Appendix

A Data Appendix

A.1 Coverage of the IFLS and INPRES program

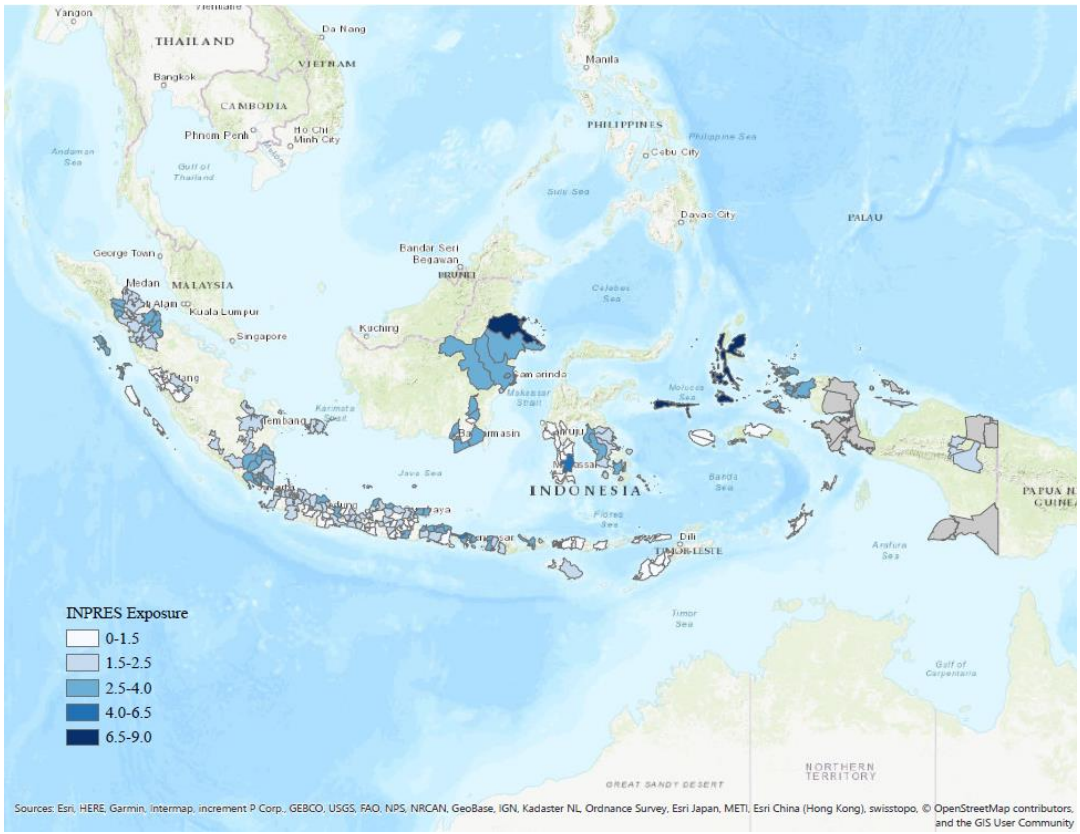
Program intensity

We compare the intensity of the INPRES school construction project in the IFLS and IFLS-E against the national record. The IFLS provinces include 13 out of Indonesia's 26 provinces in 1993. They include: North Sumatra, West Sumatra, South Sumatra, Lampung, Jakarta, West Java, Central Java, Yogyakarta, East Java, Bali, West Nusa Tenggara, South Kalimantan, and South Sulawesi. The IFLS-E provinces include the following 7 provinces in 2012: East Nusa Tenggara, East Kalimantan, Southeast Sulawesi, Maluku, North Maluku, West Papua, and Papua. The IFLS and IFLS-E include almost 300 of Indonesia's 519 districts. Figure A.1 shows the intensity of the INRES program at the national level while Figure A.2 shows the intensity of the INPRES program in the IFLS and IFLS-E districts. A comparison of Figures A.1 and A.2 shows that the IFLS and IFLS-E include both high and low intensity program districts.

