we calculate the bias-adjusted effect of childhood poverty on adult income under the assumption of equal selection and double selection on unobservables relative to observed controls and we find an impact of -0.050 in the first case and -0.048 in the second, confronted with an unadjusted effect of -0.052.

Secondly, 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 $\Gamma = 2$ a 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 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)).[29]

Results on Rosenbaum's bounds for the p-values from Wilcoxon's signed rank test show that for an increase between 1.1 and 1.2 in Γ, the lower bound of the p-value increases to a level above the usual 0.05 threshold. [30] This means that if we were to take two children with the same observed characteristics, but one with parents with high ability ($A = 1$) and one with parents with low ability ($A = 0$), the child whose parents are of low ability type would have to have 1.2 times the risk of experiencing poverty for our effect on adult income to become insignificant. Another way of seeing it, the impact (γ) of a dichotomous parental ability variable should be larger than the impact of an observable variable like maternal employment and as large as the impact of parental migration background.

The results of the Mantel-Haenszel test for the binary outcome risk of poverty, 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 parents' ability, caused the odds of experiencing poverty in childhood to be 1.4 times higher for the children of low able parents, conditional on all the observable characteristics. The impact of secondary education is the most robust as it would remain significant even if the unobserved variable would cause the odds of experiencing poverty in childhood to be 1.7 times higher. This impact should thus be more than the impact of migration status of the father and the presence of an extra child in the household and almost as much as maternal education, respectively.

Finally, 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

[29]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_1s and N_0s be the numbers of treated and control units in stratum s, where $N_s = N_0s + N_1s$. Then, let us define Y_1s and Y_0s the number of successful participants and non-participants, respectively, with $Y_s = Y_0s + Y_1s$. The Q_{MH} test statistics is given by: $Q_{MH} = \frac{|Y_1 - \sum_{s=1}^{S}(\frac{N_1sY_s}{N_s})| - 0.5}{\sqrt{\sum_{s=1}^{S}\frac{N_{1s}N_{0s}Y_s(N_s - Y_s)}{N_s^2(N_s-1)}}}$.

[30] In this part of our analysis we use the program `rbounds` provided by Gangl (2004), which allows to run sensitivity analyzes for continuous outcomes, while to deal with binary outcomes we use the module built by Becker and Caliendo (2007) `mhbounds`. The complete results can be found in the Appendix A.1 in tables A1.9, A1.10 and A1.11 for income, poverty and secondary education, respectively.

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). [31] 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.[32] In the first one, the p_{ij} are set so as to let our simulated parental ability A mimic the behaviorof 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* con-

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

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

founder). In tables 8, 9 and 10 the results of these sensitivity checks are presented, for poverty

Table 8: Sensitivity Analysis: p_{ij} values and odds ratio-Poverty

	Father's Educ-Calibrated	Mother's Educ-Calibrated	Killer
p_{11}	0.11	0.07	0.36
p_{10}	0.19	0.13	0.26
p_{01}	0.27	0.22	0.20
p_{00}	0.42	0.35	0.05
Outcome Effect	0.52	0.53	4.76
Selection Effect	0.32	0.26	5.03

Table 9: Sensitivity Analysis: p_{ij} values and odds ratio-Education

	Father's Educ-Calibrated	Mother's Educ-Calibrated	Killer
p_{11}	0.05	0.03	0.55
p_{10}	0.26	0.18	0.30
p_{01}	0.11	0.07	0.40
p_{00}	0.46	0.39	0.05
Outcome Effect	0.14	0.12	12.69
Selection Effect	0.36	0.30	4.22

and education, respectively. These results show that our findings are robust to the introduction of both "calibrated" and "killer" confounders. To make our results invalid, the odds of experiencing poverty in childhood need to be 5 or 4.2 times higher for high ability parents than for low ability ones, and the odds of being at risk of poverty in adulthood almost 4.8 times higher and the odds of having less than secondary education almost 12.7 times higher for non-poor children of high ability parents than the ones of low ability parents.

Table 10: ATT estimation

	Baseline	Father's Educ-Calibrated	Mother's Educ-Calibrated	Killer
Poverty	0.05	0.03	0.04	-0.00
	[0.044, 0.060]	[0.020,0.048]	[0.022, 0.051]	[-0.018, 0.011]
Education	0.11	0.07	0.07	0.00
	[0.104, 0.122]	[0.056,0.088]	[0.056, 0.090]	[-0.020, 0.022]

Note: 95% Confidence Intervals in brackets.

