Critical Periods in Child Development

Pia R. Pinger

University of Bonn, IZA

SSSI 2016

Bonn

Table of Contents

Skills and human capital matter

The returns to human capital are large and rising

- Large returns to cognitive and non-cognitive skills and health

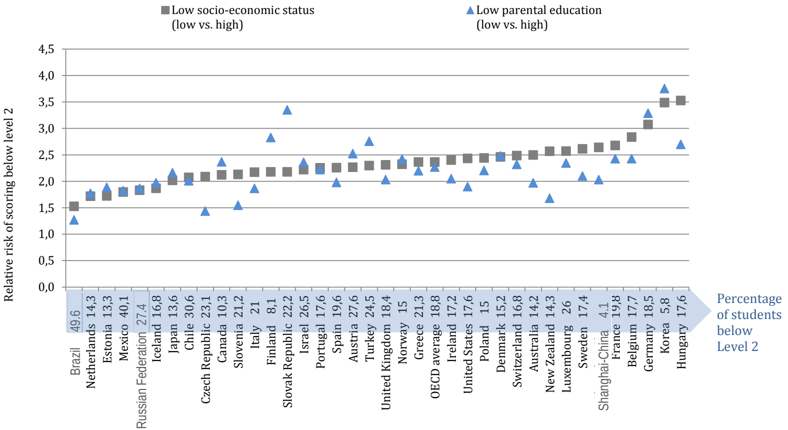
  returns to skills graph

- Large returns to education returns to education graph

Who are the “winners”? Typically children from high SES families

School performance by SES

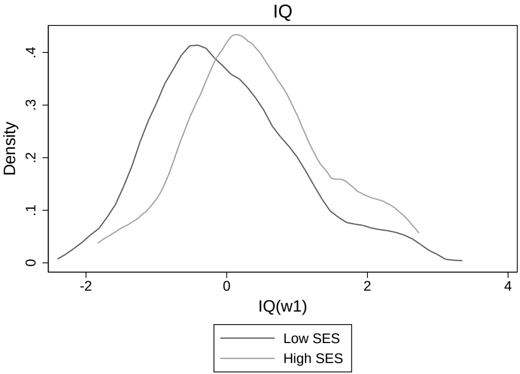
Child traits/outcomes **differ dramatically by SES**. Low SES children perform worse in PISA



Source: OECD (2010a), PISA 2009 Results: Volume II, Overcoming Social Background: Equity in Learning Opportunities and

IQ performance by SES

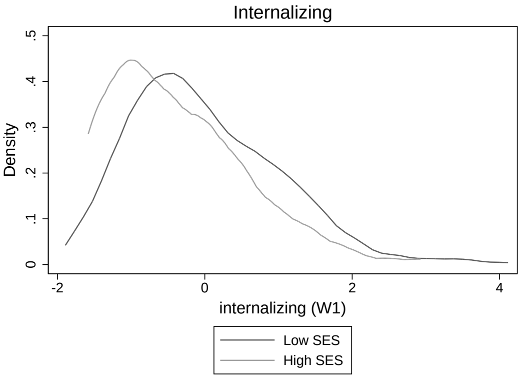
Equality of opportunity links this question to inequality. Child traits/outcomes **differ dramatically by SES**



Data Source: Bonn Intervention Panel

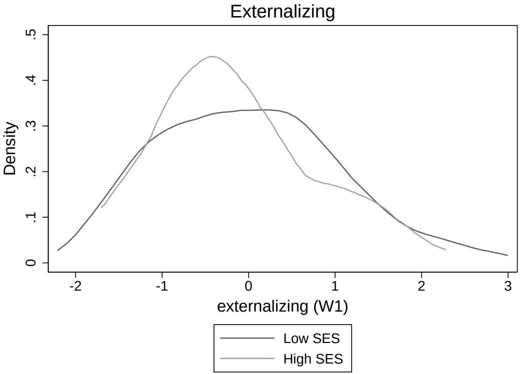
Internalizing performance by SES

Equality of opportunity links this question to inequality. Child traits/outcomes **differ dramatically by SES**



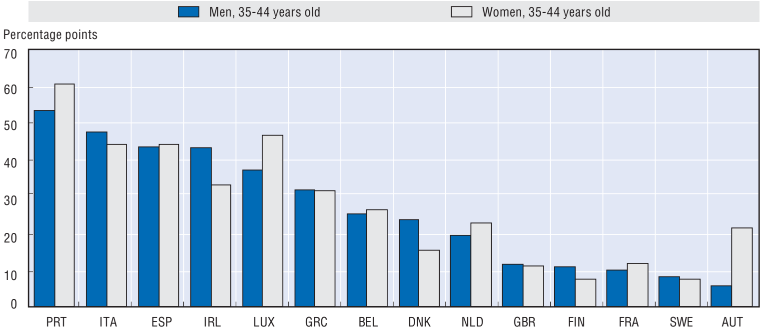
Data Source: Bonn Intervention Panel

Externalizing performance by SES



Data Source: Bonn Intervention Panel

High persistence



1. Persistence in below upper secondary education is measured as the distance between the estimated probability to achieve below upper secondary education of an individual whose father also had below upper secondary education and the probability to achieve below upper secondary education of an individual whose father had achieved tertiary education. A larger number implies a larger gap, thus stronger persistence in below upper secondary education or a lower degree of mobility across generations. For details see Causa *et al.* (2009).

Policy interventions to provide equality of opportunity

- As economists we care about the best **allocation** of scarce resources
- High **return** to a unit of investment due to optimal timing (ROI)

$\Rightarrow$ find out about important periods (**critical periods**) for human capital formation (**long-term outcomes**) and target policy accordingly

- **Efficiency**: Use the lowest amount of inputs to create the greatest amount of outputs
- **Equality of opportunity**: Remove the influence of factors over which individuals have no control (accident of birth)

Sensitive and Critical Periods: Definition

Critical period

- A critical period is a **maturational stage** at which investments are particularly fruitful and vital
- If an individual does not receive a stimulus during a critical period it may be **difficult**, ultimately **less successful**, or even **impossible**, to develop some functions later in life
- Remediation is prohibitively costly/impossible
- Sub-optimal conditions have adverse **long-run** implications (lower human capital)

Sensitive period

- A more extended period, after which learning/investment is still possible

Critical Periods and Sensitive Periods

- During critical and sensitive periods the return to investment is high
- Mostly higher than during any other period (efficiency)
- Give children the opportunity to live up to their potential (equality of opportunity)

- Sometimes researchers do not distinguish between critical and sensitive periods

Critical Periods

Examples

- Language acquisition: Native speaker up to age 5, proficiency up to puberty
- Cognitive skills: Birth up to age 8-10
- Locus of Control: Puberty
- Height/health: Up to age 4-6

Critical Periods

Critical periods in everything?

- Child traits (biological) ‣ Critical periods
  - Cognitive skills
  - Character traits/noncognitive skills/preferences
  - Health
- Decision-making (social) ‣ Critical decision periods
  - Education transitions (tracking, college)
  - Labor market entry
  - Path dependency

Long-term effects within and across generations

Critical Periods: The challenge

How can we find out about critical periods? We need

- **Data**
  - (Parental) investments
  - Variation in investments
  - Investments at different ages
  - Child outcomes (test scores, labor market outcomes)
- A **model**/hypothesis (technology of skill formation)
- An **identification strategy**
  - A (statistical) method that allows us to make causal statements about the (lack of) childhood investments at different ages on long-term outcomes.

Table of Contents

Table of Contents

Identification: The challenge

What is the causal effect of **underinvestment** during age period t?

- Assume underinvestment for individual i could be described by a binary random variable, the *treatment*.