Figure A.1: INPRES exposure - All Indonesia

Source: Authors' calculations based on Duflo (2001)

Figure A.2: INPRES exposure in the IFLS and IFLS-E districts



Source: Authors' calculations based on the IFLS, IFLS-E, and Duflo (2001)

A.2 Data construction

Indonesia is administratively divided into provinces, districts (regencies or cities), sub-districts, and villages in rural areas or townships in urban areas. The IFLS over sampled urban and rural areas outside of the main island of Java. IFLS-1 included 7,224 households residing in 13 of Indonesia's 26 provinces in 1993. These households resided in approximately 200 districts, which corresponded to 321 enumeration areas in 312 communities. A community is defined as a village in rural areas and a township in urban areas. The IFLS-E includes 2,500 households residing in seven provinces in eastern Indonesia, which corresponded to about 50 districts and 99 communities. Households in the main IFLS and IFLS-E resided in almost 300 of Indonesia's 514 districts.

Date and district of birth

To obtain the sample of first generation individuals, we begin by identifying individuals who were born between 1950 and 1972 in the IFLS and IFLS-E. In each wave, the IFLS household roster includes information on date of birth (month and year). Also, the IFLS asks respondents over the age of 15 their place of birth in the wave in which they first join the survey. Indonesia experienced district proliferation over time, so we match each district to the 1993 district code in IFLS1. INPRES school construction in the district, water and sanitation program, enrollment in 1971, number of school-aged children in 1971: We obtain these variables from Duflo (2001).

Linking the first and second generation

To identify the second generation, who are the children of the first generation individuals, we use the household relationship in the household roster and women's birth history, matched to the household roster. In each wave, the survey includes an individual's relationship to the head of the household, and an identifier for an individual's mother and father if the mother and father are in the same household. The IFLS also includes a woman's birth history, which allows us to match mothers to their children, and subsequently to children's outcomes.

Long-term outcomes for the first generation

Good self-reported health takes the value one if a respondent reported his or her self-status as "very healthy" or "healthy". The literature in epidemiology has established that self-reported health status is a valid and comprehensive health measure that is highly predictive of well-known health markers such as mortality in both high and lower income countries, even after controlling for socio-demographic factors (Idler and Benyamini, 1997; DeSalvo et al., 2005; Razzaque, Mustafa, and Streatfield, 2014). As additional adult health outcomes, we include the number of days a respondent missed his or her activities in the past 4 weeks prior to the survey. Respondents were also asked to report diagnosed chronic conditions, and we use an indicator for any condition as well as the number of conditions. These conditions include: hypertension, diabetes, tuberculosis, asthma, other respiratory conditions, stroke, heart disease, liver condition, cancer, arthritis, high cholesterol, depression/psychiatric condition.

To assess mental health, respondents were administered a series of 10 questions on how frequently they experienced symptoms of depression using the CES-D. The items include being bothered by things, having trouble concentrating, feeling depressed, feeling like everything was an effort, feeling hopeful about the future, feeling fearful, having restless sleep, feeling happy, lonely, and unable to get going. Each item includes 4 possible responses: rarely or none in the past week, 1-2 days, 3-4 days, 5-7 days. The intensity of each negative symptom is scored from 0 (rarely or none) to 3 (5-7 days a week). We recode feeling hopeful about the future and feeling happy to reflect the negative symptoms. We use the sum of the scores based on reported symptoms, where higher scores indicate a higher likelihood of having depression.

Intergenerational outcomes

Using children's height and age, we calculate height-for-age z-score using the WHO reference data.[1] Stunting takes the value one if a child's height-for-age is more than two standard deviations below the mean. Using children's hemoglobin count, sex, and age, we identify children with anemia. Specifically, anemia is defined as having a count of less than 11.5 grams of hemoglobin per deciliter (gr/dL) for children under 12 years of age. For children between the ages of 12 and 15, the threshold is 12 gr/dL. The threshold is 12 gr/dL for girls over the age of 14 and 13 gr/dL for boys over the age of 14. Self-reported health takes the value one if the child is reported as being healthy or very healthy.