Moreover we can see that here the negative effect of our calibrated parental ability on the odds of being poor as adults of non-poor children is lower than the negative effect of parental ability on the probability of experiencing poverty in childhood, while this selection effect is lower than the impact on the odds of having less than secondary education among non poor children. This could be partly explained in light of the way we calibrated our unobserved parental ability, based on their education level, which is known to have a direct impact on the education of the children. [33] Finally, while the significance of the estimated effects are robust to the introduction of the calibrated unobserved ability, the size of the effects are revised downward, suggesting an upward bias in the estimates. In other terms, the sensitivity analysis is telling us that the existence of a confounder behaving like maternal education might account for nearly 30% of the baseline estimates for poverty risk. If we were to consider this as a good approximation of the bias also on the impact on the equivalized income, we would then still have a significant impact of circa 4%.

Table 11: Average Treatment on the Treated-Making Ends Meet

	Main Outputs		Intermediate Output
	Income	Poverty	Education
Propensity score matching	-0.03	0.03	-0.09
	[-0.044, -0.025]	[0.023 ,0.045]	[-0.098, -0.072]
Doubly-Robust Estimation	-0.04	0.03	-0.09
	[-0.041, -0.033]	[0.028 ,0.042]	[-0.093, -0.078]
N.	106448	106448	105679

Note: 95% Confidence Intervals in brackets.

To conclude this part, we address a couple of concerns that might arise regarding our definition of childhood poverty and the specification used to estimate the probability of being poor in childhood. In table 11 we show the estimates for a different definition of childhood poverty based on the ability of making ends meet. The results are consistent in terms of significance and only slightly smaller in terms of point estimates, as might have been expected given that those two variables are highly correlated ($\rho = 0.73$) and the percentage of treated with this definition is higher (around 17%). In table 12 we provide the results when we estimate the probability of being poor as a child using interaction terms between all the variables and the country dummies and between education and employment status of the parents. Our average treatment effects are not significantly affected by different specifications of the model.

Table 12: Average Treatment on the Treated (pscore with interactions)

	Main Outputs		Intermediate Output
	Income	Poverty	Education
All	-0.04	0.05	-0.10
	[-0.053, -0.030]	[0.032 ,0.059]	[-0.119, -0.087]
Education and Employment	-0.04	0.05	-0.11
	[-0.054, -0.025]	[0.035 ,0.059]	[-0.125, -0.093]
N.	108255	108255	107466

Note: 95% Confidence Intervals in brackets.

[33] See among others (Piopiunik, 2014) and the references therein.

8.2 Sensitivity Analysis on the Mediation Effect

As previously mentioned in section 2, both the standard unconfoundedness assumption (*Assumption 1*) and the sequential ignorability assumption (*Assumption 2*) rely on 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 and report the indirect effects as a function of ρ. When ρ is 0, assumption 2 is satisfied, i.e., 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 $\rho = 0.3$, that means that the true mediated effect could be 0 if there were a modest violation of the SI. As the sensitivity parameter itself is rather difficult to interpret directly, we show here an alternative approach, expressing the degree of sensitivity as a function of R^2, that is the usual coefficient of determination. In the 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.[34] 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). [35] 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 5 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 to 0. In particular, the

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

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

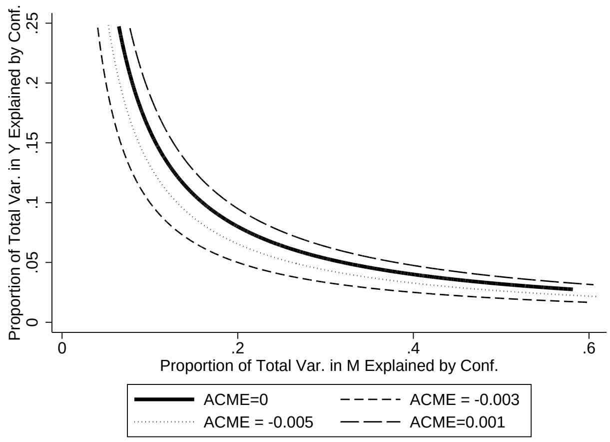


Figure 5: Sensitivity Analysis

true ACME would change sign when the product of the proportions is greater than 0.02. For example, we might think that pre-existing 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, we can say that our results will be

Table 13: 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

insignificant if the unobserved confounder U_i was as important in explaining the probability of acquiring at least a secondary education degree and the household income in adulthood as maternal education, which in both our outcome and mediator estimation has one of the most important impacts in terms of magnitude. [36] We think that the existence of such a confounder is quite unlikely given the number and type of variables we are able to control for in our analysis and given that none of those alone would explain that much of the variance.

To sum up, as reported in table 13 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.

