Intergenerational poverty transmission in Europe: the role of education

Luna Bellani, Michela Bia

Kostanz University, LISER

luna.bellani@uni-konstanz.de, michela.bia@liser.lu

10 December 2020

Outline

Why should we care?

In the US, roughly 1 in 5 children live in poverty compared with 1 in 8 adults (US Census Bureau 2014).

About 2 out of every 5 children spent at least one year in poverty before turning 18.

Persistently poor children are then 43% less likely to finish college than non poor children, and 13% less likely to finish high school (Ratcliffe, 2015)

Why should we care?
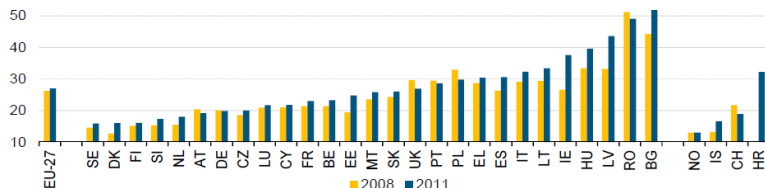
Note: For 2008: HR data is not available. For 2011: EU-27 Eurostat estimation, IE data is for 2010.

Figure: Children at risk of poverty or social exclusion, (%), 2008 and 2011. Source: Eurostat

Why should we care?

Recent research on intergenerational mobility agrees that growing up in a poor family raises the probability of falling below the poverty threshold in adulthood.

The key contentious question for policy here is whether this association is truly causal.

Related Contributions

- Level estimates (see Jenkins and Siedler, 2007 for a survey)
- Instrumental variables approach (IV) (e.g. Shea, 2000)
- Models exploiting sibling differences (see among others Ermisch et al., 2004)

Results are 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.

Related Contributions

- Level estimates (see Jenkins and Siedler, 2007 for a survey)
- Instrumental variables approach (IV) (e.g. Shea, 2000)
- Models exploiting sibling differences (see among others Ermisch et al., 2004)
- Decomposition of the mobility coefficient conditional on some mediating variables (Blanden et al., 2007)

Cognitive and non-cognitive skills accounting for 20% and 10% of intergenerational earning persistence in the UK.

Our Contribution

Because of the complexity of the process, different statistical techniques have been used, each of which relies on a different set of assumptions.

i) We apply a potential outcomes approach for causal inference to quantify the impact of experiencing financial difficulties while growing up and we provide a sensitivity analysis on the unobserved parental ability;

ii) we analyze the channel of this poverty transmission, introducing individual human capital accumulation as an intermediate variable and we provide a sensitivity analysis on unobserved confounders also for this mediation analysis.

On average, over the 27 European countries considered, growing up poor leads to an increase of the risk of being poor and to a decrease in the adult equivalent income.

Moreover, we find that experiencing poverty during childhood will more likely translate into an exclusion from secondary education and that education accounts for a substantial share of the total effect on adult income.

Estimation Strategy (1/4)

We consider a set of N individuals, and denote each of them by subscript i: $i = 1, \ldots, N$. Let T_i indicate whether a child was growing up in a poor household, $T_i = 1$, or not, $T_i = 0$.

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.

We consider a set of N individuals, and denote each of them by subscript i: $i = 1, \ldots, N$. Let T_i indicate whether a child was growing up in a poor household, $T_i = 1$, or not, $T_i = 0$.

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.

Central assumption: "assignment to treatment" is unconfounded given the set of observable variables: $Y_i(0) \perp T_i | X_i$.

Estimation Strategy (1/4)

We consider a set of N individuals, and denote each of them by subscript i: $i = 1, \ldots, N$. Let T_i indicate whether a child was growing up in a poor household, $T_i = 1$, or not, $T_i = 0$.

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.

Central assumption: "assignment to treatment" is unconfounded given the set of observable variables: $Y_i(0) \perp T_i | X_i$.

Rosenbaum and Rubin (1983) show that if the potential outcome $Y_i(0)$ is unconfounded given X_i, it also independent conditional on the propensity score $p(x) = Pr(T = 1 | X = x)$: $Y_i(0) \perp T_i | p(X_i)$.

Estimation Strategy (2/4)

We apply a propensity score based method to select a control group of non-treated individuals (non poor as a child) who are very similar to treated individuals conditional on a set of observable characteristics.

The matched samples of poor and non-poor children will be used to assess the average impacts of being poor as a child on adulthood outcomes.

Formally, given the population of units i, if we know $p(X_i)$, we can estimate the average effect of being poor on the population (ATT):

$$\tau = E[Y_{1i} - Y_{0i}|T_i = 1] =$$

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

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

We also consider a causal mediation analysis (Hicks and Tingley, 2011).