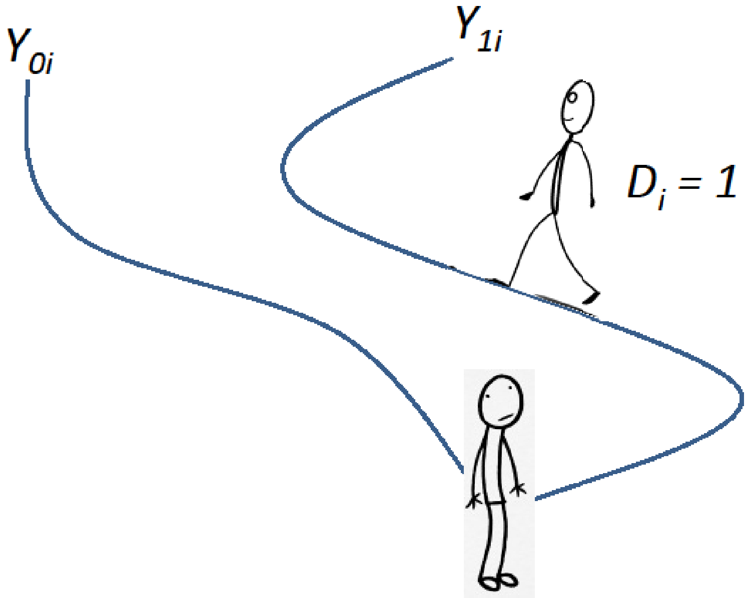
$$D_i = \{0, 1\} \tag{1}$$

- What would have happened to someone who received very little investments if he had received more?
- Hence, two

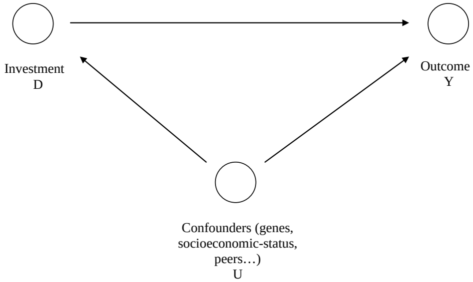
$$\text{potential outcomes} = \begin{cases} Y_{i1} \text{ if } D_i = 1 \\ Y_{i0} \text{ if } D_i = 0 \end{cases} \tag{2}$$

$$\text{observed outcome} = Y_{0i} + (Y_{1i} - Y_{0i})D_i \tag{3}$$

Identification: The challenge



Identification: The challenge



$$Y_i = \alpha + \rho D_i + U_i$$

Many other factors (confounding variables) exit, which influence both parental investments and the outcome ($Cov(D, U) \neq 0$)

Identification: The challenge

IDENTIFICATION PROBLEM: investments and later life outcomes jointly driven by unobserved factors

- Endogeneity is everywhere ($Cov(D, U) \neq 0$)
- Ceteris paribus changes rarely exist
- Experiments are sometimes prohibitively costly, unpractical, or unethical
- It takes a long time to observe long-term outcomes
- Individuals do not remember early investments (recall bias)

Identification: The challenge

Strategies

1. **Contextual variation and IVs**: find an instrument!
2. **Matching on observables/unobservables**: model the unobserved heterogeneity (U)
3. **Experiments**: randomize s.t. unobservables are the same (RCT/ECI)!
4. **Structural models**: find a theoretical model and fit it to data

Table of Contents

Contextual variation and IVs

Economists are quite creative in finding exogenous **contextual changes at the macro-level** (instruments) that are **exogenous at the individual level**

1. Historical incidences
   - Landmine explosions (Camacho, 2008, AER), Influenza (Almond, 2006, JPE), Rainfall (Maccini and Yang, 2009, AER), Malaria (Barreca, 2010, JHR)
2. Variation in macroeconomic conditions
   - Recessions (van den Berg et al., 2006, AER), Famines (Lumey, Susser & Stein, 2011; Neelsen and Stratmann, JHE, 2011; Lindeboom, Portrait & van den Berg, JHE, 2010)
3. Policy changes
   - The introduction of food stamps (Hoynes et al., 2016, AER), Change in the EITC (Dahn and Lochner, 2012, AER); Alcohol

Contextual variation and IVs

Some **advantages** of using contextual variation are

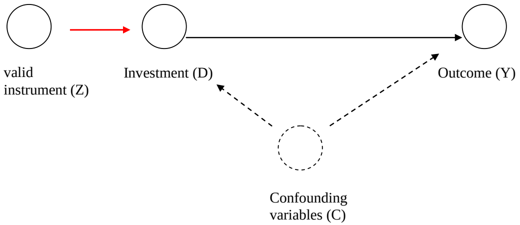
- Provide quasi-experimental variation in investments
- Large variation in inputs
- Investments/schocks can be linked to long-term outcomes
- Allows us to study the impact of investments within and across generations

Some **disadvantages** of using contextual variation are

- Have to take “whatever is there” as an instrument (compliers)
- Oftentimes measures of **compliance** are absent
- Information is often retrospective and may not be remembered correctly
- Selective fertility/selective mortality

Contextual variation and IVs

- An instrument (Z) exogenously changes the probability that we observe a certain D (strong shifter)
- It should influence Y only through its effect on D (exclusion restriction)



Instrument/Exogenous Variation

Contextual variation and IVs

- Z thus needs to satisfy two conditions:

1. (**Exogeneity**) Z is uncorrelated with U in the outcome equation (as good as randomly assigned):

$$Cov(Z, U) = 0 \tag{4}$$

2. (**Relevance**) In a linear projection of D on all the exogenous variables the coefficient on Z is nonzero:

$$\delta_1 \neq 0 \tag{5}$$

i.e. Z is partially correlated with D (once the other exogenous variables have been netted out).

Contextual variation and IVs

- Z thus needs to satisfy two conditions:

1. (**Exogeneity**) A historical incidence/policy change needs to (randomly) affect some individuals but not others. Variation may occur
   - Across time (e.g. one year) $\Rightarrow$ compare cohorts
   - Across space (e.g. one region) $\Rightarrow$ compare individuals in different locations
   - Across cutoffs (e.g. school grades) $\Rightarrow$ compare individuals at the cutoff
   - By luck (e.g. lottery) $\Rightarrow$ quite ideal

2. (**Relevance**) A historical incidence/policy change needs to have a strong effect on investments (the treatment).

Contextual variation and IVs

Then there are three causal relationships we may care about.

- The **structural equation** (outcome equation)

$$Y_i = \alpha + \rho D_i + U_i$$

where $U_i = c_i + \epsilon_i$ is a structural error term, not a regression residual.

- **First stage**: The regression of the treatment on the instrument (causal effect 1)

$$D_i = \alpha_1 + \delta_1 Z_i + v_{i1}$$

- **Reduced form**: The regression of earnings on the instrument is called the reduced form (causal effect 2)

$$Y_i = \alpha_2 + \delta_2 Z_i + v_{i2}$$

Contextual variation and IVs

There is a relationship between the three equations that you should be aware of.

$$
\begin{aligned}
Y_i &= \alpha + \rho D_i + U_i \\
&= \alpha + \rho[\alpha_1 + \delta_1 Z_i + v_{i1}] + U_i \\
&= (\alpha + \rho\alpha_1) + \rho\delta_1 Z_i + (\rho v_{i1} + U_i) \\
&= \alpha_2 + \delta_2 Z_i + v_{i2}
\end{aligned}
$$

Hence, the reduced form coefficients are:

$$
\begin{aligned}
\alpha_2 &= \alpha + \rho\alpha_1 \\
\delta_2 &= \rho\delta_1
\end{aligned}
$$

Contextual variation and IVs

The IV estimate is then equal to the ratio of the reduced form coefficient on the instrument to the first stage coefficient.

$$\rho = \frac{\delta_2}{\delta_1}$$

With a binary treatment and a binary instrument we can compute the causal effect using the Wald estimator:

$$\begin{aligned}\rho_{IV} &= \frac{cov(Y_i, Z_i)}{cov(D_i, Z_i)} \\ &= \frac{\mathrm{E}[Y|Z=1] - \mathrm{E}[Y|Z=0]}{\mathrm{E}[D|Z=1] - \mathrm{E}[D|Z=0]}.\end{aligned}$$

Contextual variation and IVs

- Sometimes, the reduced form effect is interesting in its own right.
- When contextual variation is used, a measure of compliance $\mathrm{E}[D|Z=1] - \mathrm{E}[D|Z=0]$ is often not available. Limitation:
  - Overcome endogeneity problem but cannot measure causal effect of undernutrition (imperfect compliance)
  - $\Rightarrow$ **Reduced form**/aggregate effect; but how are *individuals* affected?