[36] Starting from the estimation of the mediator and the outcome model, we calculated the $\tilde{R}^2_M$ and $\tilde{R}^2_Y$ for maternal education, which are 0.56 and 0.04, respectively.

9 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 thorough sensitivity analyzes augment the credibility of our identifying assumptions.

Our analysis is based on the 2011 module on intergenerational transmission of EU-SILC data. Our results show that, on average, over the 28 European countries considered, growing up in financial distress leads to a significant 1.4 times higher probability of being at risk of poverty in adulthood, and to a significant decrease of 4% in the adult equivalent income (circa 610 €on average), in our most conservative estimates. Our results are particularly relevant for Mediterranean and Central and Easter European Countries.

Our estimates are robust to the presence of unobservable variables bias. We find that in order to drive our effects to zero: i) the explanatory power of the unobservables would need to be 20 times as important as the one of the observables, following the method proposed by Oster (2016); ii) parental unobservable ability would have to increase the parents' probability of experiencing financial problems by 20% (in our most restrictive case) even conditional on all the other observable parental characteristics, following the method developed by Rosenbaum (2002); and iii) the odds of experiencing poverty would need to be almost 5 times higher for high ability parents than for low ability ones and the odds of being at risk of poverty in adulthood almost 5 times higher for non-poor children of high ability parents than for those of low ability parents, following the approach suggested by Ichino et al. (2008).

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. Achieving (at least) a secondary level of education does not completely overcome the detrimental impact of childhood poverty.

Moreover, experiencing poverty during childhood will more likely translate into an exclusion from further education, as the probability of completing at least secondary education is 1.5 times higher for children who were not affected by financial problems. We find that education plays a substantial role as intermediate variable in the causal effect estimation of childhood poverty on income and poverty risk in adulthood, accounting for more than 30% of the total effect.

Some policy implications can be derived by our study. Our analysis reinforces the need for policies devoted to eliminate the source at the basis of this increased risk, reducing childhood poverty. Our results are also 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).

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. Among others, 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 (Heckman et al., 2006). Reducing the stress and anxiety of the parents, targeting intensive health, nutrition and care supports on particularly deprived households or areas might also be 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.

Aizer, A., Eli, S., Ferrie, J., and Lleras-Muney, A. (2016). The long-run impact of cash transfers to poor families. *American Economic Review*, 106(4):935–71.

Akerlof, G. and Yellen, J. (1985). Unemployment through the filter of memory. *The Quarterly Journal of Economics*, 100.(3):747–773.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy*, 113(1):151–184.

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 (2017). *Monitoring Social Inclusion in Europe*. Eurostat Statistical. Forthcoming.

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

Bellani, L. and Bia, M. (2017). The impact of growing up poor in europe. In Atkinson, T., Guio, A.-C., and Marlier, E., editors, *Monitoring Social Inclusion in Europe*. Eurostat Statistical. Forthcoming.

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. and Gibbons, S. (2006). The persistence of poverty across generations. Technical report, The Policy Press.

Blanden, J., Gregg, P., and Macmillan, L. (2007). Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal*, 117:C43–C60.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics*, 129(4):1553–1623.

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 A*, 167:69–01.

Esping-Andersen, G. (1990). *The Three Worlds of Welfare Capitalism.* Cambridge: Polity Press.

Eurostat (2016). Children were the age group at the highest risk of poverty or social exclusion in 2014. Statistics Explained 3/2016, Eurostat.

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.

Gal, J. (2010). Is there an extended family of mediterranean welfare states? *Journal of European Social Policy*, 20(4):283–300.

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.

Hafeman, D. and Schwartz, S. (2009). Opening the black box: a motivation for the assessment of mediation. *International Journal of Epidemiology*, 38:838–845.

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

Hardt, J. and Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *Journal of Child Psychology and Psychiatry*, 45(2):260–273.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from kenya. *The Quarterly Journal of Economics*.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482.

Hemerijck, A. (2012). *Changing Welfare States*. Oxford University Press.

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

Hill, J., Waldfogel, J., and Brooks-Gunn, J. (2003). Sustained effects of high participation in an early intervention for low-birth-weight premature infants. *Developmental Psychology*, 38:730–744.

Hong, G. (2010). Ratio of mediator probability weighting for estimating natural direct and indirect effects. *2010 Proceedings of the American Statistical Association, Biometrics Section*, pages 2401–2415.

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

Huber, M., Lechner, M., and Mellace, M. (2016). The finite sample performance of estimators for mediation analysis under sequential conditional independence. *Journal of Business and Economic Statistics*, 34:139–160.

