The Intergenerational Transmission of Cognitive Skills:

An Investigation of the Causal Impact of Families on Student Outcomes

Eric A. Hanushek[1] Babs Jacobs[2] Guido Schwerdt[3]
Rolf van der Velden[2] Stan Vermeulen[2] Simon Wiederhold[4]

[1]Hoover Institution, U Stanford [2]Research Centre for Education and the Labour Market (ROA)
[3]U Konstanz [4]KU Eichstaett-Ingolstadt

March 24, 2022

Motivation

- Parental inputs are remarkably related to the lifetime outcomes of children
- Strong **intergenerational persistence** in:
  - Income/Wealth (e.g., Solon (1999), Bjoerklund and Jaentti (2011), Black et al. (2020))
  - Education (e.g., Adermon, Lindahl, and Palme (2021), Dodin et al. (2021), Lundborg, Plug, and Wuertz Rasmussen (2021))
  - IQ (e.g., Black, Devereux, and Salvanes (2009), Anger and Heineck (2010))
- What are the **fundamental causes** of the persistence of economic outcomes across generations?
- We provide **causal evidence** on the transmission of **cognitive skills** within families

This Paper

1. We use **unique linked data** to estimate how skills of parents — measured by math and language tests — affect skills of their children in similar tests
2. We exploit **within-family between-subject variation in cognitive skills** to provide arguably causal estimates of the key intergenerational skill transmission parameter
3. We eliminate remaining concerns of bias by **isolating determinants of parent cognitive skills outside the family** in novel IV estimations

Data

- We develop the Intergenerational Transmission of Skills (ITS) database
- It combines **survey data on parents** gathered in the 1970's and 1980's with **register data on children** available at Statistics Netherlands
- Parent surveys
  - **Large, nationally representative panels** of students
  - Two cohorts sampled in the first year of secondary education (1977 and 1989), one cohort sampled in the last year of primary education (1982)
  - Information on **parent characteristics** (e.g., gender, migration background, number of siblings), **grandparent characteristics** (e.g., education, social background), and **municipality** at the time of testing
- Linked dataset contains information on math and language skills of **25,483 parents and 41,774 of their children**

Cognitive Skills

- CITO (Central Institute for Test Development) test is taken in the **final year (grade 6)** of primary education
  - CITO test is **high-stakes** as it determines track allocation
- Statistics Netherlands has **register data of all schools** that participated in the CITO test from 2005/2006 to 2018/2019
- Some **sample selectivity** in the parent data due to limited time window of the children data
  $\longrightarrow$ 1977 cohort: parents relatively old when they had children
  $\longrightarrow$ 1989 cohort: parents relatively young when they had children
- Test of parents was a **shortened version of the children's CITO test**, administered at the same age as the children's test

Descriptive statistics

Do These Skills Matter Later in Life?

	Higher Education	STEM field of study	Log hourly wage	Personal income	Household income	Household wealth
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
Math skills	0.144	0.021	0.114	5.453	4.074	5.204
	(0.002)	(0.003)	(0.002)	(0.109)	(0.111)	(0.113)
R-squared	0.238	0.111	0.277	0.317	0.079	0.133
Observations	61,756	28,264	41,928	53,099	55,320	53,963
	Panel B: Language					
Language skills	0.132	0.001	0.102	4.656	3.265	4.133
	(0.002)	(0.003)	(0.002)	(0.109)	(0.112)	(0.114)
R-squared	0.224	0.109	0.262	0.308	0.071	0.120
Observations	61,756	28,264	41,928	53,099	55,230	53,963
	Panel C: Math and language					
Math skills	0.101	0.031	0.084	4.149	3.280	4.219
	(0.002)	(0.003)	(0.002)	(0.134)	(0.138)	(0.139)
Language skills	0.073	-0.017	0.053	2.256	1.369	1.696
	(0.002)	(0.003)	(0.002)	(0.133)	(0.138)	(0.140)
R-squared	0.254	0.112	0.287	0.320	0.081	0.135
Observations	61,756	28,264	41,928	53,099	55,230	53,963

Notes: Least squares regressions. Sample: Pooled sample of all individuals (parents and non-parents) in the three survey cohorts. All wage, income, and wealth variables are measured 30 years after the skill assessment took place; higher education degree completion is based on the highest educational degree obtained by the individual observed in the survey data. Cognitive skills are standardized to SD 1 in full sample in each survey cohort. All regressions control for individual's gender, migration background, number of siblings, age of parents at the time of individual's birth, survey indicators, education and social background of grandparents as well as municipality fixed effects. Standard errors clustered at the individual level in parentheses. *Data sources:* Administrative data; pooled survey data.

Do These Skills Matter Later in Life?

	Higher Education	STEM field of study	Log hourly wage	Personal income	Household income	Household wealth
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
Math skills	0.144	0.021	0.114	5.453	4.074	5.204
	(0.002)	(0.003)	(0.002)	(0.109)	(0.111)	(0.113)
R-squared	0.238	0.111	0.277	0.317	0.079	0.133
Observations	61,756	28,264	41,928	53,099	55,320	53,963
	Panel B: Language					
Language skills	0.132	0.001	0.102	4.656	3.265	4.133
	(0.002)	(0.003)	(0.002)	(0.109)	(0.112)	(0.114)
R-squared	0.224	0.109	0.262	0.308	0.071	0.120
Observations	61,756	28,264	41,928	53,099	55,230	53,963
	Panel C: Math and language					
Math skills	0.101	0.031	0.084	4.149	3.280	4.219
	(0.002)	(0.003)	(0.002)	(0.134)	(0.138)	(0.139)
Language skills	0.073	-0.017	0.053	2.256	1.369	1.696
	(0.002)	(0.003)	(0.002)	(0.133)	(0.138)	(0.140)
R-squared	0.254	0.112	0.287	0.320	0.081	0.135
Observations	61,756	28,264	41,928	53,099	55,230	53,963

Notes: Least squares regressions. Sample: Pooled sample of all individuals (parents and non-parents) in the three survey cohorts. All wage, income, and wealth variables are measured 30 years after the skill assessment took place; higher education degree completion is based on the highest educational degree obtained by the individual observed in the survey data. Cognitive skills are standardized to SD 1 in full sample in each survey cohort. All regressions control for individual's gender, migration background, number of siblings, age of parents at the time of individual's birth, survey indicators, education and social background of grandparents as well as municipality fixed effects. Standard errors clustered at the individual level in parentheses. *Data sources:* Administrative data; pooled survey data.

Do These Skills Matter Later in Life?