Example: **"What is the average *causal* effect of a nutritional shortage in infancy and childhood on health (adult height) later in life?"** (with G. van den Berg and J. Schoch, EJ, 2016)

Contextual variation and IVs

1. Relate aggregate data (exogenous variation) to information on individual suffering and measure causal effect

$\rightarrow$ Use IV strategy

   - Instrument: famine (trade blockades in 3 countries)
   - Treatment: nutritional shortage
   - Outcome: adult height

2. Obtain estimate of "compliance" to the instrument
3. Deal with the problem of imperfect recall

Contextual variation and IVs

SHARE: Survey of Health, Aging, and Retirement in Europe (50+)

- 1st and 2nd wave of data: survey data
- 3rd wave: retrospective interviews about life-course
- Information on rural/urban at birth
- Parental occupation, SES, father absent (at age 10)
- Birth cohorts 1920–1955 from Germany, Netherlands, Greece
- 2511 males, 2859 females

Contextual variation and IVs

For different age windows:

- **Treatment** Report period of severe hunger (undernutrition): *Looking back at your life, was there a distinct period during which you suffered from hunger?*
- **Instrument** uses information on famines in three European countries (from birth year, birth region and location history):
  1. West of the Netherlands: November 1944– April 1945
  2. Greece: May 1941– June 1942
  3. Germany: June 1945–June 1948

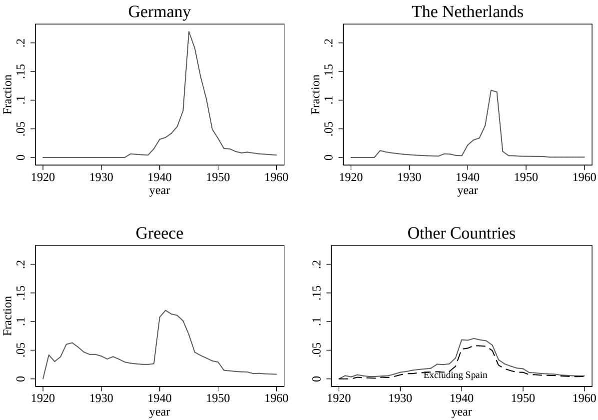
Contextual variation and IVs

For different age windows:

- **Treatment** Report period of severe hunger (undernutrition): *Looking back at your life, was there a distinct period during which you suffered from hunger?*
- **Instrument** uses information on famines in three European countries (from birth year, birth region and location history):
    1. West of the Netherlands: November 1944– April 1945 (500kcal)
    2. Greece: May 1941– June 1942 (300-600kcal)
    3. Germany: June 1945–June 1948 (1330, 1083, 1050, 900kcal)

Contextual variation and IVs

Figure: Probability for Episode of Hunger by Calendar Year



Contextual variation and IVs

First Approach – Impacts at 6 to 16

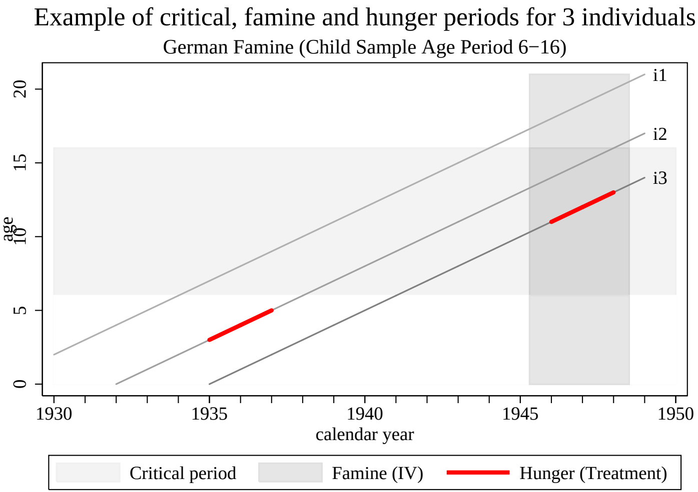
- Construct binary treatment variable:

$$D_i = \begin{cases} 1 & \text{if } i \text{ reports hunger aged } [6, 16) \\ 0 & \text{otherwise} \end{cases}$$

- Construct binary famine IV:

$$Z_i = \begin{cases} 1 & \text{if } i \text{ experienced famine aged } [6, 16) \\ 0 & \text{otherwise} \end{cases}$$

Figure: Examplary Treatment Definition



Contextual variation and IVs

Compute local average treatment effect (LATE):

$$LATE = \mathrm{E}[Y_{D=1} - Y_{D=0} | D_{Z=1} > D_{Z=0}]$$

1. Nonparametric Wald estimator

$$LATE = \frac{\int \mathrm{E}[Y|X=x, Z=1] - \mathrm{E}[Y|X=x, Z=0]\, f(x)dx}{\int \mathrm{E}[D|X=x, Z=1] - \mathrm{E}[D|X=x, Z=0] f(x)dx}.$$

2. 2SLS

Contextual variation and IVs

Figure: First stage: Coefficients for probability of reporting hunger (age 6-16)

	Males		Females	
Experienced famine being 6-16 (1 = yes)	0.207***	0.179***	0.135***	0.097***
	(0.020)	(0.024)	(0.018)	(0.025)
German Sample		2.367		4.707*
		(3.964)		(2.665)
Dutch Sample		2.923		2.539
		(2.877)		(2.660)
Lived in rural area at age 6		-0.025**		-0.026***
		(0.010)		(0.009)
Year of birth		-0.002*		-0.002**
		(0.001)		(0.001)
Year of birth × Dutch		-0.002		-0.001
		(0.001)		(0.001)
Year of birth × German		-0.001		-0.002*
		(0.002)		(0.001)
Constant	0.030***	4.213*	0.034***	4.786**
	(0.005)	(2.130)	(0.006)	(1.910)
R^2	0.106	0.116	0.051	0.070
F-Stat.	102.432	27.079	55.315	21.187
N	2511	2511	2859	2859

Contextual variation and IVs

For age 6–16, we find

- No robust significant effects (reduced form + IV)

Males

	Reduced Form	Instrumental Variables Models		
	Famine at age 6 – 16	2SLS	cond. Wald	cond. Wald – Trend corrected
Effect	0.269	1.502	-4.647	0.803
(S.E.)	(0.325)	(1.808)	(1.522)	(1.273)
t-stat.	0.829	0.831	-3.054	0.631

Females

	Reduced Form	Instrumental Variables Models		
	Famine at age 6 – 16	2SLS	cond. Wald	cond. Wald – Trend corrected
Effect	0.229	2.373	-6.696	1.291
(S.E.)	(0.258)	(2.773)	(2.050)	(1.117)
t-stat.	0.887	0.856	-3.267	1.156

Contextual variation and IVs

Impacts at 0 to 4 or in utero: recall bias for early childhood

- Let Y and Z be an outcome and a valid IV, respectively
- Let D^* be the true treatment of hunger in early life

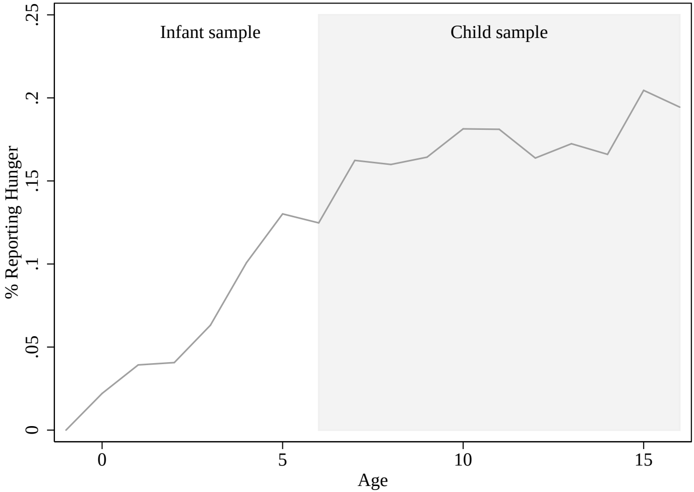
Want to estimate using IV:

$$Y = \theta D^* + \epsilon$$

PROBLEM: D^* is not recalled correctly

Contextual variation and IVs

Figure: Probability to report hunger by age



Contextual variation and IVs

IDEA: Two stage-procedure We estimate:

1. Nonparametric Wald estimator

$$LATE = \frac{\int \mathrm{E}[Y|X=x, Z=1] - \mathrm{E}[Y|X=x, Z=0]\, f(x)dx}{\int \mathrm{E}[D|X=x, Z=1] - \mathrm{E}[D|X=x, Z=0] f(x)dx}.$$

2. 2-sample-2SLS (Arellano & Meghir, 1992; Inoue & Solon, 2010)

Sample 1: Childhood, sample 2: Infancy

Contextual variation and IVs

Males

	Famine at age 0 – 4	Cond. Wald	cond. Wald – Trend corrected
Effect	-0.683	-2.174	-2.885
(S.E.)	(0.268)	(1.685)	(1.249)
t-stat.	-2.550	-1.290	-2.309

Females

	Famine at age 0 – 4	Cond. Wald	cond. Wald – Trend corrected
Effect	0.259	3.151	1.318
(S.E.)	(0.288)	(2.303)	(1.674)
t-stat.	0.899	1.368	0.787

Contextual variation and IVs

Height effects for in utero and ages 0-4 (men)

- **Reduced form**: - 0.68 cm
- **LATE**: - 2.5 cm

Consistent results of nonparametric Wald & 2-sample-2SLS estimators

$\rightarrow$ Treatment effects are 3–8 times larger than reduced form estimates

Critical periods across generations

- Shocks during critical periods affect outcomes within one generation
- How about the next generation?
- Inequalities in
  - **social status** tend to persist across several generations (Clark, 2014; Lindahl et al., 2015)
  - **health** tend to persist to some extent (Clark, 2014; Lindahl et al., 2015)
    - SES/intrauterine conditions/Barker (Case et al., 2005)
    - Birth weight (Royer, 2009)
    - Longevity/self-assessed health/BMI (Trannoy et al., 2010; Brown and Roberts, 2013)

Is any of this causal? ($\Rightarrow$ long-term effects of policies)

Inter- and Transgenerational Effects

Questions related to the transmission of traits across generations:

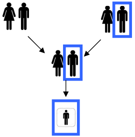
1. Can we separate biological/genetic/social mechanisms?
2. Can shocks change an individual's biology?
3. Emerging but difficult field of research

Series of papers in epidemiology

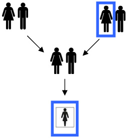
‘Överkalix study’ (Bygren et al., 2001, Kaati et al., 2002, Kaati et al., 2007, Pembrey et al., 2006, Kaati, 2010, Pemrey, 2010):

- Sole transgenerational evidence on humans
- Low food supply in Slow Growth Period (SGP) of paternal grandfather [age 9-12] (paternal grandmother [age 8-10]):
  - $\Rightarrow$ Low mortality of grandsons (granddaughters)
  - $\Rightarrow$ Low CVD mortality of grandsons
  - $\Rightarrow$ High diabetes mortality of grandsons with surfeit of food

Males



Females



Single line of research papers on humans

What do these papers imply?

- Slow growth period as a sensitive period for methylation of the male sperm?
- Potential transgenerational response to developmental conditions
- **Adverse** conditions ⇒ **improve** offspring survival capabilities
- **Favorable** conditions ⇒ **worsen** offspring survival capabilities

Huge interest in this issue in biological literature (Zeisel, 2007; Gräff and Mansuy, 2008; Masterpasqua, 2009; Francis, 2011; Grossniklaus et al., 2013)

⇒ Need for reproduction/validation

Mechanism for such transgenerational effects?

- $\Rightarrow$ Take a look at **biological literature**
  - Experiments using mice models (nutrition, methylation states last up to 4 generations, transmission through paternal/maternal line)
  - New evidence on epigenetic changes in humans (Yehuda et al., 2015, Tobi et al., 2009)
  - Not only DNA matters, but gene expression
  - **Epigenetics:** Study of heritable changes in gene expression
- The authors argue that the **slow growth period** is a sensitive period for sperm development
- Transmission to the next generation via **epigenetic imprinting**
- Behavior/experience changes methylation $\Rightarrow$ methylation heritable $\Rightarrow$ Species change quickly

Recent advances in the biological literature...

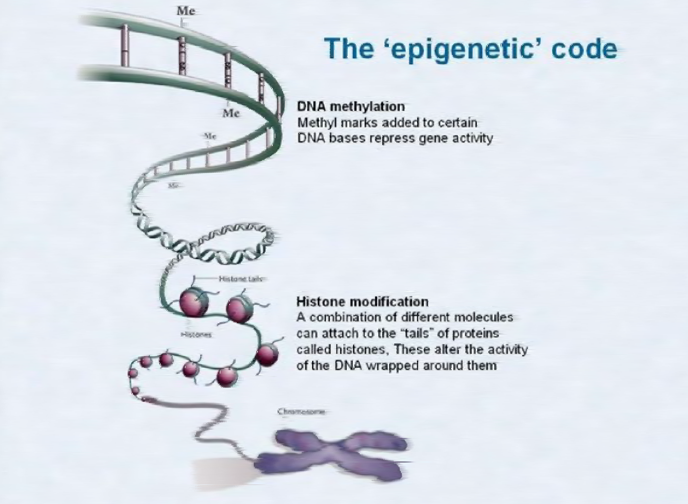
Epigenetics

- study of *heritable* changes in gene expression
- No change in the underlying DNA sequence!
- DNA methylation or histone acetylation
- Methylation determines cell fate
- Epigenetic imprinting: methyl tags from parents remain after conception ($\sim$ 1%)

$\Rightarrow$ Non-genetic inheritance (often male/female germline)

Behavior/experience changes methylation $\Rightarrow$ methylation heritable $\Rightarrow$ Species change quickly

DNA Methylation and the Epigenetic Code



Do transgenerational effects exist?

Several papers on this issue:

1. Using German data we found adaptive responses regarding mental health (with G. van den Berg)
   - *Paternal grandfather* famine during SGP $\Rightarrow$ better grandson mental health
   - *Maternal grandmother* famine during SGP $\Rightarrow$ better granddaughter mental health
   - No fading out
   - Behavioral mechanisms implausible
   - Stronger effects on male than female line
2. No validation of smoking relationship in Norwegian data (with D. Carslake, G. Davey-Smith, Pål Romundstad)
3. No strong longevity effects in a Swedish study over four generations (with G. van den Berg, Bitte Modin, Denny Vågerö)

Table of Contents

Matching

We can use matching strategies to solve the endogeneity problem

$$Y_i = \alpha + \beta X_i + \rho D_i + U_i$$

- (Condition on observable variables (Xs)$\Rightarrow$ find the statistical twin)
- Model the unobserved heterogeneity (factor models/random effects)
- Get rid of the unobserve heterogeneity (fixed effects)

Model the unobserved heterogeneity

- Wage equation, for $s = 0, 1$

$$Y_s = X_Y \beta_Y^s + \underbrace{\theta \alpha_Y^s + \varepsilon_Y^s}_{U}, \tag{6}$$

- Measurement system (imperfect measures, dimension reduction, endogeneity)

$$\begin{aligned} M_k &= \sum_{c=1}^{C_k} c\, \mathbf{1}\left(\gamma_{k,c-1} \leq M_k^\star < \gamma_{k,c}\right), \\ M_k^\star &= X_M \beta_{M_k} + \theta \alpha_{M_k} + \varepsilon_{M_k}, \end{aligned} \tag{7}$$

- Investment decision ($S = 1$)

$$\begin{aligned} S &= \mathbf{1}(S^\star > 0), \\ S^\star &= X_S \beta_S + \theta \alpha_S + \varepsilon_S, \end{aligned} \tag{8}$$

Model the unobserved heterogeneity

Latent factor:

$$\theta \sim \mathcal{N}\left(0;\, \sigma_\theta^2\right) \qquad\qquad \theta \perp\!\!\!\perp X \perp\!\!\!\perp \varepsilon$$