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 Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological Methods*, 15:309–334.

Imai, K. L., Keele, L., and Yamamoto, T. (2010b). 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.

Jo, B. (2008). Causal inference in randomized experiments with mediational processes. *Psychological Methods*, 13(1):314–336.

Jo, B., Stuart, E., MacKinnon, D., and Vinokur, A. (2011). The use of propensity scores in mediation analysis. *Multivariate Behavioral Research*, 46:425–452.

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.

Jürges, H. (2007). Unemployment, life satisfaction, and retrospective error. *Journal of the Royal Statistical Society. Series A*, 170(1):43–61.

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.

King, G., Tomz, M., and Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science*, 44:341–355.

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.

Linden, A. and Karlson, K. B. (2013). Using mediation analysis to identify causal mechanisms in disease management interventions. *Health Services and Outcomes Research Methodology*, 13(2):86–108.

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.

Oster, E. (2016). Unobservable selection and coefficient stability: Theory and validation. *Journal of Business Economics and Statistics*.

Pearl, J. (2001a). The causal mediation formula - a giude to the assessment of pathways and mechanisms. *Technical Report, R-379, UCLA Cognitive Systems Laboratory*, pages 411–420.

Pearl, J. (2001b). Direct indirect effects. *In Proceedings of the Seventeenth Conference in Uncertainty in Artificial Intelligence. San Francisco, Morgan Kaufman*, pages 411–420.

Peterson, M., Sinisi, S., and van der Laan, M. (2006). Estimation of direct causal effects. *Epidemiology*, pages 276–284.

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.

Robins, L. N., Schoenberg, S. P., Holmes, S. J., Ratcliff, K. S., Benham, A., and Works, J. (1985). Early home environment and retrospective recall: A test for concordance between siblings with and without psychiatric disorders. *American Journal of Orthopsychiatry*, 55(1):27–41.

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.

VanderWeele, T. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*, 20:18–26.

A Online Appendix

A.1 Additional Tables and Figures

Table A1.1: Overview of present EU Member States (EU 28)

Austria	(AT)	Estonia	(EE)	Italy	(IT)	Portugal	(PT)
Belgium	(BE)	Finland	(FI)	Latvia	(LV)	Romania	(RO)
Bulgaria	(BG)	France	(FR)	Lithuania	(LT)	Slovakia	(SK)
Croatia	(HR)	Germany	(DE)	Luxembourg	(LU)	Slovenia	(SI)
Cyprus	(CY)	Greece	(EL)	Malta	(MT)	Spain	(ES)
Czech Republic	(CZ)	Hungary	(HU)	Netherlands	(NL)	Sweden	(SE)
Denmark	(DK)	Ireland	(IE)	Poland	(PL)	United Kingdom	(UK)

Note: In 2011, only Croatia was not yet a member.

Table A1.2: T-test on the missing

	(1) All			(2) Poor			(3) Non Poor		
	Non Missing	Missing	Diff	Non Missing	Missing	Diff	Non Missing	Missing	Diff
Income	4.03	3.82	0.21***	3.96	3.82	0.14***	4.04	3.82	0.22***
Poor as adult	0.14	0.13	0.00	0.21	0.13	0.08**	0.13	0.13	-0.00
At least secondary education	0.79	0.89	-0.10***	0.60	0.89	-0.29***	0.81	0.89	-0.07**

(1) Compares Missing with all the Non Missing, (2) with only the Poor and (3) with only the Non Poor. N. of Missing: 222, Non Missing:111986, of which 11997 poor.

Table A1.3: T-test on the outcome and intermediate variables means by treatment status

	Non poor	Poor	Diff
Income	4.03	3.95	0.08***
Poor as adult	0.13	0.21	-0.08***
At least secondary education	0.82	0.59	0.22***

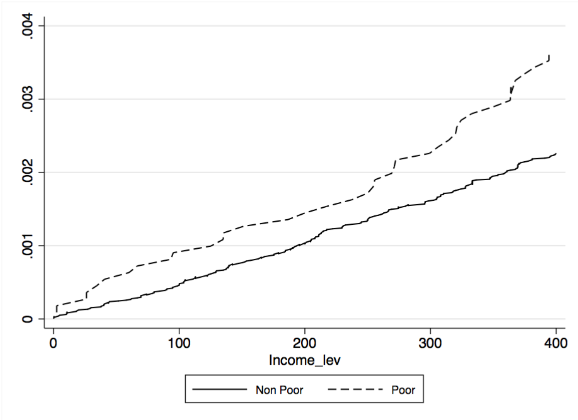
Table A1.4: Propensity Score Estimation-M.E.

