

Intergenerational mobility, intergenerational effects, the role of family background, and equality of opportunity: a comparison of four approaches

Anders Björklund

SOFI, Stockholm University

SITE, Stockholm, September 2

The four approaches

1. Intergenerational mobility (or transmission or persistence)
2. Intergenerational effects
3. The role of family background=sibling correlations
4. Equality of opportunity

My goal

Approach	Questions	Results
1. Intergenerational mobility (or transmission)		
2. Intergenerational effects		
3. The role of family background=sibling correlations		
4. Equality of opportunity		

Approach	Some key economists (except some from SOFI.....)	Top publications
1. Intergenerational mobility	Solon, Corak, Mazumder, Blanden, Machin, Chetti et al.	AER, QJE, EJ, JHR, JofPubEcon
2. Intergenerational effects	Plug, Black&Deveruex& Salvanes	AER, JPE, JEL, QJE
3. The role of family background=sibling correlations	Solon, Mazumder, Schnitzlein	JHR, Jof PubEcon, JofPopEcon
4. Equality of opportunity	Roemer, Ferreira	JofPubEcon, SoCh&Wel, RIW, JofEcIneq

My motivation



1. Academic: Four subfields that have much to learn from each other but are quite separated. Does research of today focus on the most important issues?
2. Public policy: Is family background *very* important, or only *somewhat* important? Crucial when we evaluate our societies from an egalitarian point of view.

Roadmap



1. Intergenerational mobility: approach and findings
2. Intergenerational effects: approach and findings
3. Sibling correlations: approach and findings
4. Equality of opportunity: approach and findings
5. Conclusions: what the literatures can learn from each other, and where reserach should go

My presentation is not as comprehensive as it might look:

1. I don't talk about "the structural approach" that estimates the parental production function and their investment behaviour (cf. Heckman et al.)

2. I don't consider social mobility, i.e., class mobility. I stick to results about
 - a. long-run (log) income and earnings
 - b. years of education

1. Intergenerational mobility (or transmission)

Prototypical model:

$$Y_i^{son} = \alpha + \beta Y_i^{father} + e_i$$

β : regr. coefficient or elasticity (IGE)

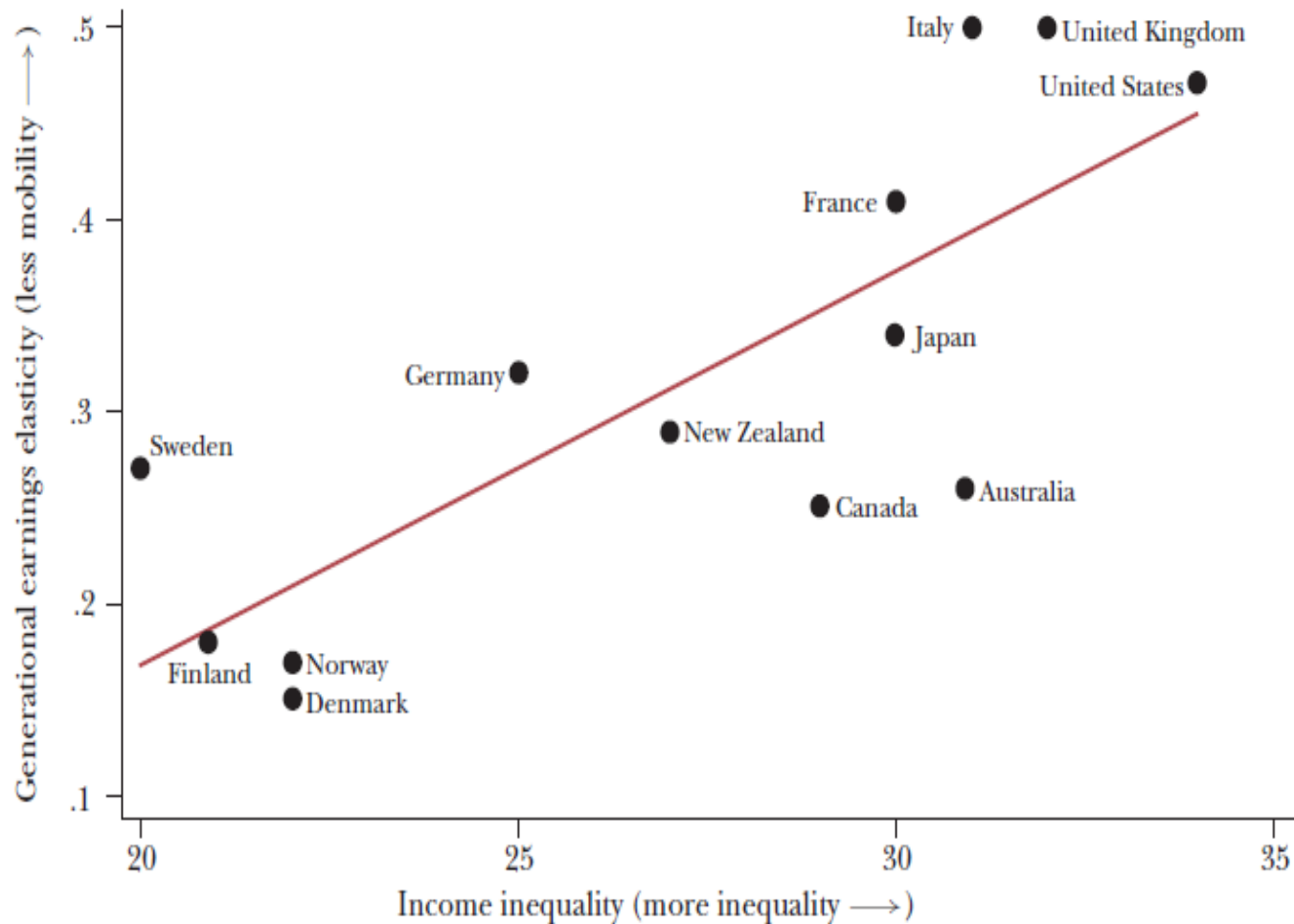
$$\text{Correlation} = \text{IGC} = \beta (\sigma^{\text{father}} / \sigma^{\text{son}})$$