	Higher Education	STEM field of study	Log hourly wage	Personal income	Household income	Household wealth
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
Math skills	0.144	0.021	0.114	5.453	4.074	5.204
	(0.002)	(0.003)	(0.002)	(0.109)	(0.111)	(0.113)
R-squared	0.238	0.111	0.277	0.317	0.079	0.133
Observations	61,756	28,264	41,928	53,099	55,320	53,963
	Panel B: Language					
Language skills	0.132	0.001	0.102	4.656	3.265	4.133
	(0.002)	(0.003)	(0.002)	(0.109)	(0.112)	(0.114)
R-squared	0.224	0.109	0.262	0.308	0.071	0.120
Observations	61,756	28,264	41,928	53,099	55,230	53,963
	Panel C: Math and language					
Math skills	0.101	0.031	0.084	4.149	3.280	4.219
	(0.002)	(0.003)	(0.002)	(0.134)	(0.138)	(0.139)
Language skills	0.073	-0.017	0.053	2.256	1.369	1.696
	(0.002)	(0.003)	(0.002)	(0.133)	(0.138)	(0.140)
R-squared	0.254	0.112	0.287	0.320	0.081	0.135
Observations	61,756	28,264	41,928	53,099	55,230	53,963

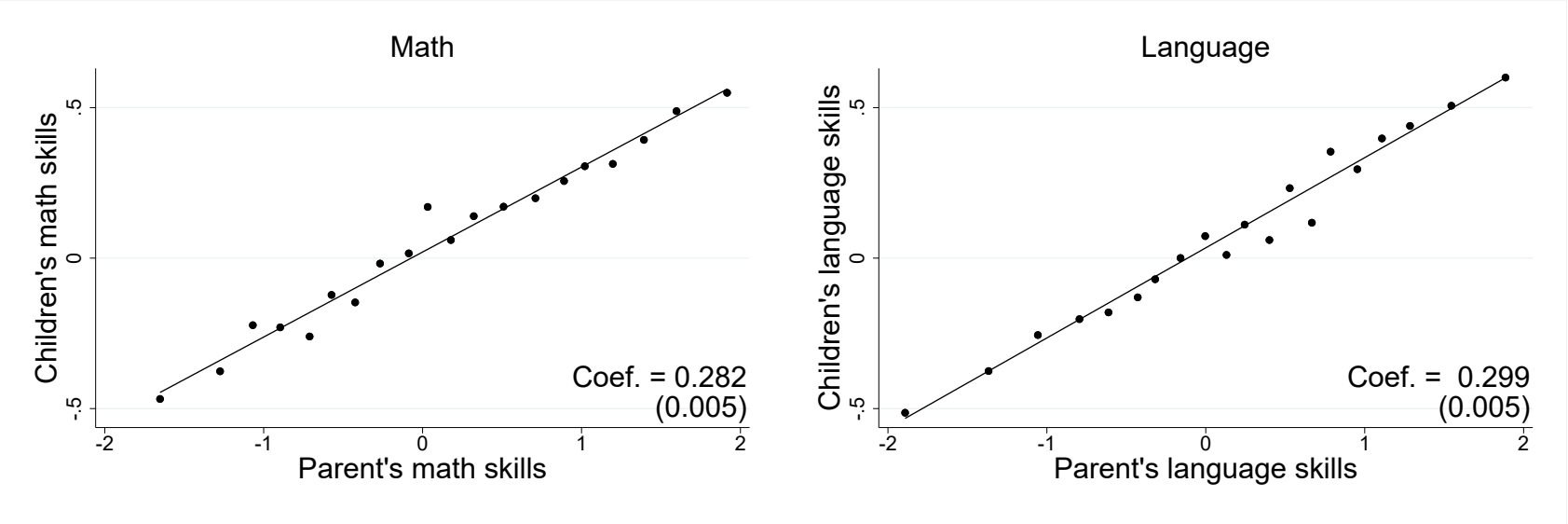
Notes: Least squares regressions. Sample: Pooled sample of all individuals (parents and non-parents) in the three survey cohorts. All wage, income, and wealth variables are measured 30 years after the skill assessment took place; higher education degree completion is based on the highest educational degree obtained by the individual observed in the survey data. Cognitive skills are standardized to SD 1 in full sample in each survey cohort. All regressions control for individual's gender, migration background, number of siblings, age of parents at the time of individual's birth, survey indicators, education and social background of grandparents as well as municipality fixed effects. Standard errors clustered at the individual level in parentheses. *Data sources:* Administrative data; pooled survey data.

Do These Skills Matter Later in Life?

	Higher Education	STEM field of study	Log hourly wage	Personal income	Household income	Household wealth
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Math					
Math skills	0.144	0.021	0.114	5.453	4.074	5.204
	(0.002)	(0.003)	(0.002)	(0.109)	(0.111)	(0.113)
R-squared	0.238	0.111	0.277	0.317	0.079	0.133
Observations	61,756	28,264	41,928	53,099	55,320	53,963
	Panel B: Language					
Language skills	0.132	0.001	0.102	4.656	3.265	4.133
	(0.002)	(0.003)	(0.002)	(0.109)	(0.112)	(0.114)
R-squared	0.224	0.109	0.262	0.308	0.071	0.120
Observations	61,756	28,264	41,928	53,099	55,230	53,963
	Panel C: Math and language					
Math skills	0.101	0.031	0.084	4.149	3.280	4.219
	(0.002)	(0.003)	(0.002)	(0.134)	(0.138)	(0.139)
Language skills	0.073	-0.017	0.053	2.256	1.369	1.696
	(0.002)	(0.003)	(0.002)	(0.133)	(0.138)	(0.140)
R-squared	0.254	0.112	0.287	0.320	0.081	0.135
Observations	61,756	28,264	41,928	53,099	55,230	53,963

Notes: Least squares regressions. Sample: Pooled sample of all individuals (parents and non-parents) in the three survey cohorts. All wage, income, and wealth variables are measured 30 years after the skill assessment took place; higher education degree completion is based on the highest educational degree obtained by the individual observed in the survey data. Cognitive skills are standardized to SD 1 in full sample in each survey cohort. All regressions control for individual's gender, migration background, number of siblings, age of parents at the time of individual's birth, survey indicators, education and social background of grandparents as well as municipality fixed effects. Standard errors clustered at the individual level in parentheses. *Data sources:* Administrative data; pooled survey data.

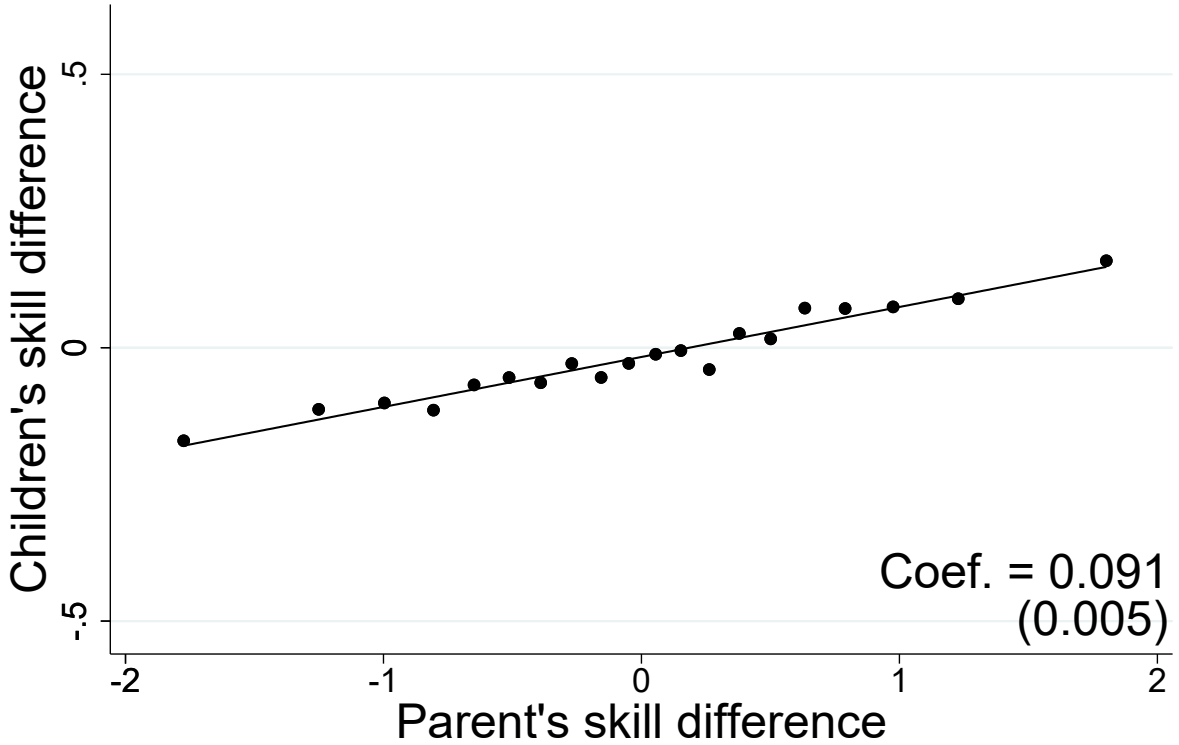
Subject-Specific Skill Transmission



Notes: The figure displays two binned scatterplots showing the strength of parent-child transmissions in math skills (left) and language skills (right). To construct the figure, we divided parent cognitive skills into 20 ranked equal-sized groups and plotted the mean of the children cognitive skills against the mean of the parent skills in each bin. The best-fit line, the coefficient, and the standard error (clustered at the parent level) are calculated from bivariate regressions on the micro data. *Data source:* ITS dataset (linked administrative and pooled survey data).