	(1) Probit	(2) Logit	(3) Linear Probability
quarter of birth	-0.001	-0.001	-0.001
	[-0.002,0.001]	[-0.002,0.001]	[-0.002,0.001]
year of birth	-0.001	-0.001	-0.000
	[-0.001,-0.000]	[-0.001,-0.000]	[-0.001,0.000]
sex	-0.004	-0.004	-0.004
	[-0.007,-0.000]	[-0.007,-0.000]	[-0.008,-0.001]
n. of adult in hh	0.013	0.013	0.018
	[0.012,0.015]	[0.011,0.015]	[0.016,0.021]
n. of children in hh	0.026	0.025	0.034
	[0.024,0.027]	[0.024,0.026]	[0.032,0.036]
n. of person in work	-0.007	-0.007	-0.009
	[-0.009,-0.005]	[-0.010,-0.005]	[-0.012,-0.006]
year of birth of father	0.001	0.001	0.001
	[0.000,0.001]	[0.000,0.001]	[0.001,0.002]
year of birth of mother	-0.002	-0.002	-0.002
	[-0.002,-0.001]	[-0.002,-0.001]	[-0.003,-0.002]
Single parent	0.124	0.119	0.158
	[0.116,0.132]	[0.112,0.127]	[0.145,0.171]
country_of_birth==EU	0.010	0.010	0.006
	[-0.002,0.021]	[-0.001,0.022]	[-0.006,0.019]
country_of_birth==OTH	0.002	0.001	0.004
	[-0.008,0.012]	[-0.009,0.010]	[-0.007,0.015]
father not born in country of residence	0.012	0.012	0.013
	[0.002,0.021]	[0.002,0.022]	[0.004,0.022]
mother not born in country of residence	0.005	0.004	0.005
	[-0.005,0.015]	[-0.006,0.014]	[-0.004,0.014]
father primary education	0.055	0.057	0.047
	[0.049,0.060]	[0.051,0.062]	[0.043,0.051]
mother primary education	0.047	0.052	0.030
	[0.042,0.053]	[0.046,0.058]	[0.026,0.034]
Tenancy_status== Owner	-0.055	-0.056	-0.061
	[-0.059,-0.052]	[-0.060,-0.052]	[-0.065,-0.056]
father self-Employed	-0.098	-0.091	-0.157
	[-0.108,-0.087]	[-0.101,-0.081]	[-0.176,-0.137]
father Employee	-0.095	-0.088	-0.153
	[-0.105,-0.085]	[-0.097,-0.079]	[-0.172,-0.134]
mother self-Employed	0.007	0.007	0.009
	[0.001,0.013]	[0.001,0.013]	[0.001,0.017]
mother Employee	-0.021	-0.021	-0.021
	[-0.026,-0.016]	[-0.026,-0.016]	[-0.026,-0.015]
Country Fixed Effect	Yes	Yes	Yes
Observations	108355	108355	108355