Sometimes nonlinearities

Sometimes rank correlations

Sometimes transition matrixes

The Great Gatsby Curve (from Corak, JEP 2013)



Some results, years of schooling (from Hertz, BEJEAP 2008)

Country	Regression coefficient	Correlation
USA	.46	.46
Denmark	.49	.30
Finland	.48	.33
Norway	.40	.35
Sweden	.58	.40
Great Britain	.71	.30
The Netherlands	.58	.36
Belgium	.41	.40
Italy	.67	.40

Conclusions:

- We have learnt some about country differences.
- Associations are not that strong. Correlations from 0.2 to 0.5 imply R-squares of 0.04-0.25.
 - Scatter plots reveal a lot of mobility!

2. Intergenerational effects

- Meaning:
 - What are the causal effects of thought interventions that change parents' income or education?
 - This is something (potentially very) different from descriptive intergenerational mobility patterns

2. Intergenerational effects

Prototypical model

$$Y_i^{offspring} = \alpha + \beta_1 Y_i^{father} + \beta_2 Y_i^{mother} + e_i$$

Intergenerational effects: empirical strategies.

Strategies to get the "causal effect":

1. Twin-differences: $\Delta Y^{\text{cousins}} = \alpha + \beta \Delta Y^{\text{twinparents}}$
 - takes some genetics and common environment out
2. Adoptive parents
 - eliminates genetic transmission (if random assignment)
3. IV, reforms
 - gives exogenous variation in parental resources

Some key European studies:

Norway: 1. Black, Devereux & Salvanes (2005). 2. Loken (2010). 3. Haegeland et al. (2010). Sweden: 1. Holmlund, Lindahl, Plug (JEL, 2011) 2. Amin, Lundborg & Rooth (2013). 3. Björklund, Lindahl, Plug (2006). US: several.

General pattern in the results

Estimates of causal effects in general in the range 0-60% of the IG associations

My own overall estimate: 33% of the IG-coefficients

In some contexts, however, the causal income (education) effects might be very large

3. The sibling correlation

$$y_{ij} = a_i + b_{ij}$$

a_i common to all siblings in family i

b_{ij} unique to individual j in family i

a_i and b_{ij} orthogonal by construction. Thus:

$$\sigma_y^2 = \sigma_a^2 + \sigma_b^2$$

The family share of the outcome variance is:

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_b^2}$$

ρ is also the sibling correlation

A sibling correlation captures more than an intergenerational correlation (IGC)

Sibling correlation = $(IGC)^2$ + other shared factors that are uncorrelated with parental y

An omnibus measure! Captures both observed and unobserved family background (and neighborhood) factors

Yet it is a lower bound!

