Critical Periods in Child Development

Pia R. Pinger

University of Bonn, IZA

SSSI 2016

Bonn
# Table of Contents

1. Critical periods and why they matter

2. Identification
   - The Fundamental Problem of Identification
   - Contextual variation and IVs
   - Matching
   - Experiments
   - (Structural models)

3. Mechanisms
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.
IQ performance by SES

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

![IQ distribution by SES](image)

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

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)

⇒ 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)
  - Cognitive skills
  - Character traits/noncognitive skills/preferences
  - Health

- Decision-making (social)
  - 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

1 Critical periods and why they matter

2 Identification
   • The Fundamental Problem of Identification
   • Contextual variation and IVs
   • Matching
   • Experiments
   • (Structural models)

3 Mechanisms
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*.

$$D_i = \{0, 1\} \quad (1)$$

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

- Hence, two

$$\text{potential outcomes} = \begin{cases} Y_{i1} & \text{if } D_i = 1 \\ Y_{i0} & \text{if } D_i = 0 \end{cases} \quad (2)$$

$$\text{observed outcome} = Y_{0i} + (Y_{1i} - Y_{0i})D_i \quad (3)$$
Identification: The challenge

\[ Y_{0i} \quad Y_{1i} \]

\[ D_i = 1 \]
Identification: The challenge

\[ Y_i = \alpha + \rho D_i + U_i \]

Many other factors (confounding variables) exist, which influence both parental investments and the outcome \((\text{Cov}(D, U) \neq 0)\)
Identification: The challenge

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

- Endogeneity is everywhere ($\text{Cov}(D, U) \neq 0$)
- 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

1. Critical periods and why they matter

2. Identification
   - The Fundamental Problem of Identification
   - Contextual variation and IVs
   - Matching
   - Experiments
   - (Structural models)

3. Mechanisms
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 availability policy (Nilsson, forthcoming, JPE)
Contextual variation and IVs

Some **advantages** of using contextual variation are

- Provide quasi-experimental variation in investments
- Large variation in inputs
- Investments/shocks 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

Investment (D)  Outcome (Y)
Confounding variables (C)
valid instrument (Z)
```

Instrument/Exogenous Variation