Subject-Specific OLS Results

Child and Parent Skill Differences



Notes: The figure displays two binned scatterplots showing the strength of parent-child transmissions in math skills (left) and language skills (right). To construct the figure, we divided parent cognitive skills into 20 ranked equal-sized groups and plotted the mean of the children cognitive skills against the mean of the parent skills in each bin. The best-fit line, the coefficient, and the standard error (clustered at the parent level) are calculated from bivariate regressions on the micro data. *Data source:* ITS dataset (linked administrative and pooled survey data).

Between-Subject Results Histograms

Between-Subject Model: Lower Bound Estimates

- We are **not estimating the total effect of families** on child outcomes.
- We **eliminate all other family factors** with a common (causal) impact across subjects (e.g., motivation, access to learning opportunities, general cognitive factors, etc.)
- Between-subject model provides a **lower bound** for the total effect of a more broadly defined set of cognitive factors

Between-Subject IV Model

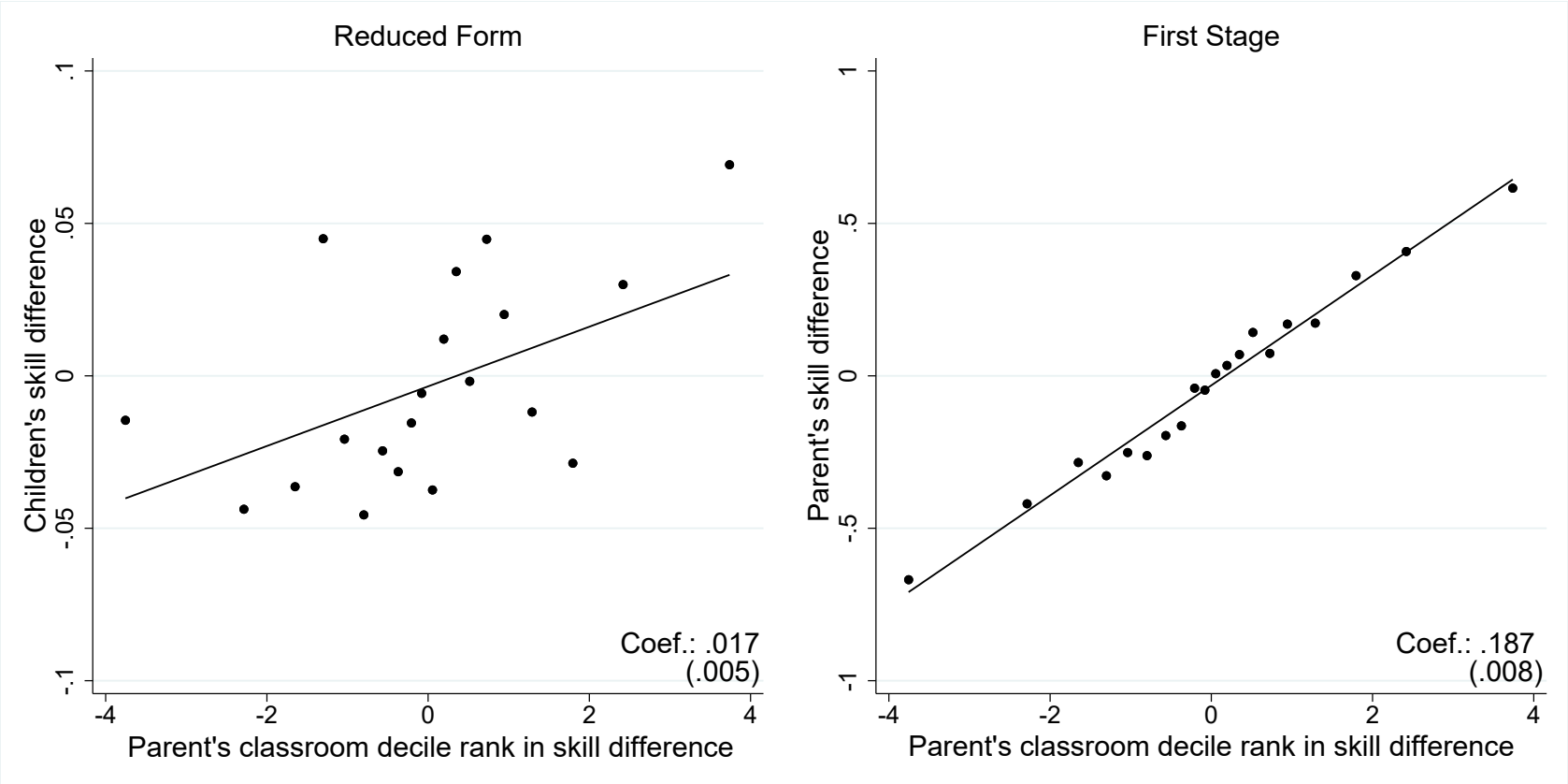
Potential remaining confounders:

- Long-lasting **dynastic predispositions or genetic advantages** in the production of skills in a certain subject within dynasties

Solution:

- We use only variation of within-parent skill differences driven by cross-subject differences in the **quality of formal education environments** of parents
- Exploiting the **class sampling** in the parent survey data, we construct an instrument based on the **difference in class ranks across subjects**

Differences in Parent Classroom Quality



Notes: The figure displays two binned scatterplots showing the strength of the relationship between differences in parent classroom quality and differences in children's cognitive skills (left) and differences in parent cognitive skills (right), respectively. To construct the scatterplots, we divided differences in parent classroom quality into 20 ranked equal-sized groups and plotted the mean of the differences in parent classroom quality against the mean of the differences in children/parent cognitive skills in each bin. The best-fit line, the coefficient, and the standard error (clustered at the classroom level) are calculated from bivariate regressions on the micro data. *Data source:* ITS dataset (linked administrative and pooled survey data).