[1] We use the 2007 WHO growth chart, which is applicable to children between 0 and 19 years of age.

INPRES exposure variable

For the first generation, the IFLS asks respondents over the age of 15 their place of birth in the wave in which they first joined the survey. Additionally, in 2000, IFLS-3 asked all respondents over the age of 15 their district of birth. We combine both sources of information to obtain the respondents' district of birth.

For the second generation, we identify mother-child and father-child pairs based on the relationships within the household. We use mother-child (father-child) pairs by including respondents identified as the biological child of adult female (male) respondents who were born between 1950 and 1972. Additionally, in cases where the child's place and/or date of birth is missing from the household roster, we use women's pregnancy history to identify children born to women who were born between 1950 and 1972.

Table A.1: Summary statistics

	(1) Mean	(2) SD	(3) N
Panel A. First generation			
Male	0.502	0.500	12,137
Born between 1963-1972	0.568	0.495	12,158
INPRES schools per 1,000	2.138	1.251	12,158
Javanese	0.455	0.498	12,158
Year of birth	1,963	6.392	12,158
Primary completion	0.674	0.469	12,158
Self-reported healthy	0.719	0.450	10,801
Days missed activities	1.841	3.267	10,429
Any chronic condition	0.422	0.494	10,738
No. of chronic conditions	0.683	1.007	10,738
Mental health screening score	5.513	4.694	10,254
Panel B. Second generation			
First child	0.330	0.402	14,421
Male child	0.514	0.500	14,421
Child's year of birth	1,992	7.551	14,421
Javanese	0.434	0.496	14,421
Mother born 1963-1972	0.556	0.497	13,322
Father born 1963-1972	0.478	0.500	12,569
Height for age z-score	-1.649	1.096	25,482
Stunted	0.366	0.482	25,482
Anemia	0.249	0.433	21,952
Self-reported health	0.879	0.326	22,748

Notes: Summary statistics for the expanded sample, which includes first generation individuals born between 1950-1972 and their children. The summary statistics for the health outcomes correspond to individuals observed in the Wave 5 of the IFLS and IFLS-E. Second generation height captures multiple observations per child as the IFLS measures height in all waves.

A.3 Data Comparison

Comparison of non-coresident and coresident second generation individuals in the IFLS

The longitudinal nature of the main IFLS allows us to track non-coresident children who make up the second generation. Non-coresident children in the IFLS account for about 45% of the total sample of the second generation individuals in our sample. Including only coresident children in the second generation individuals may introduce some sample selection. To address this, we explore the characteristics of coresident and non-coresident children (Table A.2). We begin by comparing the raw sample mean, followed by the adjusted differences. The adjusted difference takes into account mother's district of birth (in columns 9-10), and then we include mother's year of birth (in columns 11-12). These comparisons suggest that coresident children are more likely to be younger, come from households with a higher asset index, and their parents have higher education. Additionally, co-resident children are more likely to have mothers who were older at the time of their first birth.

Comparison of coresident second generation individuals in the IFLS and a nationally representative survey

We also examine the representativeness of the coresident children's characteristics in the IFLS by comparing the IFLS against the 2014 Socioeconomic survey, 2014 Susenas (Table A.3). Comparing the characteristics of the national sample to the sample restricted to the IFLS and IFLS-E provinces (columns 4-6 of Table A.3) suggests that the child characteristics are similar to children's characteristics in the national sample. Additionally, when we restrict the 2014 Susenas to the IFLS provinces (columns 7-9 of Table A.3), the children's characteristics are similar to the co-resident children in the IFLS in columns 4-6 of Table A.2.