Latent factors: normally distributed or mixture of normals

$$\varepsilon \perp\!\!\!\perp X, \qquad \varepsilon \perp\!\!\!\perp \theta, \qquad \varepsilon_j \perp\!\!\!\perp \varepsilon_{j'}, \tag{9}$$

Error terms: standard normal (probits and ordered probits) or mixture of normals

Model the unobserved heterogeneity

- With $\mathrm{E}(\theta) = 0, \quad \mathrm{V}(\theta) = \sigma_\theta^2 \ll \infty, \quad \theta \perp\!\!\!\perp X$,
- The model is identified from the covariance matrix:

$$\Omega^\star \equiv \mathrm{V}(S^\star, Y, M^\star \mid X), \tag{10}$$

Factor loadings can be identified from the ratios of observed covariances

$$\frac{\mathrm{Cov}\left(M_1^\star; M_2^\star \mid X\right)}{\mathrm{Cov}\left(M_2^\star; M_3^\star \mid X\right)} = \frac{\alpha_{M_1}\alpha_{M_2}\sigma_\theta^2}{\alpha_{M_2}\alpha_{M_3}\sigma_\theta^2} = \frac{\alpha_{M_1}}{\alpha_{M_3}}. \tag{11}$$

Either set $\alpha_{M_1} = 1$ or $\sigma_\theta^2 = 1$ to set the scale

Model the unobserved heterogeneity

- With discrete items the covariance matrix is not directly observed
- Use polychoric correlations (normality of error term)
- Distributional assumptions

$$\theta \sim \mathcal{N}\left(0;\, \sigma_\theta^2\right), \qquad (\varepsilon_S, \varepsilon_D^0, \varepsilon_D^1, \varepsilon_{M_1}, \ldots, \varepsilon_{M_K})' \sim \mathcal{N}(0;\, \Sigma), \tag{12}$$

Mixture for error term of the wage equation:

$$\varepsilon_Y^s \sim \sum_{h=1}^{H_s} \pi_h^s \mathcal{N}\left(\mu_h^s;\, (\omega_h^s)^2\right), \qquad \mathrm{E}\left(\varepsilon_Y^s\right) = \sum_{h=1}^{H_s} \pi_h^s \mu_h^s = 0, \tag{13}$$

Model the unobserved heterogeneity

Derive the likelihood

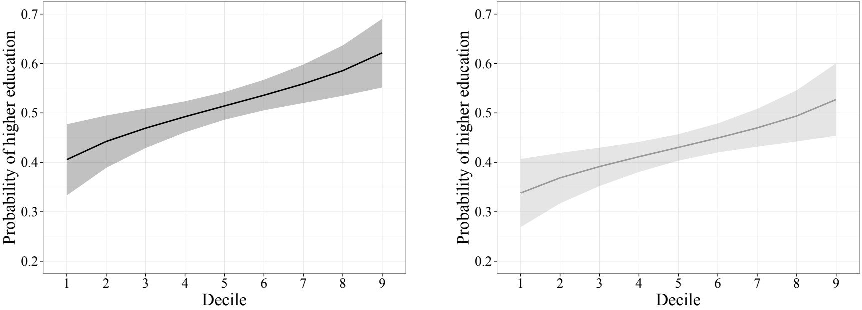
$$\mathcal{L}(\psi \mid S, Y, M, X) = \int_{\Theta} \prod_{s=0}^{1} \Pr(S = s | X, \theta, \psi)^{\mathbf{1}(S=s)} \times \prod_{s=0}^{1} f(Y_s | X, \theta, \psi)^{\mathbf{1}(S=s)} \times \prod_{k=1}^{K} f(M_k | X, \theta, \psi) \,\mathrm{d}F(\theta),$$

Estimate it with frequentist or Bayesian methods

Two-step approach possible (adjust standard errors + bias)

Model the unobserved heterogeneity

Simulation study (effect on education)



Model the unobserved heterogeneity

Simulation study (effect on wages)

Figure: Higher education

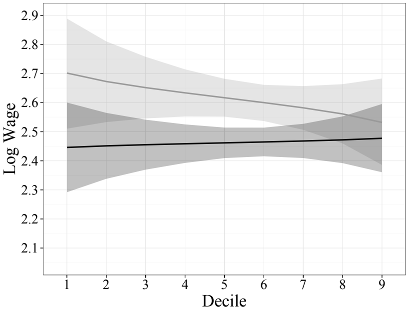
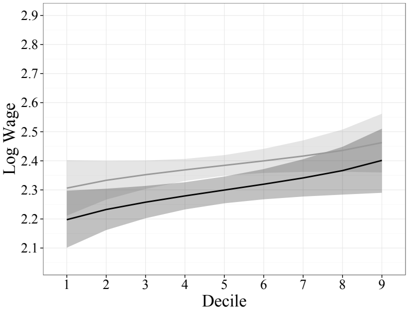


Figure: No higher education



Get rid of the unobserved heterogeneity

Get rid of the unobserve heterogeneity (fixed effects)

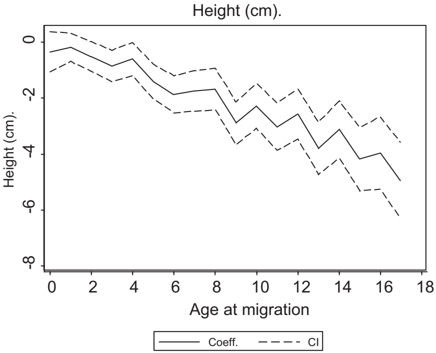
$$Y_i = \alpha + \beta X_i + \rho D_i + \underbrace{c_i + \varepsilon_Y^s}_{U}$$

E.g. van den Berg et al. (2014) "Critical Periods during Childhood and Adolescence"

- Compare siblings who enter Sweden from poor countries
- Siblings enter Sweden at different ages (family FE)
- The family fixed effect wipes out all family-specific unobserved determinants
- The family fixed effect wipes out shocks that are common to both siblings

Get rid of the unobserved heterogeneity

Results



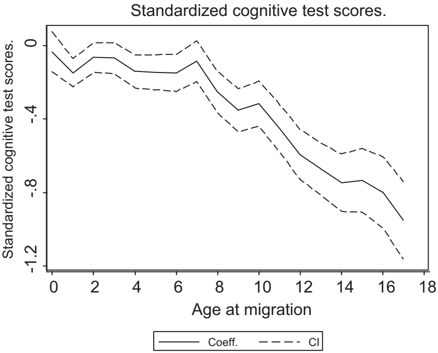


Table of Contents

Take me on Your Shoulders!

The Effect of Child Mentoring on Education Outcomes

Armin Falk   Fabian Kosse   Pia Pinger   Hannah Schildberg-Hörisch

Evidence from an experiment

Inequality of opportunity is a major concern in the presence of

1. High and increasing returns to education
2. Educational achievement being determined by SES

- Risk factors are
  - **Low parental education** (Heineck and Riphahn, 2007; Lundborg et al., 2014)
  - **Low income** (Duncan et al., 1998; Dahl and Lochner, 2012)
  - **Single parenthood** (Krein and Beller, 1988; Ermisch and Francesconi, 2001)

Evidence from an experiment

Much recent work focuses on how opportunities of children from low SES backgrounds can be improved

- **Preschool education/Intervention programs** (Deming, 2009; Heckman et al., 2010; Campbell et al., 2014; Gertler et al., 2014; Attanasio et al., 2015)
- **Mentoring**: advising, helping parents, personal assistance (Lavecchia et al., 2014; Oreopoulos, 2014; Fryer, 2016)

Improved equity and potential for large societal returns

Evidence from an experiment

- Preschool education/ECIPs affect the formation of human capital
- **Mentoring** provides
  - Information/advice
  - Role models
  - Character traits prosociality/patience/grit
  - Help for overcoming self-control problems
  - **Substitutes for parental time and encouragement**

$\Rightarrow$ Improve outcomes during critical decision periods (childhood & adolescence)

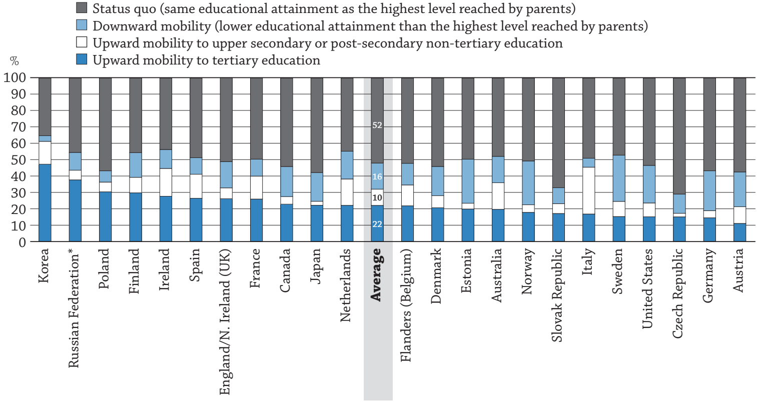
This project: evidence from a mentoring RCT

Does a low-cost weekly mentoring program during elementary school affect secondary **school track choice** in Germany?