IV Results

	Between-subject model	First stage IV	Reduced form	Second stage IV	
	(1)	(2)	(3)	(4)	(5)
Parent cognitive skills	0.084			0.095	0.098
	(0.008)			(0.028)	(0.029)
Parent classroom quality		0.187	0.018		
		(0.007)	(0.005)		
Further controls					yes
F-statistic excluded instrument				656.21	614.70
R-squared	0.08	0.32	0.05	0.08	0.08
Observations	24,536	24,536	24,536	24,536	24,536

Notes: Least squares and two-stage least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language. Dependent variables: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year in columns (1), (3), (4), and (5); parent cognitive skills in column (2). Column (1) replicates baseline between-subject model (see column 1 of Table 4) in the IV sample. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Parent classroom quality is measured by the rank of math and language classrooms within the parent's education cohort. Further controls include grandparent education, grandparent social background based on the occupation type of the main breadwinner in the parent household, and municipality fixed effects (all referring to the time when parents took the skill test). All regressions additionally control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

IV Results: Controlling for Children's School Quality

	Between-subject model	First stage IV	Reduced form	Second stage IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Parent cognitive skills	0.083			0.085	0.089	0.088
	(0.008)			(0.028)	(0.028)	(0.028)
Parent classroom quality		0.186	0.016			
		(0.007)	(0.005)			
Children's school quality (ranks)	0.053	0.005	0.053	0.053	0.052	
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	
Children's school quality (absolute)						0.489
						(0.036)
Further controls					yes	yes
F-statistic excluded instrument				658.66	612.95	612.56
R-squared	0.11	0.32	0.08	0.11	0.12	0.12
Observations	24,482	24,482	24,482	24,482	24,482	24,482

Notes: Least squares and two-stage least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language; children with missing school information are excluded. Dependent variables: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year in columns (1), (3), (4), (5), and (6); parent cognitive skills in column (2). Column (1) replicates baseline between-subject model (see column 1 of Table 4) in the IV sample. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Parent classroom quality is measured by the rank of math and language classrooms within the parent's education cohort. Children's school quality is measured by the rank in math and language, respectively, of the school attended by the child (leave-out mean) in the national test score distribution in a given year. Percentile ranks are divided by 10 to ease exposition. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

One-classroom schools

Robustness Checks

- Dropping one parent in two-parent sample Results
- Between-subject model: Estimates for each cohort Results
- Between-subject model: Ranks Results
- IV: Different definitions of classroom quality Results
- IV: Addressing correlated intergenerational peer composition Results
- IV: Addressing between- or within-school sorting of parents Results

Other Findings

Effect heterogeneity: Results

- Effect does not vary by the gender match of parents and children
- Effect tends to be lower for children with low-educated grandparents
- No effect heterogeneity with respect to grandparents' social background

Potential mechanisms: Results

- Role of economic factors in explaining the skill transmission is very limited

Long-term outcomes: Results

- Pattern of parent skills directly relates to STEM choices of children

Conclusions

- **Evidence on cognitive skill transmission within the family**
  - Dutch assessment data on cognitive skills of parents linked with register data on their children's skills in the same subjects (math and language) elicited by similar tests at about the same age
- **Strong causal impact of parent skills on the skills of children**
  - One SD increase in parent cognitive skills leads to 0.1 SD increase in child cognitive skills
- **Family cognitive skills also influence long-run career patterns**
  - High math skills of parents promote greater choice of STEM paths by children
- **Policy implication**
  - Improving education quality today entails spill-over effects to future generations

Thank you for your attention!


Rick Hanushek


Babs Jacobs


Guido Schwerdt


Rolf van der Velden


Stan Vermeulen


Simon Wiederhold

guido.schwerdt@uni.kn

website

Empirical Strategy: Conceptual Model

$$T^c_{ida} = \alpha + \beta_a F_{ida} + \gamma_a S_{ida} + \eta_{ida}$$

T^c_{ida}	test score of child *i* of dynasty *d* in subject *a*
F_{ida}	family factors
S_{ida}	environmental factors (or simply school factors)
η_{ida}	contains all unobservable influences on child test scores

Empirical Strategy: Conceptual Model II

$$F_{ida} = T^{p}_{ida} + \psi_{ida} + \xi_{ida}$$

T^{p}_{ida}	test score of parent of child *i* of dynasty *d* in subject *a*
ψ_{ida}	pre-birth factors for members of dynasty *d*
ξ_{ida}	post-birth factors for members of dynasty *d*

Subject-Specific Models

$$T^c_{idm} = \alpha_m + \beta_m T^p_{idm} + \gamma_m S_{idm} + \beta_m \psi_{idm} + \beta_m \xi_{idm} + \eta_{idm}$$

$$T^c_{idl} = \alpha_l + \beta_l T^p_{idl} + \gamma_l S_{idl} + \beta_l \psi_{idl} + \beta_l \xi_{idl} + \eta_{idl}$$

$\rightarrow$ We can combine these equations by assuming $\beta_a = \beta$ and $\gamma_a = \gamma$

Between-Subject Model

Simplified version:

$$\Delta T^c_{id} = \alpha + \beta \Delta T^p_{id} + \epsilon_{id}$$

where $\epsilon_{id} = \gamma \Delta S_{id} + \beta \Delta \psi_{id} + \beta \Delta \xi_{id} + (\eta_{idm} - \eta_{idl})$

Remaining concern: Are ϵ_{id} and ΔT^p_{id} correlated?

Descriptive Statistics I

Variables		Pooled	Cohort 1977	1982	1989
		(1)	(2)	(3)	(4)
Child characteristics					
Math skills	Mean	0.048	0.120	0.009	-0.123
	SD	0.97	0.95	0.97	0.99
Language skills	Mean	0.066	0.149	0.019	-0.132
	SD	0.95	0.92	0.96	1.00
Math-language skill difference	Mean	-0.017	-0.029	-0.010	0.009
	SD	0.78	0.78	0.77	0.78
Course profile	STEM	0.34	0.37	0.34	0.24
	Non-STEM	0.46	0.46	0.46	0.44
Field of study	STEM	0.23	0.28	0.20	0.11
	Non-STEM	0.48	0.57	0.43	0.27
Gender	Female	0.50	0.50	0.51	0.51
Parent characteristics					
Math skills	Mean	0.100	0.220	0.027	-0.173
	SD	0.98	0.97	0.99	0.97
Language skills	Mean	0.105	0.202	0.036	-0.092
	SD	0.96	0.93	0.99	0.98
Math-language skill difference	Mean	-0.006	0.018	-0.009	-0.081
	SD	0.86	0.82	0.93	0.84
Personal income percentile	Mean	58.08	63.29	56.05	43.42
	SD	25.86	24.82	25.33	24.51
Household income percentile	Mean	72.50	74.38	72.18	66.54
	SD	21.84	21.54	21.64	22.18
Household wealth percentile	Mean	63.29	66.36	61.67	55.72
	SD	28.84	28.77	28.65	27.79
Gender	Female	0.53	0.48	0.57	0.63
Education	Low	0.24	0.21	0.25	0.30
	Medium	0.44	0.48	0.41	0.40
	High	0.25	0.28	0.25	0.17
Migration background	Yes	0.08	0.07	0.08	0.15
Number of siblings	0 siblings	0.06	0.06	0.05	0.05
	1 sibling	0.37	0,34	0.41	0.40
	2 siblings	0.28	0.30	0.26	0.23
	>2 siblings	0.23	0.25	0.21	0.19