Table A.2: Comparison of non-coresident and coresident children in the IFLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Out of Household			In Household			Raw		Mother's birth place FE		Mother's birth place, year of birth FE	
	Mean	SD	Obs	Mean	SD	Obs	Diff	SE	Mean Diff	SE	Mean Diff	SE
Child male	0.50	0.50	7,214	0.52	0.50	5,475	0.021**	0.01	0.02**	0.01	0.02*	0.01
Child age in 2014	27.00	5.48	6,524	19.40	6.76	5,475	-7.599***	0.13	-7.58***	0.14	-5.65***	0.12
Log HH expenditure	14.59	0.77	6,558	14.82	0.72	5,461	0.223***	0.02	0.19***	0.02	0.20***	0.02
Asset index	-0.12	1.05	6,574	0.13	0.85	5,468	0.253***	0.02	0.21***	0.02	0.23***	0.02
Household size	4.76	2.74	6,716	5.95	2.35	5,475	1.190***	0.07	1.19***	0.07	1.28***	0.07
Child: primary complete	0.65	0.48	7,080	0.58	0.49	5,461	-0.073***	0.01	-0.07***	0.01	-0.04***	0.01
Child: secondary complete	0.32	0.47	7,005	0.31	0.46	5,459	-0.015	0.01	-0.01	0.01	0.01	0.01
Mother: primary complete	0.67	0.47	7,213	0.75	0.43	5,475	0.079***	0.01	0.07***	0.01	0.02**	0.01
Father: primary complete	0.75	0.44	6,646	0.80	0.40	5,191	0.052***	0.01	0.05***	0.01	0.02***	0.01
Mother: years of schooling	6.62	4.12	7,213	7.61	4.30	5,475	0.982***	0.10	0.87***	0.09	0.31***	0.08
Father: years of schooling	7.69	4.25	6,646	8.52	4.31	5,191	0.833***	0.10	0.73***	0.09	0.37***	0.09
Urban	0.60	0.49	6,716	0.61	0.49	5,475	0.017	0.01	-0.02*	0.01	-0.02	0.01
Mother born 1963-72	0.44	0.50	7,214	0.69	0.46	5,475	0.248***	0.01	0.24***	0.01	0.00	0.00
Mother's year of birth	1,961.30	5.75	7,214	1,964.73	5.58	5,475	3.427***	0.15	3.31***	0.14	-0.00	0.00
Father’s year of birth	1,955.95	7.87	6,654	1,960.22	7.64	5,195	4.262***	0.20	4.09***	0.20	0.70***	0.12
Ever moved	0.08	0.28	6,707	0.03	0.18	5,473	-0.050***	0.01	-0.05***	0.01	-0.05***	0.01
Mother's age at first birth	22.65	4.59	2,604	25.54	5.48	1,288	2.894***	0.20	2.81***	0.21	3.61***	0.22

Notes: Observations in cols. 1-6 weighted by IFLS cross sectional weight. Standard errors clustered at mother's district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table A.3: Coresident children characteristics in Susenas 2014

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All provinces			IFLS and IFLS-E provinces			IFLS provinces		
	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs
Child male	0.56	0.50	177,436	0.55	0.50	135,698	0.55	0.50	105,948
Child primary complete	0.86	0.34	174,130	0.87	0.34	133,075	0.87	0.33	104,485
Child secondary complete	0.64	0.48	174,130	0.65	0.48	133,075	0.65	0.48	104,485
Child year of birth	1,993.48	7.34	177,436	1,993.44	7.37	135,698	1,993.34	7.37	105,948
Household size	5.23	1.85	177,436	5.21	1.84	135,698	5.15	1.80	105,948
Mother born 1963-72	0.65	0.48	177,436	0.65	0.48	135,698	0.65	0.48	105,948
Mother's age at marriage	19.76	4.01	177,433	19.72	4.03	135,695	19.62	4.01	105,945
Mother's year of birth	1,964.22	5.83	177,436	1,964.20	5.83	135,698	1,964.18	5.84	105,948
Mother: primary complete	0.75	0.43	160,555	0.75	0.43	122,495	0.75	0.43	97,105
Mother: secondary complete	0.38	0.49	160,555	0.38	0.49	122,495	0.38	0.49	97,105
Mother: years of schooling	7.49	3.79	160,288	7.48	3.79	122,296	7.47	3.80	96,950
Urban	0.54	0.50	177,436	0.57	0.50	135,698	0.59	0.49	105,948

Notes: Observations weighted by Susenas cross sectional weight.