1. What is the **overall effect** of mentoring shortly before a critical education decision (tracking)?
2. Which **groups** benefit most?
   - Household risk factors (poverty, low education, single parenthood)
   - Child characteristics (age, sex, ability)
3. Why?

The German setting

In Germany educational mobility is low despite 100% free education



OECD (2015), Education at a Glance 2015: OECD Indicators, OECD Publishing, Paris.

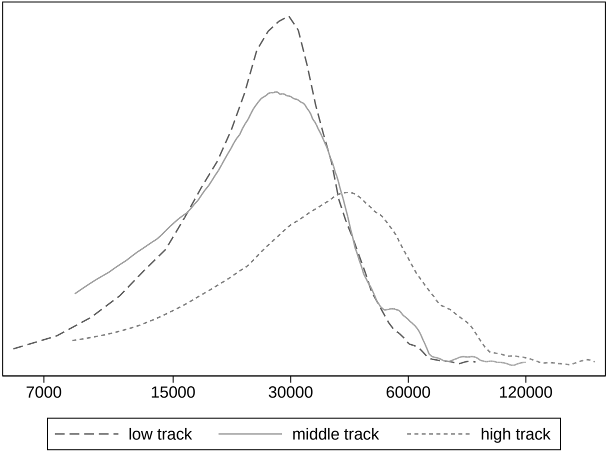
The German setting

Early tracking as one reason for low mobility (Bauer and Riphahn, 2005; Pekkarinen et al., 2009)

- After 4th grade:
  - high track: upper secondary school degree (Gymnasium, 42%)
  - middle track: secondary school degree (Realschule, 21%)
  - low track: lower secondary school degree (Hauptschule, 4.3%)
- High track allows for university studies (upper secondary school certificate)
- Teacher recommendation after first half of 4th grade (mandatory or non-mandatory)

The German setting

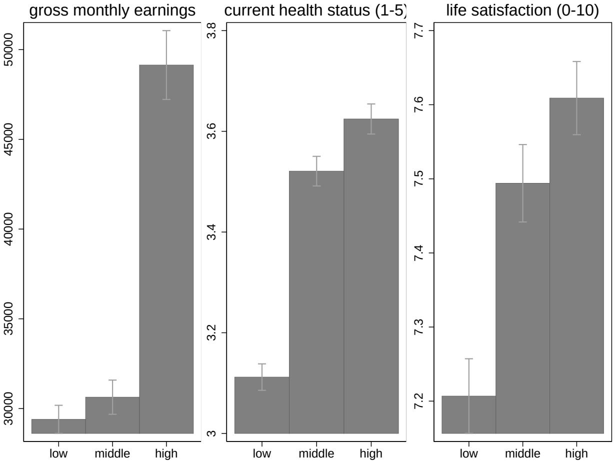
Graduation from a certain track is predictive of wages



SOEP, 2015, kernel density plot of gross annual wages (ft employed), logarithmic scale, own calculations.

The German setting

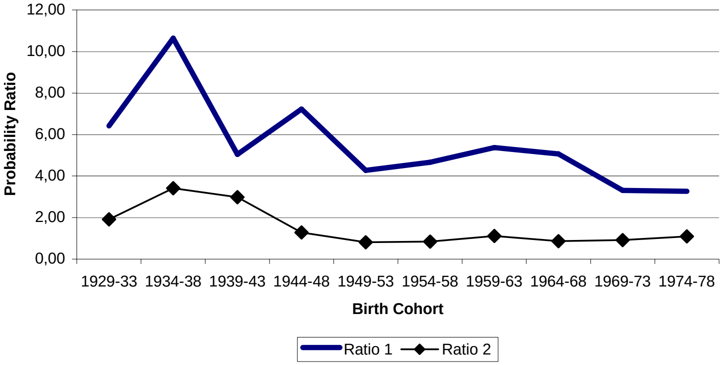
Graduation from a certain track is predictive of other life outcomes



SOEP, 2015, own calculations.

The German setting

Graduation from a certain track is predictive of child track choice



Ratio 1: P (child high — parent high) / P (child high — parent low)

Ratio 2: P (child middle — parent middle) / P (child middle — parent low)

Heineck, G. and Riphahn, R.T. (2009), Intergenerational transmission of educational attainment in Germany - The last five decades. *Jahrbücher für Nationalökonomie und Statistik*, pp.36-60 (graph for males only).

The German setting

Parental background matters even after conditioning on IQ/GPA

high track	1	2	3	4	5
parental background					
poor HH	-0.198***	-0.198***	-0.197***	-0.132**	-0.132**
low educated HH	-0.295***	-0.295***	-0.293***	-0.236***	-0.215***
single parent HH	-0.103*	-0.103*	-0.102*	-0.123**	-0.071
gender and age					
sex (male=1)		-0.012	-0.014	-0.024	-0.041
grade			-0.007	-0.078	0.041
ability					
IQ				0.183***	0.100***
GPA					-0.220***
Observations	342	342	341	341	341
pseudo-R2	0.10	0.10	0.10	0.14	0.22

This table reports average marginal effects from a logit model. “Poor” indicates that a respective household earns less than the 30th quantile of the German income distribution. “1 parent” (“2 parents”) indicates that a child grows up in a single parent (two parent) household. “Low edu” indicates that a child grows up in a household where neither parent has obtained an upper secondary school certificate (highest track credential). Robust standard errors in parentheses. Treated individuals were excluded from the sample. All models contain a constant (intercept). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The mentoring RCT

- Intervention is a mentoring program (Baloo and you)
- Mentors
  - Volunteers, mainly university students
  - Meet children once per week, overall duration: one year
- Concept of the mentoring program:
  - One-to-one mentoring, Informal learning, no focus on achievement
  - Widening a child's horizon through engaging in joint activities with a new contact/attachment person, role model
- Children were in 2nd (80%) or 3rd grade (20%)
- Professional structure: online diaries, paid coordinators, bi-weekly monitoring meetings
- Moderate monetary costs: 1000EUR per child and year

The mentoring RCT

Data collection

- Family addresses from registry data
- Offers to families with
  - Children born between 09/2002 and 08/2004
  - Low income families ($<$30th percentile)
  - Low education families (neither mother nor father with upper secondary school degree)
  - Single parent families
- Stratified random treatment assignment: 14 subgroups by city and SES criteria

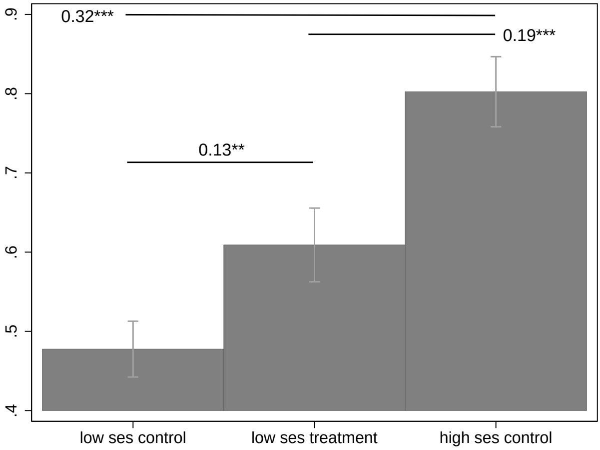
The mentoring RCT

Data come from 3 waves of interviews

- Wave 1: Central location labs (**pre-treatment**)
- Wave 2: Central location labs (**post-treatment/pre-transition**)
- Wave 3: Home interviews (**post-transition**)
  - Mothers: answered a SOEP-like questionnaire
  - Children: one-to-one questionnaires with trained interviewers
- Vast battery of questions on
  - Child characteristics
  - Parental background
  - School outcomes (track, grades, IQ)