```
valid instrument (Z)  Investment (D)  Outcome (Y)
Confounding variables (C)
```
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):

   $\text{Cov}(Z, U) = 0$  \hspace{1cm} (4)

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

   $\delta_1 \neq 0$  \hspace{1cm} (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.

\[
Y_i = \alpha + \rho D_i + U_i
\]

\[
= \alpha + \rho[\alpha_1 + \delta_1 Z_i + v_{i1}] + U_i
\]

\[
= (\alpha + \rho \alpha_1) + \rho \delta_1 Z_i + (\rho v_{i1} + U_i)
\]

\[
= \alpha_2 + \delta_2 Z_i + v_{i2}
\]

Hence, the reduced form coefficients are:

\[
\alpha_2 = \alpha + \rho \alpha_1
\]

\[
\delta_2 = \rho \delta_1
\]
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:

$$\rho_{IV} = \frac{\text{cov}(Y_i, Z_i)}{\text{cov}(D_i, Z_i)} = \frac{E[Y|Z = 1] - E[Y|Z = 0]}{E[D|Z = 1] - E[D|Z = 0]}.$$
Contextual variation and IVs

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

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
   → 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
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

[Graphs showing probability of hunger for Germany, The Netherlands, Greece (excluding Spain), and Other Countries from 1920 to 1960]
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: Exemplary Treatment Definition

Example of critical, famine and hunger periods for 3 individuals

German Famine (Child Sample Age Period 6−16)
Contextual variation and IVs

Compute local average treatment effect (LATE):

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

1. Nonparametric Wald estimator

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

2. 2SLS
## Contextual variation and IVs

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

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced famine being 6-16 (1 = yes)</td>
<td>0.207*** (0.020)</td>
<td>0.179*** (0.024)</td>
</tr>
<tr>
<td>German Sample</td>
<td>2.367 (3.964)</td>
<td>4.707* (2.665)</td>
</tr>
<tr>
<td>Dutch Sample</td>
<td>2.923 (2.877)</td>
<td>2.539 (2.660)</td>
</tr>
<tr>
<td>Lived in rural area at age 6</td>
<td>-0.025** (0.010)</td>
<td>-0.026*** (0.009)</td>
</tr>
<tr>
<td>Year of birth</td>
<td>-0.002* (0.001)</td>
<td>-0.002** (0.001)</td>
</tr>
<tr>
<td>Year of birth × Dutch</td>
<td>-0.002 (0.001)</td>
<td>-0.001 (0.001)</td>
</tr>
<tr>
<td>Year of birth × German</td>
<td>-0.001 (0.002)</td>
<td>-0.002* (0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.030*** (0.005)</td>
<td>4.213* (2.130)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.106</td>
<td>0.116</td>
</tr>
<tr>
<td>F-Stat.</td>
<td>102.432</td>
<td>27.079</td>
</tr>
<tr>
<td>N</td>
<td>2511</td>
<td>2511</td>
</tr>
</tbody>
</table>

Note: Number of respondents reporting hunger = 189; control group = 2322. Significance computed using standard errors clustered by country-year cells. Significance computed using bootstrap (500 replications) when applying Wald estimators. Control Variables are a dummy for whether the accommodation at age 6 has been in rural area, country fixed effects, and a country specific trend in year of birth.
Contextual variation and IVs

For age 6–16, we find

- No robust significant effects (reduced form + IV)

### Table 6: Effects of Hunger at age 6 – 16 on Adult Height in cm (males)

<table>
<thead>
<tr>
<th></th>
<th>Reduced Form</th>
<th>Instrumental Variables Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Famine at age 6 – 16</td>
<td>2SLS</td>
</tr>
<tr>
<td>Effect</td>
<td>0.269</td>
<td>1.502</td>
</tr>
<tr>
<td>(S.E.)</td>
<td>(0.325)</td>
<td>(1.808)</td>
</tr>
<tr>
<td>t-stat.</td>
<td>0.829</td>
<td>0.831</td>
</tr>
</tbody>
</table>

### Table 7: Effects of Hunger at age 6 – 16 on Adult Height in cm (females)

<table>
<thead>
<tr>
<th></th>
<th>Reduced Form</th>
<th>Instrumental Variables Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Famine at age 6 – 16</td>
<td>2SLS</td>
</tr>
<tr>
<td>Effect</td>
<td>0.229</td>
<td>2.373</td>
</tr>
<tr>
<td>(S.E.)</td>
<td>(0.258)</td>
<td>(2.773)</td>
</tr>
<tr>
<td>t-stat.</td>
<td>0.887</td>
<td>0.856</td>
</tr>
</tbody>
</table>
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

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

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

Sample 1: Childhood, sample 2: Infancy
## Contextual variation and IVs

### Males

<table>
<thead>
<tr>
<th>Famine at age 0 – 4</th>
<th>Cond. Wald</th>
<th>cond. Wald – Trend corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
<td>-0.683</td>
<td>-2.174</td>
</tr>
<tr>
<td>(S.E.)</td>
<td>( 0.268)</td>
<td>( 1.685)</td>
</tr>
<tr>
<td>t-stat.</td>
<td>-2.550</td>
<td>-1.290</td>
</tr>
</tbody>
</table>

Note: Sample size = 2511; thereof famine-exposed at age 0 – 4: = 398; thereof treated / famine-exposed at age 6 – 16 = 189 / 580. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.

### Females

<table>
<thead>
<tr>
<th>Famine at age 0 – 4</th>