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 $M = m$.

The average causal mediation effect on the treated (ACMET) is defined as:
$$\bar{\tau}_t = E[Y_i(t, M_i(1)) - Y_i(t, M_i(0))|T_i = 1] .$$

Similarly, the average direct effect among the treated (ADET) as:
$$\bar{\gamma}_t = E[Y_i(1, M_i(t)) - Y_i(0, M_i(t))|T_i = 1].$$

Estimation Strategy (4/4)

In our study the mediating variable is binary and a probit model is used:

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

where

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

The outcome variable is continuous and a linear regression model is implemented:

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

The error terms are *iid*, $\epsilon_{i2} \sim \mathcal{N}(0, 1)$, $\epsilon_{i3} \sim \mathcal{N}(0, \sigma_3^2)$ and

$(\epsilon_{i2}, \epsilon_{i3}) \sim \mathcal{N}(0, \rho\sigma_3^2)$, where ρ is the correlation between the two error terms.

Variables Description

The analysis is based on the module on intergenerational transmission of 2011 of the European Union Statistics on Income and Living Conditions (EU-SILC).

Our treatment variable is based on:

- the financial situation of the household (very bad or bad)

The outcomes in adulthood that we are interested in are

- the equivalized income,
- the probability of being poor and
- the highest level of education attained (*intermediate outcome*).

Our pre-treatment variables are:

- the country of residence and the country of birth,
- the year and quarter of birth,
- family composition,
- year of birth, highest level of education, main activity and main occupation of the father and of the mother,
- number of siblings.
- n. of adult in the household,
- n. of persons in the household in work,
- country of birth, citizenship and managerial position of the father and of the mother.

Average Treatment Effect

Table: Average Treatment on the Treated

	Main Oucomes		Intermediate Outcome
	Income (in %)	Poverty (in pp)	Education (in pp)
Propensity score matching	-5.7	5.7	-11.9
	[-7.1, -4.3]	[4.3, 7.1]	[-13.4, -10.4]
Doubly-Robust Estimation	-4.9	5.2	-11.3
	[-5.4, -4.4]	[4.4, 6.0]	[-12.2, -10.4]
N.	108355	108355	107566

Note: 95% Confidence Intervals in brackets. pp=percentage points.

What is the problem?

What is the problem?

Consider for example unobserved parental genetic ability (A).

Sensitivity Analysis

What is the problem?

Consider for example unobserved parental genetic ability (A).

if $Pr(T = 1|X, A) \neq Pr(T = 1|X)$

Sensitivity Analysis

What is the problem?

Consider for example unobserved parental genetic ability (A).

if $Pr(T = 1|X, A) \neq Pr(T = 1|X) \Rightarrow$ selection into treatment

Sensitivity Analysis

What is the problem?

Consider for example unobserved parental genetic ability (A).

if $Pr(T = 1|X, A) \neq Pr(T = 1|X) \Rightarrow$ selection into treatment

if $Pr(Y = 1|T, X, A) \neq Pr(Y = 1|T, X)$

Sensitivity Analysis

What is the problem?

Consider for example unobserved parental genetic ability (A).

if $Pr(T = 1|X, A) \neq Pr(T = 1|X) \Rightarrow$ selection into treatment

if $Pr(Y = 1|T, X, A) \neq Pr(Y = 1|T, X) \Rightarrow$ outcome

Sensitivity Analysis I (Rosenbaum, 2002)

Table: Rosenbaum bounds for Income

Γ	Wilcoxon's test significance level upper bound	lower bound	Hodges-Lehmann point estimate upper bound	lower bound
1	0	0	-.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

Γ = log odds of differential assignment due to unobserved factors

Sensitivity Analysis II (Rosenbaum, 2002)

Table: 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

Γ = odds of differential assignment due to unobserved factors

Let us assume that parental unobserved ability (A) can be expressed as a binary variable taking value 1=High Ability, 0=Low Ability.

Sensitivity Analysis I (Ichino et al., 2008)

Let us assume that parental unobserved ability (A) can be expressed as a binary variable taking value 1=High Ability, 0=Low Ability.