Back

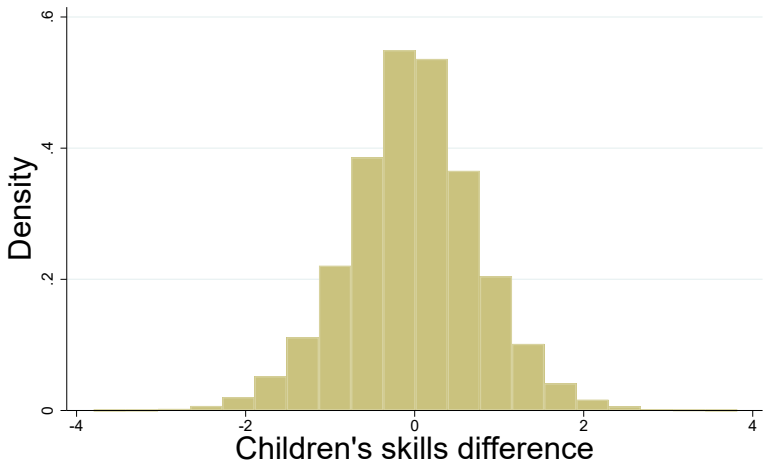
Descriptive Statistics II

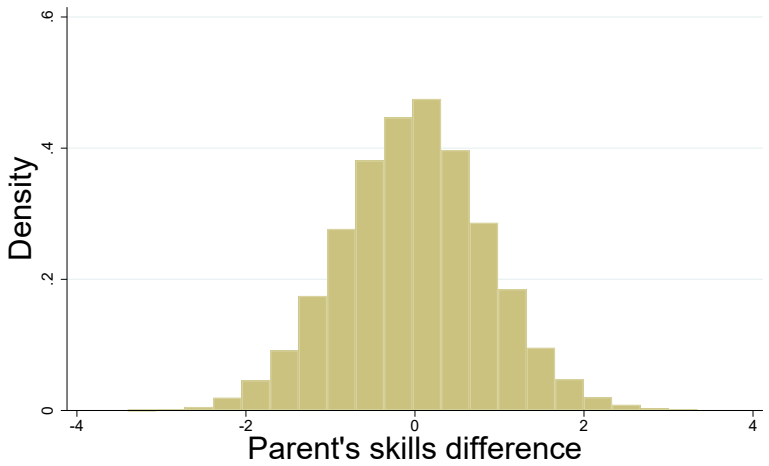
Grandparent Characteristics					
Education	Primary Education	0.19	0.14	0.26	0.20
	Lower secondary education	0.31	0.30	0.34	0.27
	Higher secondary education	0.29	0.33	0.19	0.34
	Tertiary education	0.17	0.19	0.14	0.14
Social background	Blue collar worker	0.28	0.28	0.29	0.28
	Employer - without staff	0.08	0.09	0.07	0.05
	Employer - with staff	0.05	0.06	0.05	0.04
	Lower white collar worker	0.11	0.12	0.11	0.09
	Middle white collar worker	0.19	0.21	0.16	0.17
	Professionals	0.12	0.13	0.11	0.12
	Other	0.13	0.12	0.13	0.21
Age at time of birth grandfather	Mean	30.57	31.47	29.76	29.06
Age at time of birth grandmother	Mean	27.99	28.81	27.36	26.42
Observations	Total number	41,774	22,417	12,930	6,427

Notes: Table reports means, SD, and shares for the variables indicated in column (1) for the pooled sample as well as the three education cohorts separately. The type of statistic reported is indicated in column (2). If neither Mean, SD, or Total number is specified, the reported statistic refers to the share with in the sample indicated in the top row. Children's cognitive skills are standardized with mean zero and SD one in the full sample of children taking the test in their cohort based on administrative data. Parent cognitive skills standardized with mean zero and SD one in the full sample of participants from each education cohort. Children's gender, course profile, and field of study are taken from administrative data. Students are designated as following a STEM course profile if they take the Technical or Agriculture profile (low academic track) or the Nature & Technical or Nature & Health profile (middle/high academic track). STEM study choice is determined based on the 1-digit ISCED97 fields of education classification (UNESCO, 2003). Study programs in the Science, Mathematics and Computing, Engineering, Manufacturing and Construction, Agriculture, and Medicine and Nursery were classified as a STEM choice of study. Students who chose a 'combination' course profile, where its' STEM-component is unknown, have been coded as non-STEM. Not all students can be assigned a STEM/non-STEM course profile/field of study as they have not progressed far enough into the education system. Household income is based on the percentile of the household in the Dutch distribution in terms of yearly spendable income. Parent personal income is based on the percentile of the parent in the Dutch income distribution (sources include: labor income, owned companies, unemployment benefits and social security). Household wealth is percentile of the household in the Dutch distribution in terms of the household's total wealth. Income and wealth data are taken from the administrative data in the child's test-taking year. Parent education is measured as the highest educational degree obtained by the parent observed in the survey data. In parent education, "low" denotes maximum lower secondary education (ISCED 1 or 2); "medium" denotes higher secondary or upper secondary vocational education (ISCED 3 or 4); "high" denotes tertiary education, consisting of higher vocational education and university (ISCED 5 and above). Grandparent education is the highest level of education of both grandparents. Social background is based on the occupation type of the main breadwinner in the parent household at the time of the parents' skill assessment. For expositional reasons, mean age of grandparents at the time of the parent's birth is shown in the table; in the regressions, we control for the following age groups: below 21, 21-25, 26-30, 31-35, 36-40, 41 and above. Apart from income and wealth, which are taken from administrative data, all (grand-)parent characteristics stem from the survey datasets. (Grand-)parent characteristics are reported at the child level. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Histogram of Math-Language Skill Differences





Notes: The figure depicts the difference between math and language skills for children (left) and parents (right). *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Coefficients on Control Variables in Between-Subject Model

Variables	(1)	Variables	(2)
Parent cognitive skills	0.096	No Answer	0.003
	(0.005)		(0.037)
Language	-0.168	*Grandparent characteristics*	
	(0.225)	Age grandfather at time of parent birth: 21-25	-0.033
Parent characteristics			(0.042)
Gender	-0.034	Age grandfather at time of parent birth: 26-30	-0.017
	(0.009)		(0.042)
Migrant	0.009	Age grandfather at time of parent birth: 31-35	-0.022
	(0.016)		(0.043)
Number of siblings: 1	0.000	Age grandfather at time of parent birth:36-40	-0.017
	(0.018)		(0.045)
Number of siblings: 2	0.010	Age grandfather at time of parent birth: 41 and above	-0.012
	(0.019)		(0.048)
Number of siblings: 3 or more	-0.031	Age grandfather at time of parent birth: missing	0.004
	(0.019)		(0.084)
Number of siblings: missing	0.032	Age grandmother at time of parent birth: 21–25	0.019
	(0.031)		(0.022)
Grandparent education		Age grandmother at time of parent birth: 26-30	0.022
Grandparent education: lower secondary	0.016		(0.024)
	(0.013)	Age grandmother at time of parent birth:21-25	0.019
Grandparent education: upper secondary	0.018		(0.022)
	(0.014)	Age grandmother at time of parent birth: 26-30	0.022
Grandparent education: tertiary	0.045		(0.024)
	(0.017)	Age grandmother at time of parent birth: 21-25	0.019
Grandparent education: missing	0.034		(0.022)
	(0.034)	Age grandmother at time of parent birth: 26-30	0.022
Grandparent social background			(0.024)
Blue-collar worker	0.061	Age grandmother at time of parent birth: 31-35	0.049
	(0.018)		(0.027)
Employer	0.047	Age grandmother at time of parent birth: 36-40	0.018
	(0.024)		(0.031)
Low white-collar worker	0.083	Age grandmother at time of parent birth: 41 and above	0.032
	(0.020)		(0.043)
Medium white-collar worker	0.081	Age grandmother at time of parent birth: missing	-0.366
	(0.019)		(0.245)
Professionals	0.072		
	(0.021)		
Other	0.060		
	(0.021)		
Municipality fixed effects		yes	
R-squared	0.067	Observations	83,548

Back

Subject-Specific OLS Results

	Panel A: Math			
	(1)	(2)	(3)	(4)
Parent cognitive skills	0.273	0.257	0.255	0.258
	(0.005)	(0.005)	(0.005)	(0.005)
R-squared	0.090	0.094	0.096	0.120
Observations (students)	41,774	41,774	41,774	41,774
	Panel B: Language			
Parent cognitive skills	0.286	0.264	0.261	0.261
	(0.005)	(0.005)	(0.005)	(0.005)
R-squared	0.102	0.109	0.110	0.135
Observations (students)	41,774	41,774	41,774	41,774
	Control variables in Panels A+B			
Grandparent education		yes	yes	yes
Grandparent social background			yes	yes
Municipality fixed effects				yes