Some estimates of brother correlations in long-run earnings

Country	Estimate	Study
USA	.49	Mazumder (2008)
Denmark	.23	Björklund et al. (2002)
Finland	.26	Björklund et al. (2002)
Norway	.14	Björklund et al. (2002)
Sweden	.25	Björklund et al. (2002)
Sweden	.32	Björklund, Jäntti & Lindquist (2009)
Germany	.43	Schnitzlein (2013)

Some estimates of sibling correlations in years of schooling

Country	Sibling type	Estimate	Study
USA	Mixed sexes	.60	Mazumder (2008)
Norway	Mixed sexes	.41	Björklund & Salvanes (2010)
Sweden	Brothers	.43	Björklund & Jäntti (2012)
Sweden	Sisters	.40	Björklund & Jäntti (2012)
Germany	Brothers	.66	Schnitzlein (2013)
Germany	Sisters	.55	Schnitzlein (2013)

These quite high numbers are only lower bounds. What is missing?

1. Full siblings have only about half of (initial) genes in common. But each individual has 100% of her (initial) genes from the parents.
2. Not all environmental experience and "shocks" are shared, only some. Thus some environmental stuff is missing.
3. Differential treatment by parents. Will not be captured if it creates differences, but is part of family background.

Raising the lower bound: MZ-twins?

1. They share all (initial) genes (GOOD)
2. They share more environment and more "shocks" (GOOD)
3. They might interact more and affect each other in ways that have no counterpart in the general population (BAD)

Because of (3), an MZ-correlation might be an upper bound of family background

Sibling correlations for MZ-twins vs. full siblings: Swedish results

Outcome	Sibling type	Full sibling	MZ-twins
Earnings	Brothers	.22	.73
Earnings	Sisters	.16	.40
Schooling	Brothers	.44	.75
Schooling	Sisters	.40	.73

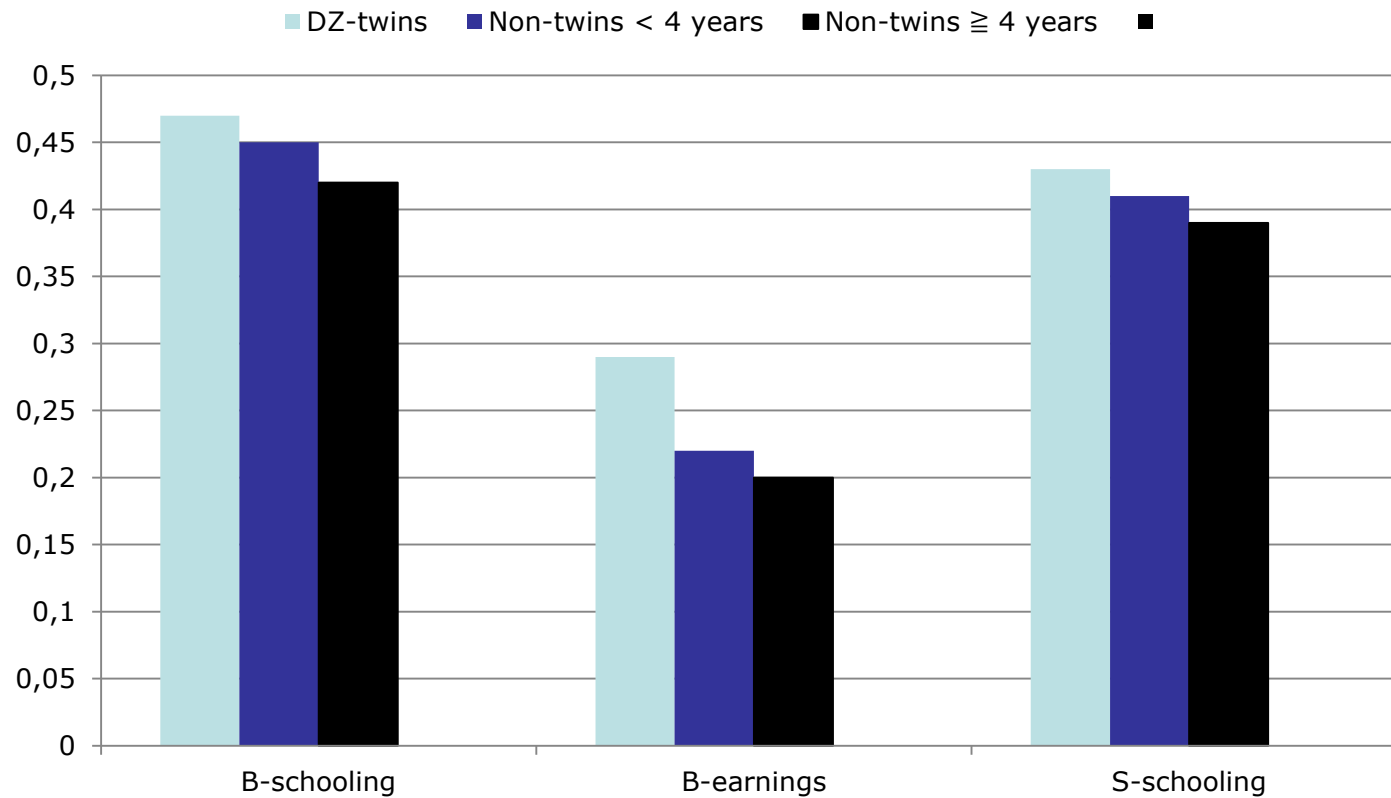
But how important is interaction among siblings for the family component?