<th>Cond. Wald</th>
<th>cond. Wald – Trend corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect</td>
<td>0.259</td>
<td>3.151</td>
</tr>
<tr>
<td>(S.E.)</td>
<td>( 0.288)</td>
<td>( 2.303)</td>
</tr>
<tr>
<td>t-stat.</td>
<td>0.899</td>
<td>1.368</td>
</tr>
</tbody>
</table>

Note: Sample size = 2859; thereof famine-exposed at age 0 – 4: = 417; thereof treated / famine-exposed at age 6 – 16 = 175 / 606. Regressions include control country fixed effects, a dummy for urbanization of birthplace and country specific trends in year of birth (trends cannot be included when using Wald estimator). Significance for Wald estimator computed using a bootstrap (500 repetitions). All standard errors clustered by country-year cells.
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

→ 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? (⇒ 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

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

- ⇒ 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 ⇒ methylation heritable ⇒ 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 (∼ 1%)

⇒ Non-genetic inheritance (often male/female germline)

Behavior/experience changes methylation ⇒ methylation heritable ⇒
Species change quickly
DNA Methylation and the Epigenetic Code

The ‘epigenetic’ code

**DNA methylation**
Methyl marks added to certain DNA bases repress gene activity

**Histone modification**
A combination of different molecules can attach to the “tails” of proteins called histones. These alter the activity of the DNA wrapped around them.
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

1 Critical periods and why they matter

2 Identification
   • The Fundamental Problem of Identification
   • Contextual variation and IVs
   • Matching
   • Experiments
   • (Structural models)

3 Mechanisms
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 $s = 0, 1$
  \[ Y_s = X_Y \beta_Y^s + \theta \alpha_Y^s + \varepsilon_Y^s, \tag{6} \]

- Measurement system (imperfect measures, dimension reduction, endogeneity)
  \[ M_k = \sum_{c=1}^{C_k} c \mathbf{1}(\gamma_{k,c-1} \leq M_k^* < \gamma_{k,c}), \]
  \[ M_k^* = X_M \beta_{M_k} + \theta \alpha_{M_k} + \varepsilon_{M_k}, \tag{7} \]

- Investment decision ($S = 1$)
  \[ S = \mathbf{1}(S^* > 0), \]
  \[ S^* = X_S \beta_S + \theta \alpha_S + \varepsilon_S, \tag{8} \]
Model the unobserved heterogeneity

Latent factor:

\[ \theta \sim \mathcal{N}(0; \sigma^2_\theta) \quad \theta \perp \perp X \perp \perp \varepsilon \]

Latent factors: normally distributed or mixture of normals

\[ \varepsilon \perp \perp X, \quad \varepsilon \perp \perp \theta, \quad \varepsilon_j \perp \perp \varepsilon_{j'}, \quad (9) \]

Error terms: standard normal (probits and ordered probits) or mixture of normals
Model the unobserved heterogeneity

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

\[
\Omega^* \equiv V(S^*, Y, M^* \mid X),
\]  

(10)

Factor loadings can be identified from the ratios of observed covariances

\[
\frac{\text{Cov}(M_1^*; M_2^* \mid X)}{\text{Cov}(M_2^*; M_3^* \mid X)} = \frac{\alpha_{M_1} \alpha_{M_2} \sigma^2_{\theta}}{\alpha_{M_2} \alpha_{M_3} \sigma^2_{\theta}} = \frac{\alpha_{M_1}}{\alpha_{M_3}}.
\]  

(11)

Either set $\alpha_{M_1} = 1$ or $\sigma^2_{\theta} = 1$ 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

\[
\theta \sim \mathcal{N}\left(0; \sigma^2_\theta\right), \quad (\varepsilon_S, \varepsilon_D^0, \varepsilon_D^1, \varepsilon_M, \ldots, \varepsilon_M^k)' \sim \mathcal{N}(0; \Sigma), \quad (12)
\]

Mixture for error term of the wage equation:

\[
\varepsilon^s_Y \sim \sum_{h=1}^{H_s} \pi^s_h \mathcal{N}\left(\mu^s_h, (\omega^s_h)^2\right), \quad \mathbb{E}(\varepsilon^s_Y) = \sum_{h=1}^{H_s} \pi^s_h \mu^s_h = 0, \quad (13)
\]
Model the unobserved heterogeneity

Derive the likelihood

\[
\mathcal{L}(\psi \mid S, Y, M, X) = \int_{\Theta} \prod_{s=0}^{1} \Pr(S = s \mid X, \theta, \psi) 1(S=s) \\
\times \prod_{s=0}^{1} f(Y_s \mid X, \theta, \psi) 1(S=s) \\
\times \prod_{k=1}^{K} f(M_k \mid X, \theta, \psi) dF(\theta),
\]

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)

\[ Y_i = \alpha + \beta X_i + \rho D_i + c_i + \varepsilon_i^Y \]

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
Results

![Height and Standardized Cognitive Test Scores vs Age at Migration](image)

Figure 1. Sibling fixed-effects estimates of height and standardized cognitive test scores as a function of age at migration. The figure plots the estimates of the effect of each age at migration from 0 to 17 on height and standardized cognitive test scores, corresponding to the estimates in columns (1) and (4) of Table 2. All fixed effect estimates are from models controlling for age at enlistment and being first born. The dotted lines represent 95% confidence intervals.

The sample used for this particular analysis only consist of foreign-born siblings and the reference category is thus siblings immigrating to Sweden before the age of 1.