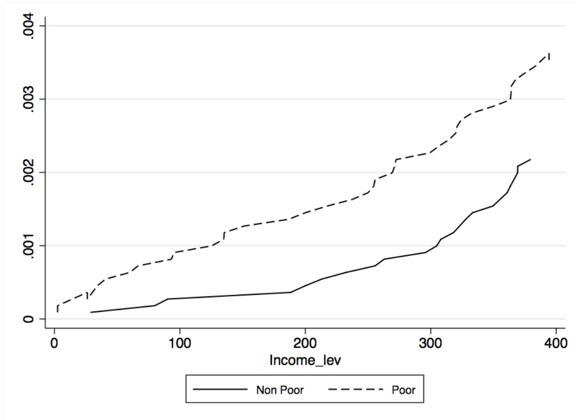
95% confidence intervals in brackets

Table A1.5: Balancing Table

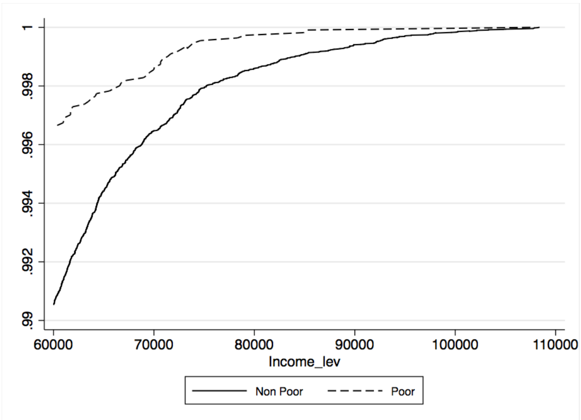
	Mean			Variance			Skewness			Std difference	
	treated	control	weighted cont	treated	control	weighted cont	treated	control	weighted cont	treat&control	treat& weighted cont
quarter of birth	1.9	1.9	1.9	2.6	2.7	2.6	-.48	-.45	-.48	.049	5.7e-05
year of birth	1964	1966	1964	34	35	34	.3	.011	.31	-.25	.0078
sex	1.5	1.5	1.5	.25	.25	.25	-.093	-.13	-.092	-.016	7.0e-05
n. of adult in hh	2.7	2.5	2.7	1.5	.88	1.5	1.4	1.9	1.5	.18	1.2e-04
n. of children in hh	2.9	2.3	2.9	2.3	1.4	2.3	.77	1.1	.83	.4	1.1e-04
n. of person in work	1.8	1.9	1.8	1	.74	1	1.4	1.6	1.6	-.098	2.8e-05
year of birth of father	1934	1936	1934	72	73	72	-.079	-.21	-.076	-.21	.0052
year of birth of mother	1937	1939	1937	69	68	69	.075	-.15	.052	-.23	.0054
Single parent	.081	.031	.081	.075	.03	.075	3.1	5.4	3.1	.19	-1.9e-05
country_of_birth==EU	.04	.034	.04	.038	.033	.038	4.7	5.2	4.7	.03	-1.6e-06
country_of_birth==OTH	.071	.047	.071	.066	.045	.066	3.3	4.3	3.3	.092	9.6e-07
father not born in country of residence	.13	.1	.13	.12	.093	.12	2.2	2.6	2.2	.085	-1.1e-05
mother not born in country of residence	.13	.1	.13	.11	.092	.11	2.2	2.6	2.2	.08	-1.1e-05
father primary education	.83	.6	.83	.14	.24	.14	-1.8	-.42	-1.8	.61	3.8e-04
mother primary education	.89	.67	.89	.1	.22	.1	-2.4	-.71	-2.4	.68	5.3e-04
Tenancy_status== Owner	.61	.74	.61	.24	.19	.24	-.47	-1.1	-.47	-.26	2.3e-05
father self-Employed	.22	.18	.22	.17	.15	.17	1.4	1.7	1.4	.093	6.0e-05
father Employee	.73	.81	.73	.2	.16	.2	-1	-1.6	-1	-.17	-2.6e-05
mother self-Employed	.14	.1	.14	.12	.09	.12	2.1	2.7	2.1	.12	5.1e-05
mother Employee	.33	.5	.33	.22	.25	.22	.72	-.015	.72	-.37	-1.6e-04
country==AT	.049	.032	.049	.047	.031	.047	4.2	5.3	4.2	.078	-6.4e-06
country==BE	.022	.029	.022	.021	.028	.021	6.6	5.6	6.6	-.053	-1.1e-05
country==BG	.012	.036	.012	.012	.035	.012	8.9	5	8.9	-.22	-1.9e-04
country==CY	.054	.021	.054	.051	.02	.051	4	6.7	4	.15	2.8e-05
country==CZ	.023	.036	.023	.022	.035	.022	6.4	5	6.4	-.091	-8.5e-06
country==DE	.04	.054	.04	.038	.051	.038	4.7	3.9	4.7	-.073	-3.0e-05
country==DK	.007	.014	.007	.0069	.014	.0069	12	8.2	12	-.089	-3.4e-05
country==EE	.0088	.025	.0088	.0087	.024	.0087	11	6.1	11	-.17	-1.6e-04
country==EL	.037	.023	.037	.036	.023	.036	4.9	6.3	4.9	.073	2.5e-05
country==ES	.081	.075	.081	.075	.069	.075	3.1	3.2	3.1	.022	5.0e-05
country==FR	.047	.05	.047	.045	.048	.045	4.3	4.1	4.3	-.018	2.2e-05
country==HR	.057	.024	.057	.054	.024	.054	3.8	6.2	3.8	.14	2.1e-05
country==HU	.064	.069	.064	.06	.064	.06	3.6	3.4	3.6	-.02	7.1e-06
country==IE	.011	.015	.011	.011	.014	.011	9.4	8.1	9.4	-.034	-8.0e-06
country==IT	.099	.11	.099	.089	.1	.089	2.7	2.4	2.7	-.048	4.4e-05
country==FI	.0075	.016	.0075	.0075	.016	.0075	11	7.7	11	-.1	-3.8e-05
country==LU	.032	.034	.032	.031	.033	.031	5.3	5.1	5.3	-.01	-2.6e-06
country==LV	.011	.03	.011	.011	.029	.011	9.2	5.6	9.2	-.17	-1.2e-04
country==MT	.021	.02	.021	.021	.02	.021	6.7	6.8	6.7	.0045	1.7e-05
country==NL	.012	.031	.012	.012	.03	.012	9	5.4	9	-.17	-1.2e-04
country==PL	.058	.068	.058	.055	.063	.055	3.8	3.4	3.8	-.041	-1.6e-05
country==PT	.082	.025	.082	.075	.024	.075	3.1	6.1	3.1	.21	5.5e-05
country==RO	.053	.032	.053	.05	.031	.05	4	5.3	4	.092	3.8e-05
country==SE	.0072	.011	.0072	.0071	.011	.0071	12	9.4	12	-.045	-2.1e-05
country==SI	.044	.022	.044	.042	.021	.042	4.4	6.6	4.4	.11	2.0e-05
country==SK	.023	.038	.023	.022	.037	.022	6.4	4.8	6.4	-.1	-3.9e-05
country==UK	.019	.03	.019	.018	.029	.018	7.1	5.6	7.1	-.08	-1.4e-05