Comparison of primary school effects with a nationally representative survey

The IFLS contains rich information, but its relatively small sample may be a concern. To address this, we examine the representativeness of the IFLS and IFLS-E by comparing our data against the 2014 Socioeconomic survey, 2014 Susenas, since the fifth wave of the IFLS was administered in 2014. The Susenas is a nationally representative survey that is administered annually. The survey covers every district in Indonesia and includes some questions that are also available in the IFLS and IFLS-E. Table A.4 presents primary completion rates using the 2014 Susenas. Panel A presents estimated effects for all provinces and Panel B presents effects for the IFLS and IFLS-E provinces. The effects are similar in panels A and B, and these estimated effects are similar to our main estimates on first generation primary school completion shown in Table B.2.

Table A.4: Primary completion: 2014 Susenas

Panel A. Susenas 2014: all provinces						
	(1)	(2)	(3)	(4)	(5)	(6)
	Expanded sample			Restricted sample		
	All	Male	Female	All	Male	Female
Young cohort	0.013***	0.016***	0.011**	0.017***	0.017***	0.017***
× INPRES	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
No. of obs.	236,270	122,213	114,057	130,855	67,674	63,181
Dep. var. mean (unweighted)	0.77	0.81	0.73	0.77	0.81	0.73
Dep. var. mean (weighted)	0.76	0.80	0.72	0.76	0.80	0.72
R-squared	0.099	0.084	0.107	0.099	0.088	0.111
Panel B. Susenas 2014: restricted to the IFLS and IFLS-E provinces						
	(1)	(2)	(3)	(4)	(5)	(6)
	Expanded sample			Restricted sample		
	All	Male	Female	All	Male	Female
Young cohort	0.017***	0.019***	0.015**	0.022***	0.022***	0.025***
× INPRES	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.007)
No. of obs.	180,283	92,552	87,731	99,308	50,992	48,316
Dep. var. mean (unweighted)	0.77	0.81	0.73	0.77	0.81	0.73
Dep. var. mean (weighted)	0.76	0.76	0.72	0.76	0.76	0.72
R-squared	0.103	0.086	0.111	0.104	0.088	0.117

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. Young cohort for the expanded sample corresponds to those born between 1963 and 1972. Young cohort for the restricted sample corresponds to those born between 1968 and 1972. Covariates include: district, year of birth, month of birth, ethnicity (Javanese dummy), birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

We also compare migration between the IFLS and IFLS-E and 2014 Susenas in Table A.5. Migration is defined as currently residing in a district that is different from one's district of birth. The estimated effects are similar in Panels A and B, and the estimates are similar to our main estimates.

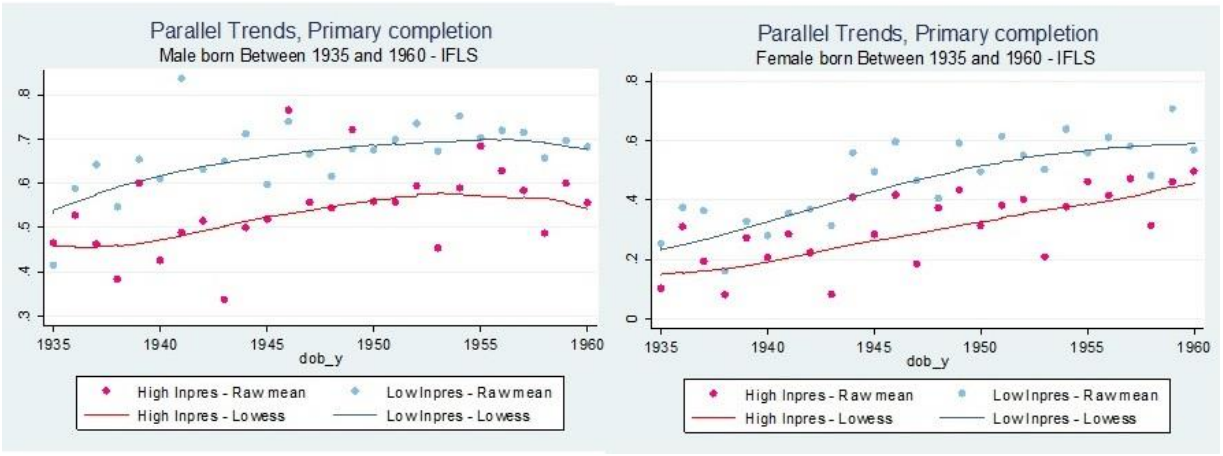
Table A.5: Potential mechanisms: Migration