- Step 1- Characterize the distribution of the binary confounding factor A

The distribution of the binary confounding factor A can be fully characterized by the choice of four parameters:

  I) $p_{11} \equiv Pr(A = 1 | T = 1, Y = 1) = Pr(A = 1 | T = 1, Y = 1, X)$

  II) $p_{10} \equiv Pr(A = 1 | T = 1, Y = 0) = Pr(A = 1 | T = 1, Y = 0, X)$

  III) $p_{01} \equiv Pr(A = 1 | T = 0, Y = 1) = Pr(A = 1 | T = 0, Y = 1, X)$

  IV) $p_{00} \equiv Pr(A = 1 | T = 0, Y = 0) = Pr(A = 1 | T = 0, Y = 0, X)$

▸ p_{ij}

Sensitivity Analysis I (Ichino et al., 2008)

Let us assume that parental unobserved ability (A) can be expressed as a binary variable taking value 1=High Ability, 0=Low Ability.

- Step 1 Characterize the distribution of the binary confounding factor A

- Step 2 A value of A is attributed to each subject, according to the p_{ij} chosen before.

Sensitivity Analysis I (Ichino et al., 2008)

Let us assume that parental unobserved ability (A) can be expressed as a binary variable taking value 1=High Ability, 0=Low Ability.

- Step 1 Characterize the distribution of the binary confounding factor A

- Step 2 A value of A is attributed to each subject, according to the p_{ij} chosen before.

- Step 3 The estimation of the propensity score and the ATT is performed including the simulated A in the set of confounders.

Sensitivity Analysis: Summary

The impacts would remain significant even if parental unobservable ability alone were to increase the probability of experiencing financial problems of about 15% and 30%, for income and poverty respectively.

To drive our results to zero:

- The odds of experiencing financial problems should be 5 times higher for high ability parents than for low ability ones.

- The odds of being at risk of poverty in adulthood 4.8 times higher for non-poor children of high ability parents than the ones of low ability parents.
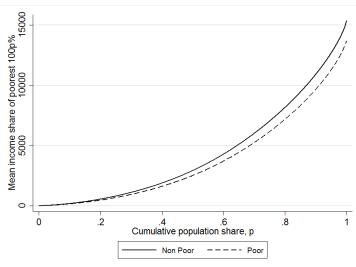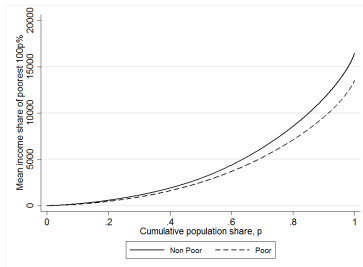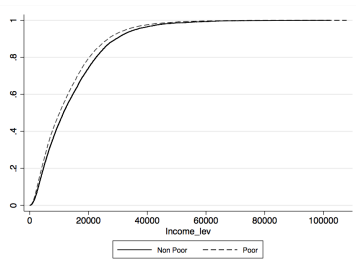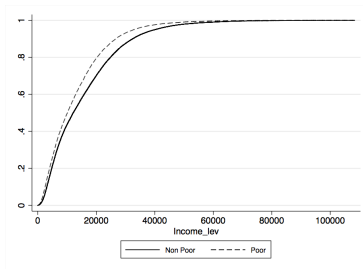
Distributional Effect

The Role of Education

Table: Mediation Causal Analysis on the Treated

	Income (in %)	Poverty (in pp)
Total Effect	-5.1	5.6
	[-5.8, -4.5]	[4.6, 6.6]
Average Causal Mediation Effect	-1.6	1.8
	[-1.8, -1.5]	[1.6, 2.1]
Direct Effect	-3.5	3.8
	[-4.1, -2.9]	[2.7, 4.7]
% of Tot.Eff. Mediated	32	33
	[29, 37]	[28, 40]

Note: 95% Confidence Intervals in brackets. pp=percentage points.

The Role of Education

Table: Mediation Causal Analysis on the Treated: poor and non poor

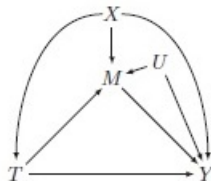|	Income (in %)	Poverty (in pp)
Poor		
Causal Mediation Effect	-1.7	2.0
	[-2.0, -1.5]	[1.7, 2.2]
Direct Effect	-3.6	3.9
	[-4.3, -3.0]	[2.8, 4.9]
% of Tot.Eff. Mediated	34	35
	[30, 39]	[30, 43]
Non Poor		
Causal Mediation Effect	-1.5	1.7
	[-1.7, -1.4]	[1.4, 2.0]
Direct Effect	-3.4	3.6
	[-4.0, -2.8]	[2.7, 4.56]
% of Tot.Eff. Mediated	30	30
	[27, 34]	[26, 37]

What is the problem?