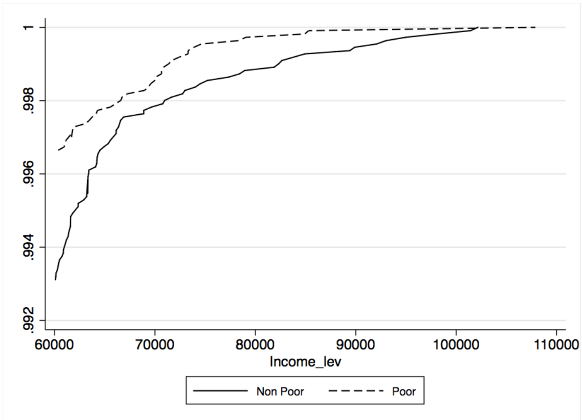
(a) Without Matching



(b) Matched



(c) Without Matching



(d) Matched

Figure A1.1: Cumulative Distribution at the bottom and top of the distribution.

Table A1.6: Classification of Countries

Welfare State Regime	Countries
Continental	Austria, Belgium, Germany, France, Luxembourg, Netherlands
Social Democratic	Denmark, Finland, Sweden
Central & Eastern European	Bulgaria, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Romania, Slovakia, Croatia, Slovenia.
Mediterranean	Cyprus, Greece, Italy, Malta, Spain, Portugal
Liberal	Ireland, United Kingdom

Note: See Hemerijck (2012) for a description of this classification and refer to Gal (2010) for the inclusion of Malta and Cyprus in the Mediterranean welfare state.

Table A1.7: T-test on the outcome and intermediate variables means by treatment status

	Continental	Social Democratic	CEE	Mediterranean	Liberal
Income	0.06	0.04	0.07	0.12	0.04
	[0.05,0.06]	[0.02,0.07]	[0.06,0.08]	[0.11,0.13]	[0.01,0.07]
Poor as adult	-0.06	-0.04	-0.09	-0.07	-0.05
	[-0.08,-0.05]	[-0.07,-0.00]	[-0.11,-0.08]	[-0.09,-0.06]	[-0.09,-0.02]

95% confidence intervals in brackets

Table A1.8: Propensity Score Estimation-M.E., by welfare state.