Notes: Least squares regressions. Sample: Pooled sample of all matched parent-children observations in the three education cohorts. Dependent variables: Math skills of children in Panel A; language skills of children in Panel B; Children's skills in math and language standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Between-Subject Results

	(1)	(2)	(3)	(4)
Parent cognitive skills	0.098	0.098	0.097	0.096
	(0.005)	(0.005)	(0.005)	(0.005)
Grandparent education		yes	yes	yes
Grandparent social background			yes	yes
Municipality fixed effects				yes
R-squared	0.088	0.090	0.089	0.067
Observations	83,548	83,548	83,548	83,548

Notes: Least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the three education cohorts pooled over math and language. Dependent variable: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Coefficients control variables Back

Effect Heterogeneity

	(1)	(2)	(3)
Parent cognitive skills	0.099	0.086	0.099
	(0.010)	(0.070)	0.009)
Parent-child gender interaction			
x Male parent &female child	-0.002		
	(0.013)		
x Female parent & male child	-0.003		
	(0.013)		
x Female parent & female child	-0.003		
	(0.013)		
Grandparent education			
x Medium		0.026	
		(0.011)	
x High		0.021	
		(0.014)	
x Missing education information		-0.011	
		(0.024)	
Parent social background			
x Employer			-0.001
			(0.016)
x Low white-collar worker			-0.005
			(0.017)
x Medium white-collar worker			0.016
			(0.015)
x Professionals			-0.015
			(0.016)
x Other			-0.018
			(0.015)
Grandparent education	yes	yes	yes
Grandparent social background	yes	yes	yes
Municipality fixed effects	yes	yes	yes
R-squared	0.051	0.067	0.067
Observations	83,548	83,548	83,548

Notes: Least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the three education cohorts pooled over math and language. Dependent variable: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. The coarser definition of grandparent education used in this table combines primary and lower secondary education to the lower education category, while upper secondary and tertiary education are referred to as medium and tertiary education, respectively. The coarser definition of parent social status lumps together "employer without staff" and "employer with staff" in the "employer" category, and the "other" and "unknown" in the "other" category. Omitted category in column (1) is male parent & male child; omitted category in column (2) is low education (at most lower secondary); omitted category in column (3) is blue collar worker. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions further control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Effect Heterogeneity (detailed)

	(1)	(2)	(3)
Parent cognitive skills	0.099	0.061	0.089
	(0.010)	(0.012)	0.018)
Parent-child gender interaction			
x Male parent & female child	-0.002		
	(0.013)		
x Female parent & male child	-0.003		
	(0.013)		
x Female parent & female child	-0.003		
	(0.013)		
Grandparent education			
x Lower secondary		0.041	
		(0.015)	
x Upper secondary		0.052	
		(0.015)	
x Tertiary		0.047	
		(0.017)	
x Missing education information		0.015	
		(0.026)	
Grandparent social background			
x Blue collar worker			-0.010
			(0.020)
x Employer			0.021
			(0.027)
x Low white-collar worker			0.005
			(0.023)
x Medium white-collar worker			0.026
			(0.021)
x Professionals			-0.004
			(0.023)
x Other			-0.011
			(0.023)
x No answer			0.001
			(0.033)
Observations	83,548	83,548	83,548

Back

Analysis of Mechanisms (Between-Subject Model)

	(1)	(2)	(3)	(4)	(5)
Parent cognitive skills	0.096	0.097	0.096	0.096	0.095
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Parent education					
Medium		-0.002			
		(0.011)			
High		0.030			
		(0.013)			
Missing		-0.012			
		(0.019)			
Parent income (/10)			-0.001		
			(0.002)		
Household income (/10)				-0.007	
				(0.002)	
Household wealth (/10)					-0.009
					(0.002)
Grandparent education	yes	yes	yes	yes	yes
Grandparent social background	yes	yes	yes	yes	yes
Municipality fixed effects	yes	yes	yes	yes	yes
R-squared	0.067	0.070	0.067	0.064	0.062
Observations	83,548	83,548	83,548	83,548	83,548

Notes: Least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the three education cohorts pooled over math and language. Dependent variable: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Parent education is measured as the highest educational degree obtained by the observed parent (omitted category: low education); low education: at most lower secondary; medium education: higher secondary and upper secondary vocational education; high education: tertiary education, consisting of higher vocational education and university. Household income is based on the percentile of the household in the Dutch household distribution in terms of yearly spendable income in the child's test-taking year. Parent personal income is based on the percentile of the parent in the Dutch personal income distribution (including income from labor, income from owned companies, unemployment and social security) in the child's test-taking year. Household wealth is based on the percentile of the household in the Dutch household distribution in terms of the household's total wealth, determined by assets minus debts in the child's test-taking year. Missing values for parent education (3.5%), parent income (6.7%), household income (1.5%), and household wealth (11.5%) are imputed (imputation dummies added to the regression models). Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions further control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

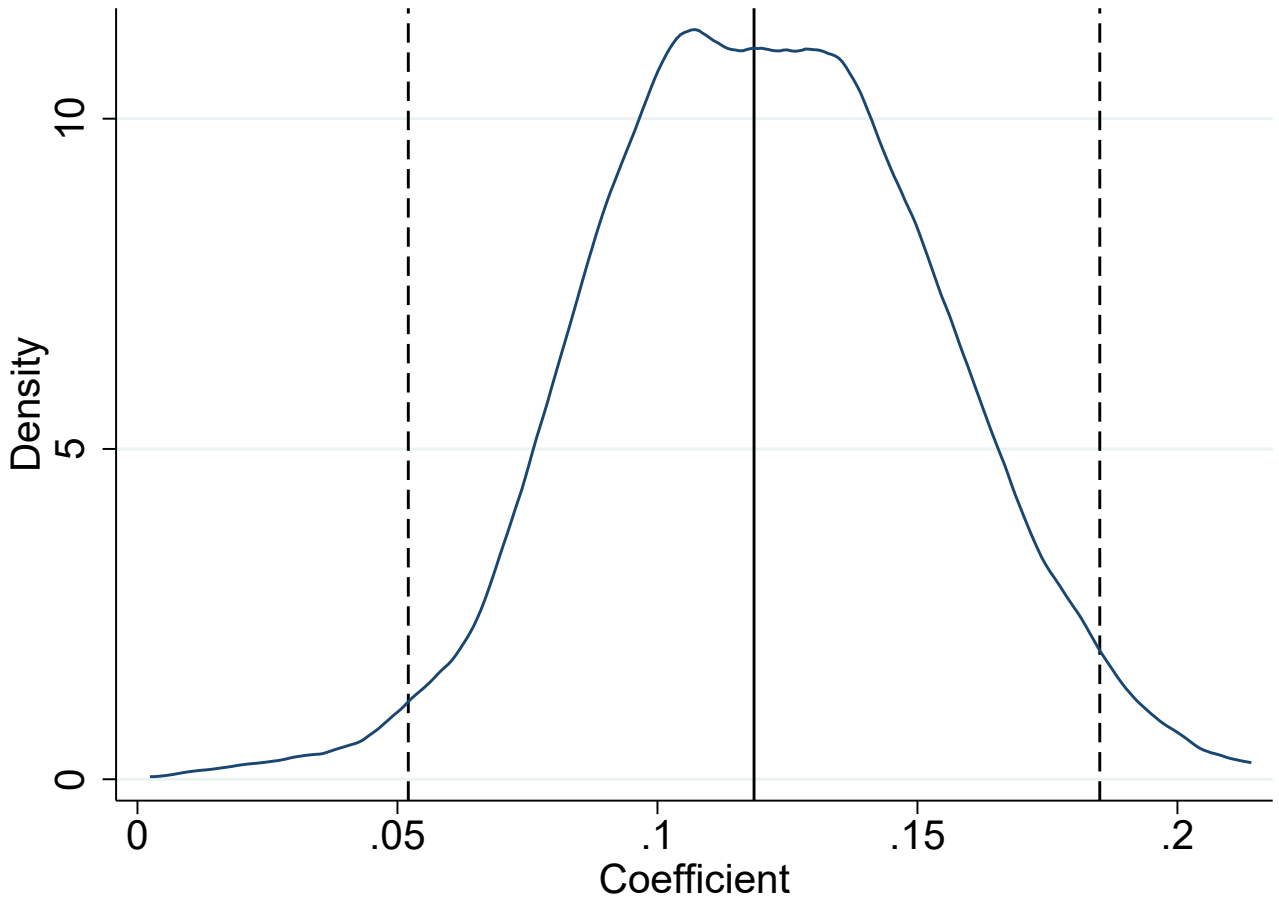
Relationship Between Parent Skills and Potential Mediators