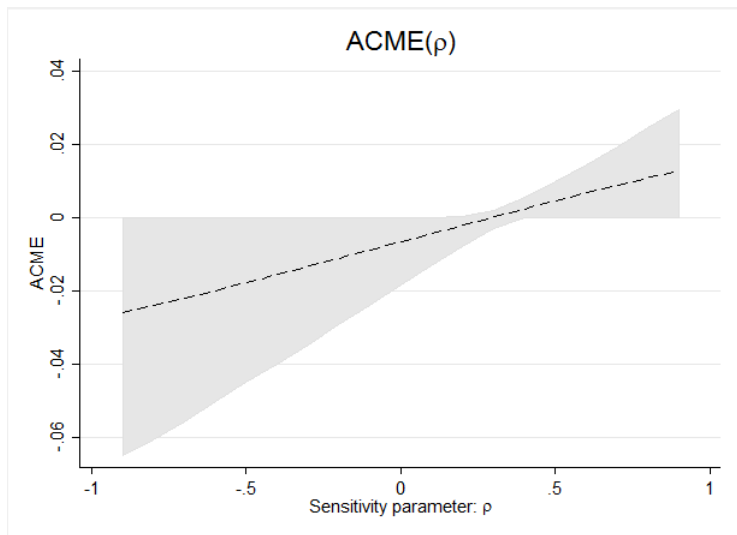# Sensitivity Analysis

What is the problem?
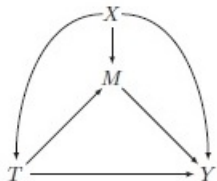
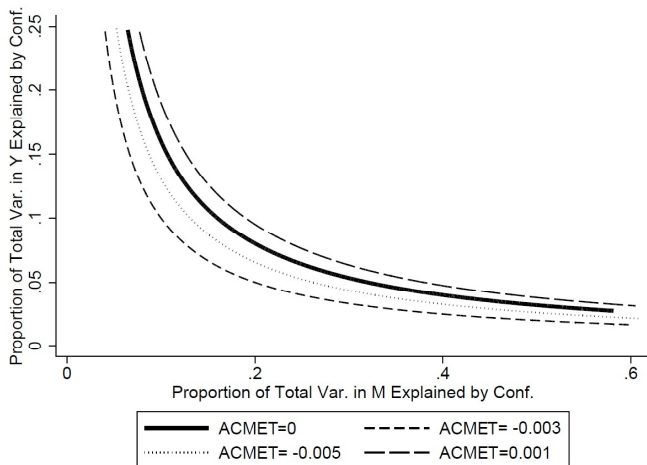(a) Example with observed covariate.   (b) Example with unobserved covariate.

E.g. 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.

Sensitivity Analysis based on ρ

Sensitivity Analysis based on R^2

Sensitivity Analysis: Summary

Table: Sensitivity results for the ACMET

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

Values of ρ different from 0 lead to violations of the Sequential Ignorability Assumption
$\tilde{R}_M^2$ and $\tilde{R}_Y^2$ are proportion of original variance that is explained by the omitted confounder in the mediator and outcome model.
R_M^{2*} and R_Y^{2*} are proportion of unexplained variance that is explained by the omitted confounder in the mediator and outcome model.

Preexisting cognitive and non-cognitive problems should explain around 14% of the variances for both of education and income for the true ACME to be 0.

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

The proportion of original variance that is explained by the omitted confounder in the mediator and outcome model must be substantially higher for the original conclusion to be changed.

Conclusions

We provide a quantitative assessment of the causal effect of poverty in childhood at the European level, performing a series of robustness checks on the potential outcome approach chosen.

⇒ -5% Disposable equivalent income
⇒ +4 p.p. Probability of being poor
⇒ -12 p.p. Probability of attaining Secondary Education

Conclusions

We provide a quantitative assessment of the causal effect of poverty in childhood at the European level, performing a series of robustness checks on the potential outcome approach chosen.

⇒ -5% Disposable equivalent income
⇒ +4 p.p. Probability of being poor
⇒ -12 p.p. Probability of attaining Secondary Education

We show that higher social welfare is reached in an hypothetical society where no one experiences poverty as a child against one where childhood poverty is common, even after achieving (at least) a secondary level of education.

Conclusions

We provide a quantitative assessment of the causal effect of poverty in childhood at the European level, performing a series of robustness checks on the potential outcome approach chosen.

⇒ -5% Disposable equivalent income
⇒ +4 p.p. Probability of being poor
⇒ -12 p.p. Probability of attaining Secondary Education

We show that higher social welfare is reached in an hypothetical society where no one experiences poverty as a child against one where childhood poverty is common, even after achieving (at least) a secondary level of education.

We analyze the impact of the interplay between growing up in poverty and individual human capital accumulation on the children outcomes later in life.

⇒ 34% of the total effect.

Thanks for your attention.