Results

Treatment effect on attending upper secondary school in grade 5



Results

Treatment effect on attending upper secondary school in grade 5

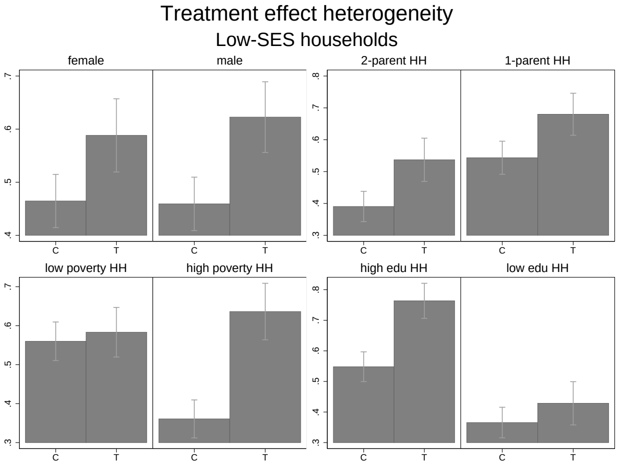


Table of Contents

Table of Contents

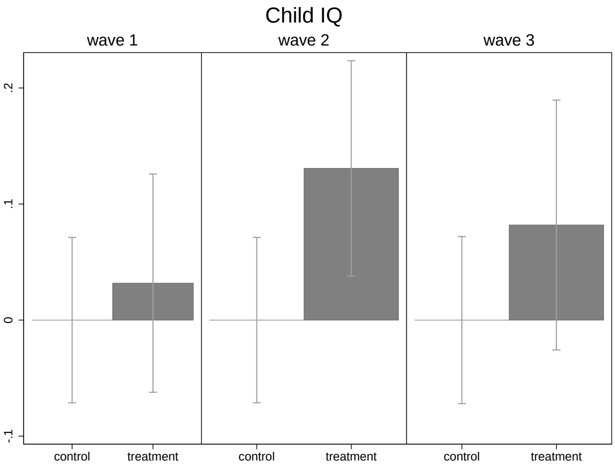
Mechanisms

Are we done?

- Overall causal effect is important
- If we only knew why...
  - we could speculate about **external validity**
  - we could design **appropriate policies**
  - we would **understand the world better**
  - we could **write a model**
- Separate the overall treatment effect into
  - Observed mechanisms (**indirect** output effect)
  - Unobserved residual effects (**direct** output effect)

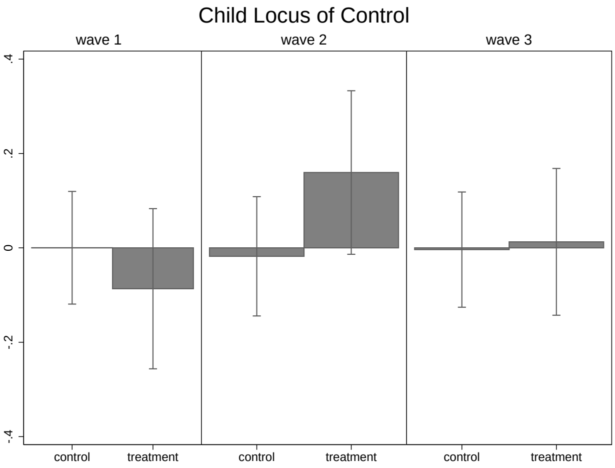
Mechanisms

Use an exogenous treatment to study effects on intermediate outcomes



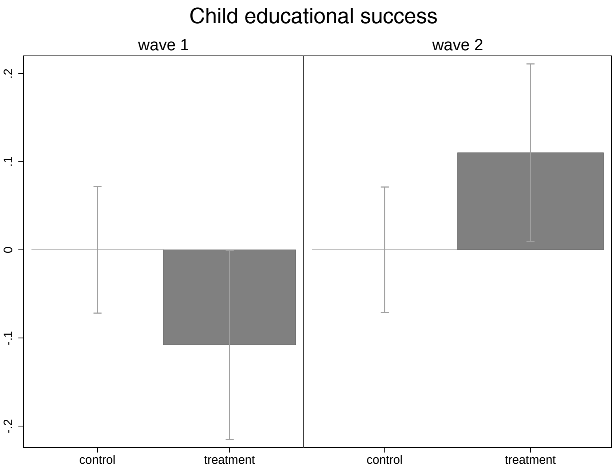
Mechanisms

Use an exogenous treatment to study effects on intermediate outcomes



Mechanisms

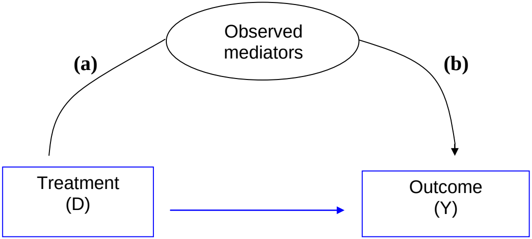
Use an exogenous treatment to study effects on intermediate outcomes



Mechanisms

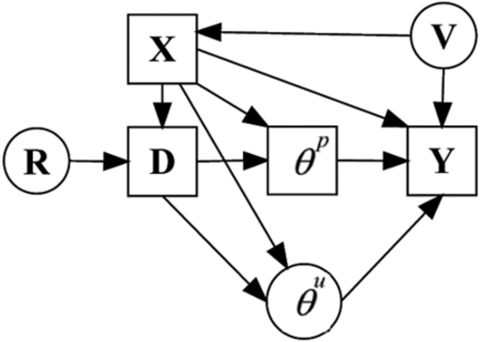
Ideally we want to

- Separate the overall treatment effect into
  - Observed mechanisms (**indirect** output effect)
  - Unobserved residual effects (**direct** output effect)
- This can be hard in the presence of unobserved mediators (Heckman and Pinto, 2015; Frölich and Huber, 2014) (b)



Mechanisms

1st case: change in unobserved mediators are **independent** of observed mediators



Mechanisms

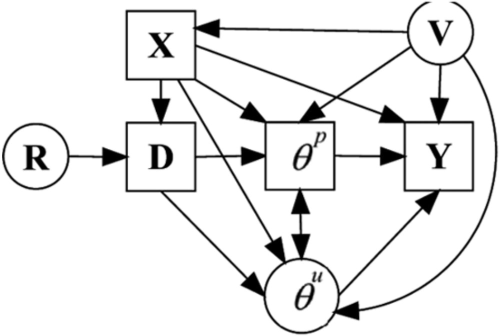
1st case: change in unobserved mediators are **independent** of observed mediators or matching on unobservables

$\Rightarrow$ Do some kind of Oaxaca-Blinder-type decomposition to decompose the ATE

$$E(Y_1 - Y_0) = \overbrace{\sum_{j=1}^{J} \theta^j E\left(M_1^j - M_0^j\right)}^{\text{observed}} + \underbrace{\tau_1 - \tau_0}_{\text{residual}}$$

Mechanisms

2nd case: change in unobserved mediators are **not independent** observed mediators



Mechanisms

2nd case: change in unobserved mediators are **not independent**of observed mediators

$\Rightarrow$ Need some kind of exogenous variation (instrument) for mediators

One possibility: estimate the following structural simultaneous equation system by two-stage least squares (**observed**=$\tau_M * \theta$ and **residual**=τ_r):

$$
\begin{aligned}
Y &= \alpha + \tau_r D + \theta M + \beta X + \epsilon_y \\
\mathbf{M} &= \alpha_M + \tau_M \mathbf{D} + \gamma_m \mathbf{Z_m} + \beta_M \mathbf{X} + \epsilon_{\mathbf{M}} \\
D &= \alpha_p + \gamma_p Z_p + \beta_p X + \epsilon_p,
\end{aligned}
$$