	Continental	Social Democratic	CEE	Mediterranean	Liberal
quarter of birth	-0.002	0.001	0.001	-0.001	
	[-0.006,0.001]	[-0.005,0.007]	[-0.002,0.003]	[-0.004,0.003]	
year of birth	-0.001	-0.002	-0.000	-0.001	0.002
	[-0.002,0.000]	[-0.004,0.000]	[-0.001,0.001]	[-0.002,-0.000]	[0.000,0.004]
sex	0.002	0.017	-0.003	-0.013	-0.001
	[-0.005,0.009]	[0.004,0.030]	[-0.008,0.002]	[-0.020,-0.006]	[-0.015,0.013]
n. of adult in hh	0.010	0.011	0.012	0.019	-0.000
	[0.006,0.015]	[0.001,0.021]	[0.010,0.015]	[0.015,0.022]	[-0.007,0.007]
n. of children in hh	0.020	0.021	0.025	0.032	0.017
	[0.017,0.022]	[0.015,0.026]	[0.023,0.028]	[0.029,0.035]	[0.011,0.024]
n. of person in work	-0.004	-0.000	-0.019	-0.004	-0.006
	[-0.009,0.001]	[-0.009,0.008]	[-0.024,-0.014]	[-0.008,-0.000]	[-0.012,0.000]
year of birth of father	0.000	-0.000	0.000	0.001	0.000
	[-0.001,0.001]	[-0.002,0.001]	[-0.000,0.001]	[0.000,0.002]	[-0.002,0.002]
year of birth of mother	-0.000	0.001	-0.002	-0.003	-0.001
	[-0.001,0.001]	[-0.001,0.003]	[-0.002,-0.001]	[-0.004,-0.002]	[-0.003,0.001]
Single parent	0.121	0.112	0.100	0.121	0.100
	[0.108,0.133]	[0.084,0.140]	[0.088,0.112]	[0.093,0.149]	[0.077,0.123]
father not born in country of residence	0.008	0.005	0.002	0.054	0.001
	[-0.007,0.022]	[-0.053,0.063]	[-0.014,0.017]	[0.025,0.084]	[-0.030,0.033]
mother not born in country of residence	-0.001	0.011	0.003	0.015	0.008
	[-0.016,0.014]	[-0.043,0.065]	[-0.013,0.019]	[-0.014,0.043]	[-0.025,0.040]
father primary education	0.037	0.029	0.049	0.094	0.024
	[0.028,0.046]	[0.015,0.044]	[0.041,0.056]	[0.079,0.108]	[0.007,0.041]
mother primary education	0.046	0.006	0.046	0.077	0.013
	[0.037,0.056]	[-0.009,0.021]	[0.038,0.054]	[0.058,0.096]	[-0.005,0.032]
Tenancy_status== Owner	-0.043	-0.049	-0.039	-0.082	-0.074
	[-0.050,-0.036]	[-0.065,-0.034]	[-0.046,-0.033]	[-0.090,-0.075]	[-0.089,-0.058]
father self-Employed	-0.103	-0.062	-0.048	-0.130	-0.070
	[-0.125,-0.080]	[-0.113,-0.011]	[-0.066,-0.031]	[-0.151,-0.109]	[-0.111,-0.029]
father Employee	-0.101	-0.060	-0.076	-0.110	-0.089
	[-0.122,-0.080]	[-0.110,-0.011]	[-0.091,-0.061]	[-0.130,-0.090]	[-0.127,-0.050]
mother self-Employed	-0.009	-0.020	-0.008	0.021	-0.001
	[-0.022,0.005]	[-0.047,0.007]	[-0.020,0.004]	[0.009,0.032]	[-0.032,0.030]
mother Employee	-0.011	-0.035	-0.032	-0.004	-0.014
	[-0.020,-0.002]	[-0.055,-0.016]	[-0.040,-0.025]	[-0.014,0.006]	[-0.030,0.003]
Observations	24733	4282	43618	31099	4623

95% confidence intervals in brackets

Table A1.9: Rosenbaum bounds for Income

Γ	Wilcoxon's test significance level		Hodges-Lehmann point estimate	
	upper bound	lower bound	upper bound	lower bound
1	0	0	-.050	-.050
1.1	0	0	-.075	-.025
1.2	0	.30	-.098	-.003
1.3	0	.99	-.119	.017
1.4	0	1	-.138	.036
1.5	0	1	-.156	.054
1.6	0	1	-.172	.071
1.7	0	1	-.188	.087
1.8	0	1	-.203	.101
1.9	0	1	-.217	.115
2	0	1	-.230	.128

Note: Γ = odds of differential assignment due to unobserved factors.

Table A1.10: Mantel-Haenszel bounds for Poverty

Γ	Test statistic		Significance level	
	over-estimation	under-estimation	over-estimation	under-estimation
1	10	10	0	0
1.1	7	12	0	0
1.2	5	15	0	0
1.3	3	17	0	0
1.4	1	19	.11	0
1.5	.63	21	.26	0
1.6	2	0	0	0
1.7	3	24	0	0
1.8	5	26	0	0
1.9	6	28	0	0
2	8	29	0	0

Note: Γ = odds of differential assignment due to unobserved factors.

Table A1.11: Mantel-Haenszel bounds for Secondary Education

Γ	Test statistic		Significance level	
	over-estimation	under-estimation	over-estimation	under-estimation
1	18	18	0	0
1.1	15	21	0	0
1.2	12	24	0	0
1.3	9	27	0	0
1.4	7	29	0	0
1.5	4	31	0	0
1.6	2	34	.002	0
1.7	.82	36	.20	0
1.8	1	38	.15	0
1.9	2	40	.002	0
2	4	41	0	0

Note: Γ = odds of differential assignment due to unobserved factors. For ease of interpretation we redefined the outcome of interest as having less than secondary education.