Suggestive evidence from sibling similarity by age difference:

1. Interaction is more common among closely spaced siblings. Thus sibling correlations should be higher for closely spaced siblings if interaction is important
2. However, closely spaced siblings are also exposed to similar "shocks"
3. Similar sibling correlations by age difference suggest that interaction is not that important. Permanent family factors would then explain similarity.

Sibling correlations by age difference



One aspect of differential treatment: birth order

- Differential treatment impossible to measure completely
- Black, Devereux & Salvanes (2005, 2011) show convincingly (within-family estimates) that first-born do better than later-born.
- Birth order can be integrated into our analysis
- A dummy for first-born
- We estimate within-family effects for first-born
- The implied variation is part of family background, but not shared by siblings!

The contribution of birth order

Out- come	Reg coeff.	Variance component (%)			
		<u>Family</u>	<u>Indiv.</u>	<u>First- born</u>	<u>Total</u>
<u>Brothers</u>					
Earnings	.024	21.6	78.3	0.0	100
Schooling	.248	44.0	55.6	0.4	100
<u>Sisters</u>					
Earnings	.027	16.1	83.9	0.0	100
Schooling	.230	40.4	59.2	0.2	100

How much do IG-mobility estimates explain?

Use:

Sibling correlation = $(IGC)^2$ + other shared factors that are uncorrelated with parental y

Swedish estimates:

Sibling correlations vs. Intergenerational correlations

Outcome	Sibling correlation	(IGC) ² =R ²	Other factors
<u>Brothers</u>			
Earnings	.24	.02	.22
Schooling	.46	.15	.31
<u>Sisters</u>			
Earnings	.21	.01	.20
Schooling	.40	.11	.29

An extension

Adding one more parent to the equation does not change the results very much.

Summing up about IG mobility and IG effects and sibling correlations:



1. IGCs capture only very little: "The tip of the iceberg".

2. Intergenerational causal effects capture even less. "The tip of the tip of the iceberg".

3. But sibling correlations reveal a large role for something in the family. Unobserved factors are very important

4. And yet siblings correlations are lower bounds
 - a. MZ-twin-correlations are possibly upper bounds, but not necessarily so. They suggest a very big role for family background
 - b. Although birth-order effects do not account for much, other differential treatment effects might be important.

4. Equality of opportunity approach

$$Y_i = \alpha C_i + \beta E_i + \varepsilon_i$$

$$E_i = \delta C_i + v_i$$

C: set of circumstances: factors beyond individual control, for which individuals should not be held responsible (such as parental resources)

E: set of effort variables: all choices for which society holds the individual accountable (such as labor supply)

Reduced form: $Y_i = (\alpha + \beta\delta)C_i + \beta v_i + \varepsilon_i$

EO-approach: implementation

- Estimate the reduced form
- Measure:
 - The fraction of variance which is explained by circumstances: R^2
 - Or derive the inequality (according to a suitable measure of inequality) that is generated by circumstances. Compare this inequality with total inequality.: $Ineq(\text{due to circ.})/Ineq(\text{total})$.
- Some empirical approaches consider the role of luck. Some try to measure effort and include it in the outcome equation. Also other nice tricks.

A first reaction from me (coming from the other literatures)

These researchers try to fill in gap between IGC and the sibling correlation:

Sibling correlation = $(IGC)^2$ + other shared factors that are uncorrelated with parental y

What circumstance variables have been used? And which are valid ones?