Conclusion

Lots of scope for **future research**

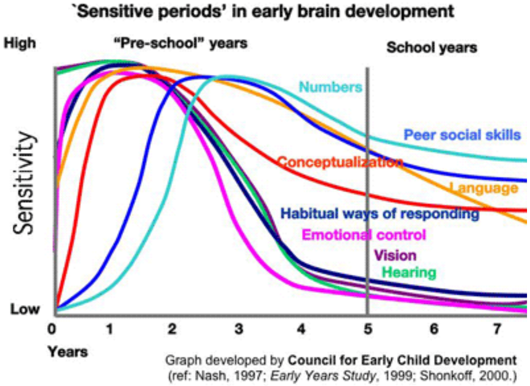
- Critical periods in everything (often country/institution specific)
- Use insights from behavioral and labor/education economics
- No unified theoretical framework for critical decision periods
- Long-term (intergenerational/transgenerational) effects
- Mechanisms

THANK YOU

pia.pinger@gmail.com

Critical periods

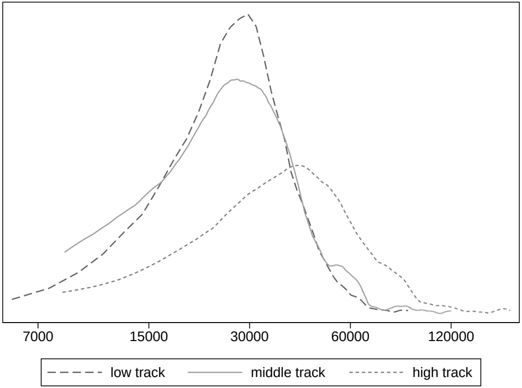
Critical periods in skill development



Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the" other 99 percent". *Science*, 344(6186), 843-851.

Critical periods

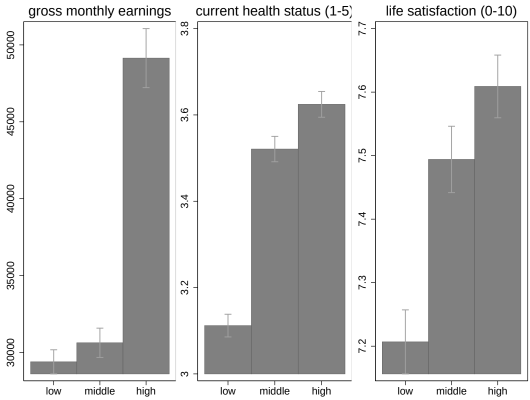
Critical decisions and wages



SOEP (own calculations)

Critical periods

Critical decisions and other outcomes



SOEP (own calculations)

Critical periods

Reasons for stopping school

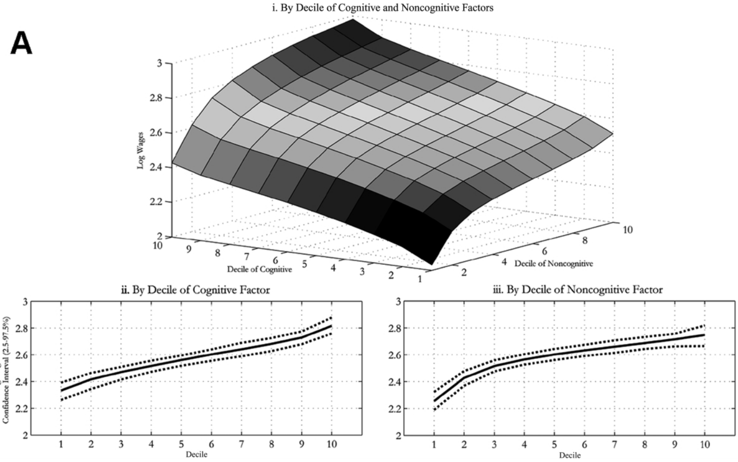
Reasons for leaving school among 16 to 25-year olds

	Fraction mentioning reason		
	Finished school immediately at minimum school-leaving age	Finished school 1–2 years after minimum school-leaving age	Finished school 3+ years after minimum school-leaving age
Had gone as far as I could	0.148	0.332	0.540
I saw no point in going on	0.295	0.172	0.193
I did not like it	0.243	0.114	0.040
I needed money	0.126	0.095	0.053
I wanted to work	0.445	0.437	0.293
Family needed money	0.039	0.034	0.013
Couldn't afford course	0.009	0.019	0.013
Had to bring up children	0.015	0.009	0.067
N	461	325	150

Notes: Sample includes 16 to 25-year olds in Britain from the 1990 Eurobarometer Youth Survey.

Oreopoulos (2007).

Skill Return



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), pp.411-482.

Cognitive Skill Return

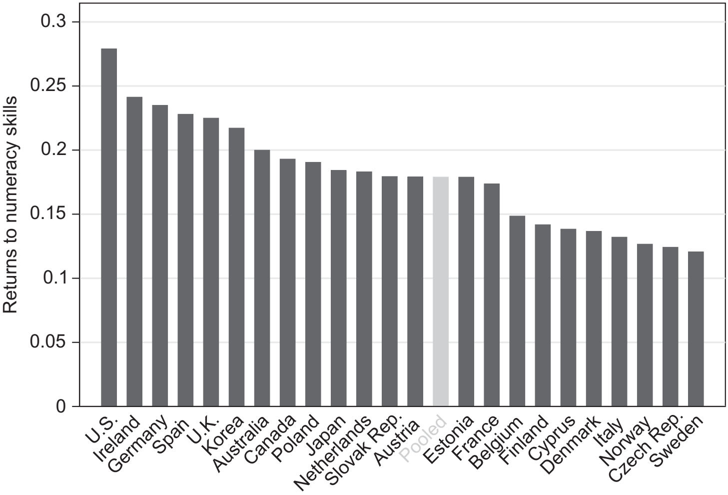
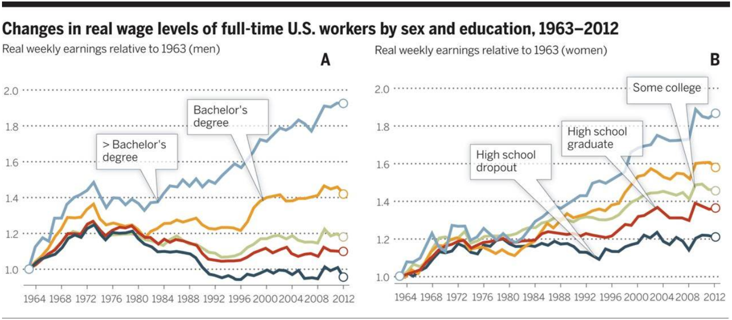


Figure 1 from Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to Skills around the World: Evidence from PIAAC. *European Economic Review*, 73, 103-130.

Returns to education

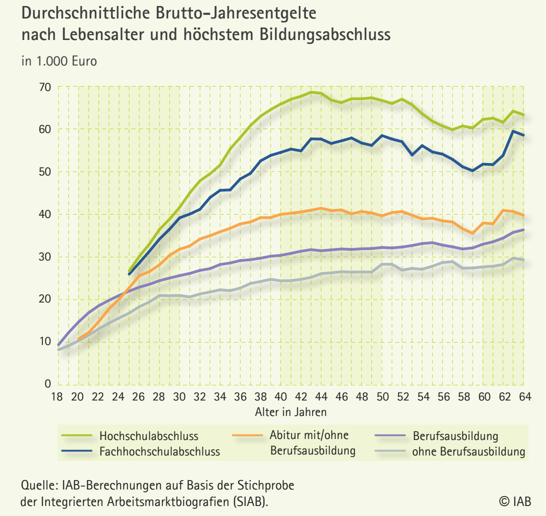
The costs of NOT going to college in the US are very high



Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the" other 99 percent". *Science*, 344(6186), 843-851.

Returns to education: unemployment

The costs of LITTLE education are also high in Germany



Returns to education: unemployment

The costs of LITTLE education are also high in Germany