	Parent education (1)	Parent income (2)	Household income (3)	Household wealth (4)
Parent skill difference	0.017	0.640	0.471	0.873
	(0.002)	(0.139)	(0.128)	(0.152)
Grandparent education	yes	yes	yes	yes
Grandparent social background	yes	yes	yes	yes
Municipality fixed effects	yes	yes	yes	yes
R-squared	0.164	0.438	0.122	0.203
Observations	40,311	38,957	41,134	36,973

Notes: Least squares regressions. Sample: Pooled sample of all matched parent-children observations in the three education cohorts. Dependent variables: Binary variable taking a value of 1 if parents obtained a degree in higher vocational education or university education; 0 otherwise (column 1). Parent income including income from labor, income from owned companies, unemployment and social security, measured as the percentile of the parent in the Dutch personal income distribution in the child's test-taking year (column 2). Sum of the personal incomes of all household members measured as the percentile of the household in the Dutch household distribution in terms of yearly spendable income in the child's test-taking year (column 3). Household wealth, measured as the percentile of the household in the Dutch household distribution in terms of the household's total wealth, determined by assets minus debts in the child's test-taking year (column 4). Parent skill difference is math – language; parent cognitive skills are standardized to SD 1 in full sample of parents in each education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions further control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Randomly Dropping One Parent in Two-Parent Sample



Notes: The figure depicts estimated coefficients on parent cognitive skills in the between-subject model when redrawing samples 1,200 times. Estimations are conducted based on 365 children for whom we observe both parents in the survey data. In each of the 1,200 iterations we randomly drop one of the parents for each child and estimate the relationship between child and parent cognitive skills. Solid vertical line indicates coefficient in the two-parent estimation, dashed lines indicate 95 percent confidence interval. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

OLS and Between-Subject Estimates by Cohort

	Panel A:Math			
	Pooled	Cohort		
		1977	1982	1989
Parent cognitive skills	0.258	0.261	0.255	0.241
	(0.005)	(0.007)	(0.010)	(0.015)
R-squared	0.120	0.130	0.134	0.140
Observations (stundents)	41,774	22,417	12,930	6,427
	Panel B: Language			
Parent cognitive skills	0.261	0.284	0.223	0.246
	(0.005)	(0.007)	(0.009)	(0.015)
R-squared	0.135	0.149	0.141	0.159
Observations (students)	41,774	22,417	12,930	6,427
	Panel C: Math and language (between-subject model)			
Parent cognitive skills	0.096	0.123	0.069	0.085
	(0.005)	(0.008)	(0.009)	(0.013)
R-squared	0.067	0.066	0.045	0.057
Observations	83,548	44,834	25,860	12,854
	Control variables in all panels			
Grandparent education	yes	yes	yes	yes
Parent social background	yes	yes	yes	yes
Municipality fixed effects	yes	yes	yes	yes

Back

Between-Subject Model in Ranks

	(1)	(2)	(3)	(4)
Parent skill rank	0.098	0.097	0.096	0.094
	(0.005)	(0.005)	(0.005)	(0.005)
Grandparent education		yes	yes	yes
Grandparent social background			yes	yes
Municipality fixed effects				yes
R-squared	0.088	0.091	0.090	0.067
Observations	83,548	83,548	83,548	83,548

Notes: Least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the three education cohorts pooled over math and language. Dependent variable: Percentile rank of test score of children in full sample of children taking the test. Parent skill rank is measured as the percentile rank of test score of parents in full sample of parents in a education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions control for parent survey indicators and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

IV: Different Definitions of Classroom Quality

	Rank class (main)	Rank class leave-out	Rank class dummy leave-out	Level class dummy leave-out	Level class leave-out	Rank class track-specific
	(1)	(2)	(3)	(4)	(5)	(6)
Parent cognitive skills	0.098	0.112	0.109	0.104	0.087	0.108
	(0.029)	(0.047)	(0.057)	(0.051)	(0.044)	(0.031)
Further controls	yes	yes	yes	yes	yes	yes
F-statistic excluded instrument	614.70	208.94	92.86	129.47	233.15	504.65
R-squared	0.083	0.084	0.084	0.084	0.082	0.084
Observations	24,536	24,536	24,536	24,536	24,536	24,536

Notes: Two-stage least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language. Dependent variables: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year; parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Instruments: Column (1): Rank of math and language classrooms within the parent's education cohort; column (2): Rank of math and language classroom peers within the parent's education cohort; column (3): Binary indicator for higher ranked classroom peers (math vs. language) within the parent's education cohort; column (4): Binary indicator for better performing classroom peers (math vs. language); column (5): Test scores in math and language of classroom peers; column (6): Like column (1), but rank of math and language classrooms in the 1989 cohort (where children were sampled in the first year of secondary school) calculated by track, distinguishing between 11 different tracks. Further controls include grandparent education, grandparent social background, and municipality fixed effects (all referring to the time when parents took the skill test). All regressions additionally control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

Back

IV: Controlling for Children's School Quality (One-Classroom Schools)

	FE model	First stage IV	Reduced form	Second stage IV		
	(1)	(2)	(3)	(4)	(5)	(6)
Parent cognitive skills	0.074			0.071	0.082	0.080
	(0.012)			(0.042)	(0.045)	(0.045)
Parent classroom quality		0.177	0.013			
		(0.010)	(0.007)			
Children's school quality	0.044	0.005	0.044	0.044	0.042	
(ranks)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	
Children's school quality						0.399
(absolute)						(0.044)
Further controls					yes	yes
F-statistic excluded instrument				315.72	279.97	280.16
R-squared	0.11	0.32	0.08	0.11	0.12	0.13
Observations	11,240	11,240	11,240	11,240	11,240	11,240

Notes: Table replicates Table 6 for children whom we observe in a school with at most 30 grade-six students in a given year; this is our proxy for one-classroom schools, as classroom identifiers are not available in the administrative CITO data. Least squares and two-stage least squares regressions with family fixed effects. Sample: Sample of all matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language in school-year combinations with 30 or less total observations; children with missing school information are excluded. Dependent variables: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year in columns (1), (3), (4), (5), and (6); parent cognitive skills in column (2). Column (1) replicates baseline between-subject model (see column 1 of Table 4) in the IV sample. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Parent classroom quality is measured by the rank of math and language classrooms within the parent's education cohort. Children's school quality is measured by the rank in math and language, respectively, of the school attended by the child (leave-out mean) in the national test score distribution in a given year. Percentile ranks are divided by 10 to ease exposition. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