Note that for hospitalizations, the sample sizes and the fraction of nonzero cases are small. Although we find a clear pattern in the coefficients, with age at migration being positively related to the probability of a hospitalization, few of the estimates at this age range are significant. In Table A.2, we find that the risk of hospitalization at ages 11–15 increases in the age at migration for males (which in this case ranges from 0 to 10). However, this is not confirmed for hospitalizations at higher ages. It should be noted that hospitalizations are imperfect measures of health since issues such as health-care seeking behavior may affect the measurement.

A natural question to ask is whether the adverse effects of age at migration on health and cognitive ability translate into adverse effects on education and income. Columns (5)–(7) of Table 3 and Figure 3 show this to be the case. Concerning education, we focus on years of schooling. Typically, education decreases in the age at migration among both males and females and all estimates beyond age 0 are significant. The effects are substantial; a female arriving at age 14, for instance, obtains on one year less of schooling on average compared to a Swedish-born sister. The corresponding estimate for males is 0.6 years of schooling.

In column (7) of Table 3 we show effects on yearly log income for men (see also the lower left panel of Figure 3). We focus on men since we cannot distinguish between labor supply and the wage rate in our data, and men are more likely to work full-time.

In Table A.3 of the Online Appendix we analyze two additional educational outcomes; the probability of entering university and the probability of studying beyond high school. Those results confirm the results for years of schooling. In particular, age at migration is strongly negatively related to the probability of entering university.

---

**Critical Periods 65 / 104**

Pia R. Pinger
Table of Contents

1 Critical periods and why they matter

2 Identification
   - The Fundamental Problem of Identification
   - Contextual variation and IVs
   - Matching
   - Experiments
   - (Structural models)

3 Mechanisms
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**

⇒ 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

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: \( P(\text{child high} | \text{parent high}) / P(\text{child high} | \text{parent low}) \)

Ratio 2: \( P(\text{child middle} | \text{parent middle}) / P(\text{child middle} | \text{parent low}) \)

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

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

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

[Graph showing the treatment effect on attending upper secondary school in grade 5.]
Results

Treatment effect on attending upper secondary school in grade 5

Treatment effect heterogeneity
Low-SES households

- Female
- Male
- 2-parent HH
- 1-parent HH
- Low poverty HH
- High poverty HH
- High edu HH
- Low edu HH

Pia Pinger
1 Critical periods and why they matter

2 Identification
   - The Fundamental Problem of Identification
   - Contextual variation and IVs
   - Matching
   - Experiments
   - (Structural models)

3 Mechanisms
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 (\textit{indirect} output effect)
  - Unobserved residual effects (\textit{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

⇒ Do some kind of Oaxaca-Blinder-type decomposition to decompose the ATE

\[ E(Y_1 - Y_0) = \sum_{j=1}^{J} \theta^j E(M^i_1 - M^i_0) + \tau_1 - \tau_0 \]
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

⇒ Need some kind of exogenous variation (instrument) for mediators

One possibility: estimate the following structural simultaneous equation system by two-stage least squares (\textit{observed}=\tau_M \ast \theta \text{ and residual}=\tau_r):

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

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

<table>
<thead>
<tr>
<th>Reason for leaving school</th>
<th>Fraction mentioning reason</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Finished school immediately at minimum school-leaving age</td>
</tr>
<tr>
<td>Had gone as far as I could</td>
<td>0.148</td>
</tr>
<tr>
<td>I saw no point in going on</td>
<td>0.295</td>
</tr>
<tr>
<td>I did not like it</td>
<td>0.243</td>
</tr>
<tr>
<td>I needed money</td>
<td>0.126</td>
</tr>
<tr>
<td>I wanted to work</td>
<td>0.445</td>
</tr>
<tr>
<td>Family needed money</td>
<td>0.039</td>
</tr>
<tr>
<td>Couldn’t afford course</td>
<td>0.009</td>
</tr>
<tr>
<td>Had to bring up children</td>
<td>0.015</td>
</tr>
<tr>
<td>N</td>
<td>461</td>
</tr>
</tbody>
</table>

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

Oreopoulos (2007).
Cognitive Skill Return