Let us look at some results

Study	Country, outcome	Circumstances	R ²	Gini	MLD, Theil (1)
Bourguignon et al. (2007)	Brazil, earnings	Race, par educ, region, father's occ status	.24-.30		.13-.34
Bourguignon et al. (2007)	Brazil, schooling	Race, par educ, region, father's occ status	.34-.43		N.a.
Ferreira et al. (2011)	Turkey, wealth	Region, Par educ, # of sibs, language	.27		.31
Björklund et al. (2012)	Sweden, earnings	Par inc (4), Par educ (3), Par sep (2), # of sibs (3), IQ age 18 (4), BMI (4)	.06	.24	.10

Study	Country, outcome	Circumstances	R ²	Gini	MLD
Niehuis & Peichl (forthc)	Germany. Long-run earnings	Gender, foreing-born, father's occ and edu, urbanization, height, year of birth			Ca. .25
Niehuis & Peichl (forthc.)	USA, Long-run earnings	Gender, foreing-born, father's occ and edu, urbanization, height, year of birth			Ca. .25

How do the results compare to those from sibling correlations?

- In general lower explanatory power than what sibling correlations predict.
- And yet sibling correlations are lower bounds of the importance of family background. What about omitted genetic influence captured by MZ-twins?
- But are all factors shared by siblings really circumstances?

Major problem: are omitted variables effort or circumstances?

For many circumstance variables there is a causal effect literature:

Variable	Results from causal effect studies	Are omitted variables circumstances?
Parental income and education	Intergenerational effects considerably lower than transmission coefficients	Maybe, because that is what is controlled away by twinning and using adoptive parents
Parental separation	Effects lower than descriptive regr coefficients	?
Family size, or number of siblings	Effects lower than descriptive regr coefficients	?

The same problem applies to sibling correlations:

- Are all factors shared by siblings really to be considered circumstances?
 - Maybe to some extent since you have not chosen your siblings.
- To find out more, we have to understand the mechanisms that siblings share.
 - We need to study the part of the iceberg that is below the water!

Circumstance variables that have not been used (so much)

- Grandparents (and other relatives)
 - Recent study that adds grandparents:
 - Earnings: R^2 from .064 to .067
 - Schooling: R^2 from .152 to .164
- School and teacher quality
- Better skill measures from early childhood, e.g. non-cognitive skills
- Health indicators from early childhood, including birth weight
- Explicit genetic information. Difficult though.

What are the prospects of coming close to sibling correlations?

- Probably low!
- My guess: very important circumstances are not only unobserved, they are also in practice unobservable.
- In that case indirect variance component approaches are needed to find out how much there is: sibling correlations including MZ-twin correlations

**Time to sum up and conclude about
the four approaches:**

1. Intergenerational mobility



- Does not directly address the inequality-of-opportunity question
- But: provides an easy-to-understand picture that the public policy debate seems to appreciate:
 - It tells us about "the rise and fall of families".
 - The cross-country pattern has received a lot of public attention
- Maybe easier to study and interpret country-differences and trends in intergenerational mobility than in the combined importance of a set of circumstances

2. Intergenerational effects:

- This approach addresses well-defined questions of high scientific and public-policy importance
- But estimated effects are small in the sense that they explain very little of inequality of income and schooling
- But effects might be large in other dimensions: the benefit-cost ratio of some interventions that change parental income and education might be high

3. Sibling correlations:

- A sibling correlation does not provide the answer to any well-defined scientific or public-policy questions
- A sibling correlation should rather be used as a warning signal (“benchmark”) whether researchers have missed important family background factors
- And this signal is very strong:
 - Considering that sibling correlations are lower bounds, the magnitudes should make all scholars with egalitarian attitudes concerned!
 - But yet the results are often neglected!

4. Equality of opportunity:

- Finding the explanatory power of circumstances is the natural correct approach to measuring inequality of opportunity!
 - But ideally: a multivariate model of circumstances' causal effects is needed.
- But many important circumstances are not observed in typical data sets. And are probably not observable even with very ambitious data collection efforts.
- But maybe one can find the most important circumstances. But that should ideally be done with a causal analysis.

Final words

- We have learnt that the family is very important as a source of inequality!
- In order to evaluate this result, we need to learn more about why.
 - Some hints about what is below the water is more valuable than an even more detailed picture of what is above!
- Workshops like this are needed!