Back

IV: Regional Movers

	W/o children of parent's classmates	Child & parent school not in same municipality	Child & parent school not in same municipality (distance >50 km)	Child & parent school not in same municipality (distance >100 km)	Child & parent school not in same province
	(1)	(2)	(3)	(4)	(5)
Parent cognitive skills	0.087	0.080	0.157	0.193	0.098
	(0.031)	(0.037)	(0.064)	(0.099)	(0.059)
Further controls	yes	yes	yes	yes	yes
F-statistic excl. instr.	541.73	339.94	66.65	43.90	103.75
R-squared	0.080	0.080	0.100	0.090	0.090
Observations	21,932	12,828	2,720	1,170	4,622

Notes: Two-stage least squares regressions with family fixed effects in the sample of matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language. Samples: Column (1): Excluding children who attend the same school and whose parents have been classmates in the education cohorts of 1982 and 1989; column (2): as in column (1), while keeping only children whose school is located in a different municipality than the parent's school in the education cohorts of 1982 and 1989; column (3) (column 4): as in column (2), while keeping only children whose school is located in a municipality that is more than 50 km (100 km) away from the municipality of the parent's school in the education cohorts of 1982 and 1989 (using the municipality centroid); column (5): as in column (1), while keeping only children whose school is located in a different province than the parent's school in the education cohorts of 1982 and 1989. Results in columns (2) and (5) contain only children with a valid municipality or province identifier (92.06 percent of the total IV sample). Results in columns (3) and (4) contain only children and parents with available municipality longitude and latitude coordinates (88.52 percent of the total IV sample). Dependent variables: children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Parent classroom quality is measured by the rank of math and language classrooms within the parent's education cohort. Further controls include grandparent education and grandparent social background (all referring to the time when parents took the skill test). All regressions additionally control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

Back

IV: School Sorting in the Parent Generation

	Main	Rural schools	One-classroom schools	Rural & one-classroom schools	Cohort 1982	Cohort 1989
	(1)	(2)	(3)	(4)	(5)	(6)
Parent cognitive skills	0.098	0.101	0.122	0.113	0.110	0.085
	(0.029)	(0.036)	(0.038)	(0.045)	(0.038)	(0.046)
Further controls	yes	yes	yes	yes	yes	yes
F-statistic excluded instrument	614.70	347.44	462.53	310.46	448.60	248.39
R-squared	0.083	0.070	0.074	0.063	0.072	0.077
Observations	24,536	11,050	13,296	7,340	11,682	12,854

Notes: Two-stage least squares regressions with family fixed effects. Samples: Column (1): Sample of all matched parent-children observations in the education cohorts of 1982 and 1989 pooled over math and language; column (2): Sample of matched parent-children observations form rural schools in the education cohorts of 1982 and 1989 pooled over math and language; column (3): Sample of matched parent-children observations from schools with exactly one classroom in the education cohorts of 1982 and 1989 pooled over math and language; column (4): Sample of matched parent-children observations from rural schools with exactly one classroom in the education cohorts of 1982 and 1989 pooled over math and language; column (5): Sample of all matched parent-children observations in the education cohort of 1982 pooled over math and language; column (6): Sample of all matched parent-children observations in the education cohort of 1989 pooled over math and language. Dependent variable: Children's cognitive skills standardized with mean zero and SD one in full sample of children taking the test in each test year. Parent cognitive skills standardized with mean zero and SD one in full sample of parents in each education cohort. Classroom quality is measured by the rank of math and language classrooms within the parent's education cohort. Further controls include grandparent education, grandparent social background, and municipality fixed effects (all referring to the time when parents took the skill test). All regressions further control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. All control variables are interacted with a subject indicator. Standard errors clustered at the classroom level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data), 1982 and 1989 cohort.

Back

Long-Term Outcomes – STEM Choice

	STEM profile choice (y/n)		STEM study choice (y/n)	
	(1)	(2)	(3)	(4)
Parent skill difference	0.027	0.025	0.011	0.017
	(0.003)	(0.005)	(0.003)	(0.005)
x child female		0.004		-0.009
		(0.006)		(0.006)
Grandparent education	yes	yes	yes	yes
Grandparent social background	yes	yes	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Baseline outcome all	0.425		0.323	
Baseline outcome female	0.359		0.217	
Baseline outcome male	0.494		0.428	
R-squared	0.046	0.064	0.040	0.088
Observations	33,414	33,414	29,686	29,686

Notes: Least squares regressions. Sample: Pooled sample of all matched parent-children observations in the three education cohorts. Dependent variables: Binary variable indicating the choice of a STEM (Science, Technology, Engineering, and Mathematics) course profile at secondary school in columns (1) and (2); binary variable indicating the choice of a STEM field of study after secondary school in columns (3) and (4). Students are designated as following a STEM-course profile if they take the Technical or Agriculture course profile (low academic track) or the Nature/Technical or Nature/Health course profile (middle or high academic track). STEM study choice is determined based on the 1-digit ISCED97 fields of education classification (UNESCO, 2003), where study programs categorized as Science, Mathematics and Computing, Engineering, Manufacturing and Construction, Agriculture, as well as Medicine and Nursery were classified as a STEM choice of study (see Section 3 for details). Baseline values are calculated based on observations with non-missing information on STEM choices. Parent skill difference is math–language; parent cognitive skills are standardized with mean zero and SD one in full sample of parents in each education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions further control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).

Back

Long-Term Outcomes – STEM Choice (Narrow Definition of STEM)

	STEM profile choice (y/n)		STEM study choice (y/n)	
	(1)	(2)	(3)	(4)
Parent skill difference	0.021	0.027	0.009	0.018
	(0.003)	(0.004)	(0.003)	(0.005)
x child female		-0.011		-0.017
		(0.005)		(0.005)
Grandparent education	yes	yes	yes	yes
Grandparent social background	yes	yes	yes	yes
Municipality fixed effects	yes	yes	yes	yes
Baseline outcome all	0.236		0.217	
Baseline outcome female	0.125		0.076	
Baseline outcome male	0.350		0.360	
R-squared	0.040	0.107	0.038	0.152
Observations	33,414	33,414	29,686	29,686

Notes: Least squares regressions. Sample: Pooled sample of all matched parent-children observations in the three education cohorts. Dependent variables: Binary variable indicating the choice of a STEM (Science, Technology, Engineering, and Mathematics) course profile at secondary school in columns (1) and (2); binary variable indicating the choice of a STEM field of study after secondary school in columns (3) and (4). Students are designated as following a STEM-course profile if they take the Technical course profile (low academic track) or the Nature/Technical profile (middle/high academic track). STEM study choice is determined based on the 1-digit ISCED97 fields of education classification (UNESCO, 2003), where study programs categorized as Science, Mathematics and Computing, and Engineering, Manufacturing and Construction were classified as a STEM choice of study (see Section 3 for details). Parent skill difference is math – language; parent cognitive skills are standardized to SD 1 in full sample of parents in each education cohort. Grandparent education is measured by four categories of the highest level of education of both grandparents. Grandparent social background is measured by seven categories of occupational status of the main breadwinner in the parent household. Grandparent education, grandparent social background, and municipality fixed effects refer to time when parents took the skill test. All regressions control for parent gender, parent migration background, number of siblings of parents, age of grandparents at the time of parent birth, parent survey indicators, and children test year fixed effects. Standard errors clustered at the parent level in parentheses. *Data source:* ITS dataset (linked administrative and pooled survey data).