Figure 1 shows clearly that returns to skills in the U.S. labor market might not adequately characterize the labor-market situation in other countries. But additionally, there are multiple reasons to believe that these estimates understate the true impact of skills across the lifecycle.

To test for age varying returns to skills, Table 3 uses the extended sample of employees aged 25–65 and allows the returns to skills to differ between entry-age (25–34), prime-age (35–54), and exit-age (55–65) workers. Results suggest that there are indeed significant age differences in returns to skills: Across all countries, returns to numeracy are only 14 percent in the baseline category of workers aged 25–34 and are 4 percentage points (some 30 percent) higher among workers aged 35–54 or 55–65. In the aggregate, there are no significant differences between prime-age and exit-age workers. In fact, estimates of returns to skills increase steadily with age until age 35 and are reasonably flat from there on, getting only slightly smaller beyond age 55 (Fig. 2).

Thus, the focus on early-career earnings in the existing literature is likely to downward bias estimated returns to skills. As earnings of prime-age workers tend to be good proxies for lifetime earnings (Haider and Solon, 2006), returns estimated in the prime-age category will more likely capture lifetime returns to skills.

Looking at individual countries, returns to skills are significantly higher for prime-age workers than for entry-age workers in 16 out of the 23 countries. In all but one country (the Slovak Republic), the interaction between skills and the dummy variable indicating prime-age workers is positive. Insignificant age interactions are observed primarily in the transition economies in Eastern Europe (the Czech Republic, Estonia, Poland, and the Slovak Republic), where transition dynamics are likely to be important in explaining earnings across workers of different ages. These transition economies experienced a radical change in their labor markets after the fall of the Iron Curtain. In the process of restructuring after privatization (e.g., Frydman et al., 1999; Mertaugh and Hanushek, 2005), workers who are in their prime age or exit age today may have lost much of their specific human capital during the early 1990s and may face a job–skill mismatch, whereas workers in their entry age today may have acquired skills more relevant for the current labor market. Without the four Eastern European transition countries, returns to skills of prime-age and exit-age workers are even 5.1 percentage points (or 38 percent) higher than among entry-age workers (13.4 percent) in the pooled sample.

4.3. Returns to skills and interaction with years of schooling

As discussed, our focus on measured skills differs from the standard Mincer approach of using years of schooling as the sole measure of human capital. When adding years of schooling to our baseline model, as shown in Table 4, both numeracy and literacy skills are significant and positive across all age groups. The returns to years of schooling are highest among prime-age workers and decrease with age, consistent with the theory of human capital accumulation. However, the interaction between years of schooling and age is not significant, indicating that the returns to schooling are not affected by the age of the worker.
Returns to education

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

Returns to education: unemployment

The costs of LITTLE education are also high in Germany

Durchschnittliche Brutto-Jahresentgelte nach Lebensalter und höchstem Bildungsabschluss

in 1.000 Euro

Quelle: IAB-Berechnungen auf Basis der Stichprobe der Integrierten Arbeitsmarktabsiografien (SIAB).
© IAB

Pia Pinger
Mentoring and Education 103 / 104
In aller Kürze
Lebensverdienste nach Qualifikation

Bildung lohnt sich ein Leben lang von Achim Schmillen und Heiko Stüber


Die sogenannte Bildungsprämie beträgt für ein Abitur (mit oder ohne Berufsausbildung) etwa 500.000 Euro, ein Fachhochschulstudium ist durchschnittlich 900.000 Euro wert und ein Hochschulstudium rund 1.250.000 Euro, jeweils im Vergleich zu Personen ohne Berufsausbildung.


- Personen ohne Berufsausbildung und ohne Abitur,
- Personen mit Berufsausbildung und ohne Abitur,
- Abiturienten mit oder ohne Berufsausbildung,
- Fachhochschulabsolventen und
- Hochschulabsolventen.

Dabei zeigt sich ganz klar: Bildung ist viel wert. So verdient ein Beschäftigter mit Berufsausbildung im Verlauf des ganzen Arbeitslebens durchschnittlich bis zum 2,7-Fachen des Seniors verdienen, was Personen ohne beruflichen Abschluss erhalten. Aber auch eine Berufsausbildung ist ihr Geld wert. Über das ganze Erwerbsleben hinweg addieren sich die Bildungsprämien zu beachtlichen Summen – für alle Berufsabschlüsse, für Männer wie Frauen und in Ost wie West. Allerdings unterscheiden sich die Bildungsprämien zwischen den betrachteten Gruppen durchaus erheblich.

Quelle: IAB-Berechnungen auf Basis der Stichprobe der Integrierten Arbeitsmarktdiagramme (SIAB).