	(1)	(2)
	Current residence different from birth place	
	Expanded sample	Restricted sample
Panel A. IFLS and IFLS-E data		
Young cohort	0.005	0.001
× INPRES	(0.010)	(0.012)
No. of obs.	13,199	7,356
Dep. var. mean	0.296	0.297
R-squared	0.25	0.25
Panel B. Susenas: All provinces		
Young cohort	0.000	0.001
× INPRES	(0.002)	(0.003)
No. of obs.	258,308	141,048
Dep. var. mean	0.250	0.254
R-squared	0.115	0.117
Panel C. Susenas: Restricted to IFLS and IFLS-E provinces		
Young cohort	0.004	0.001
× INPRES	(0.003)	(0.003)
No. of obs.	198,337	141,048
Dep. var. mean	0.229	0.234
R-squared	0.125	0.117

Notes: Expanded sample includes those born between 1950-1975. Restricted sample includes those born between 1957-1962 or 1968-1975. Young cohort for the expanded sample corresponds to those born between 1963 and 1972. Young cohort for the restricted sample corresponds to those born between 1968 and 1972. See Table 1 for covariates. Ever migrate takes the value one if respondent is not currently residing in his/her district of birth. Robust standard errors clustered at district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

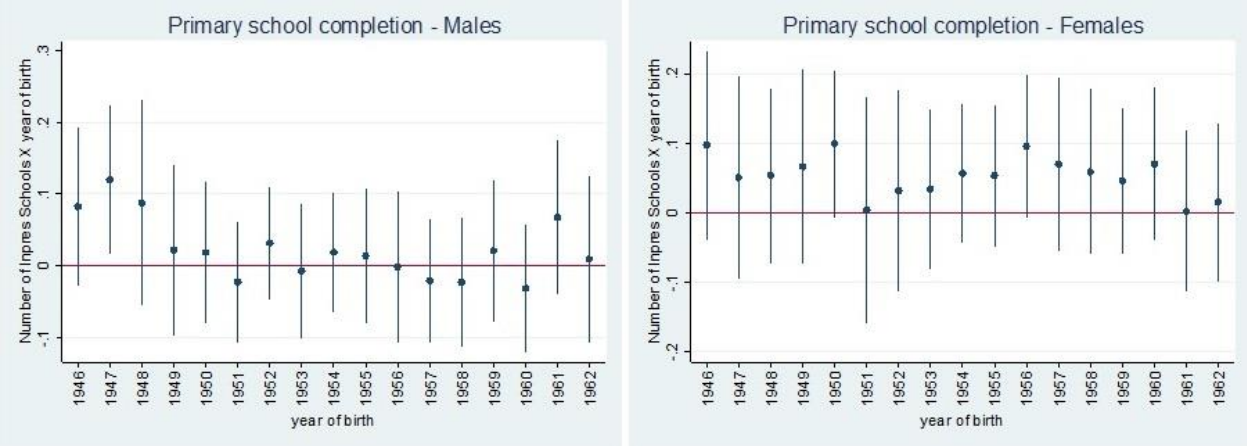
B Additional figures and tables

Figure B.1: Pre-trends raw data: Primary school completion – IFLS



Notes: Primary completion rates for cohorts born between 1935 and 1958 from the main IFLS and IFLS-E.

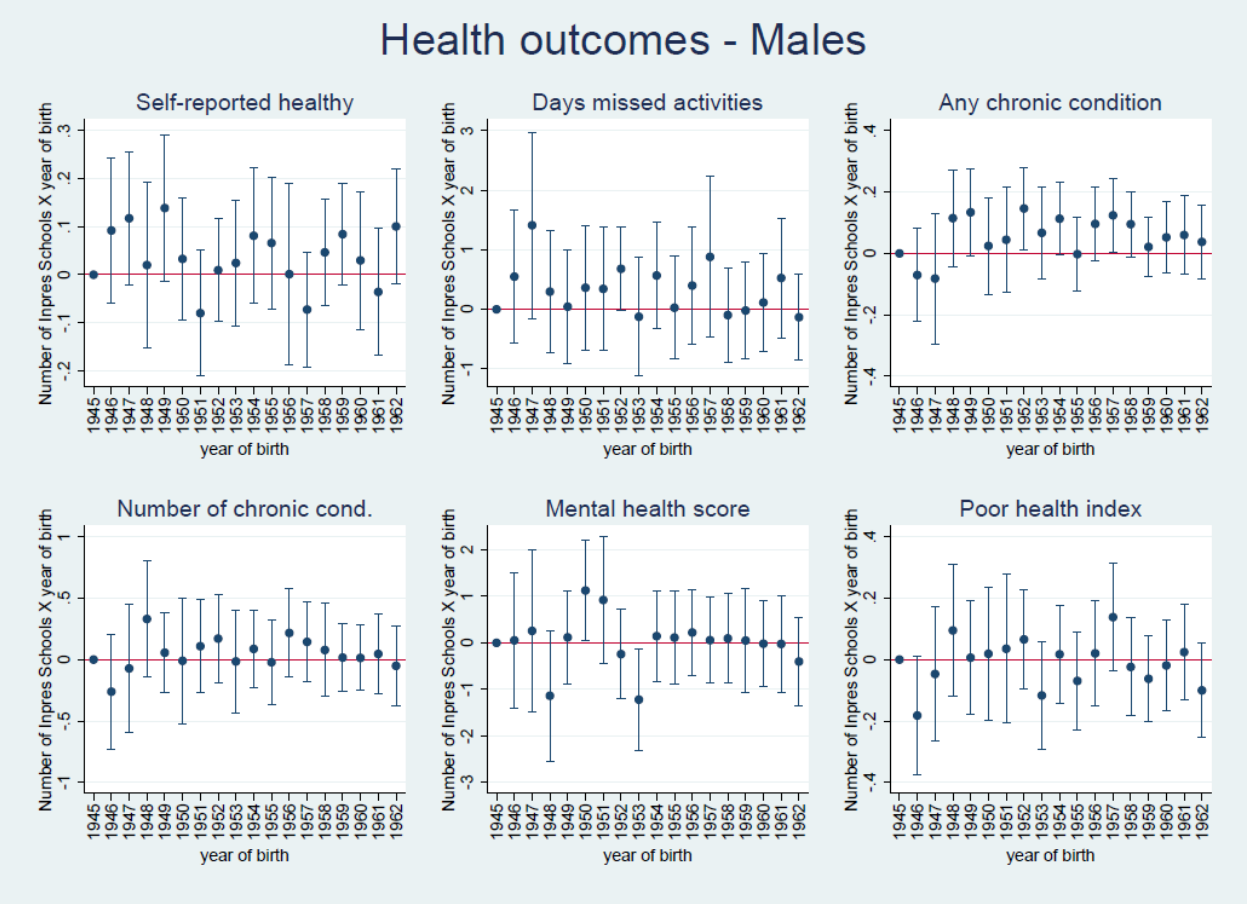
Figure B.2: Pre-trends raw regression: Primary school completion – IFLS



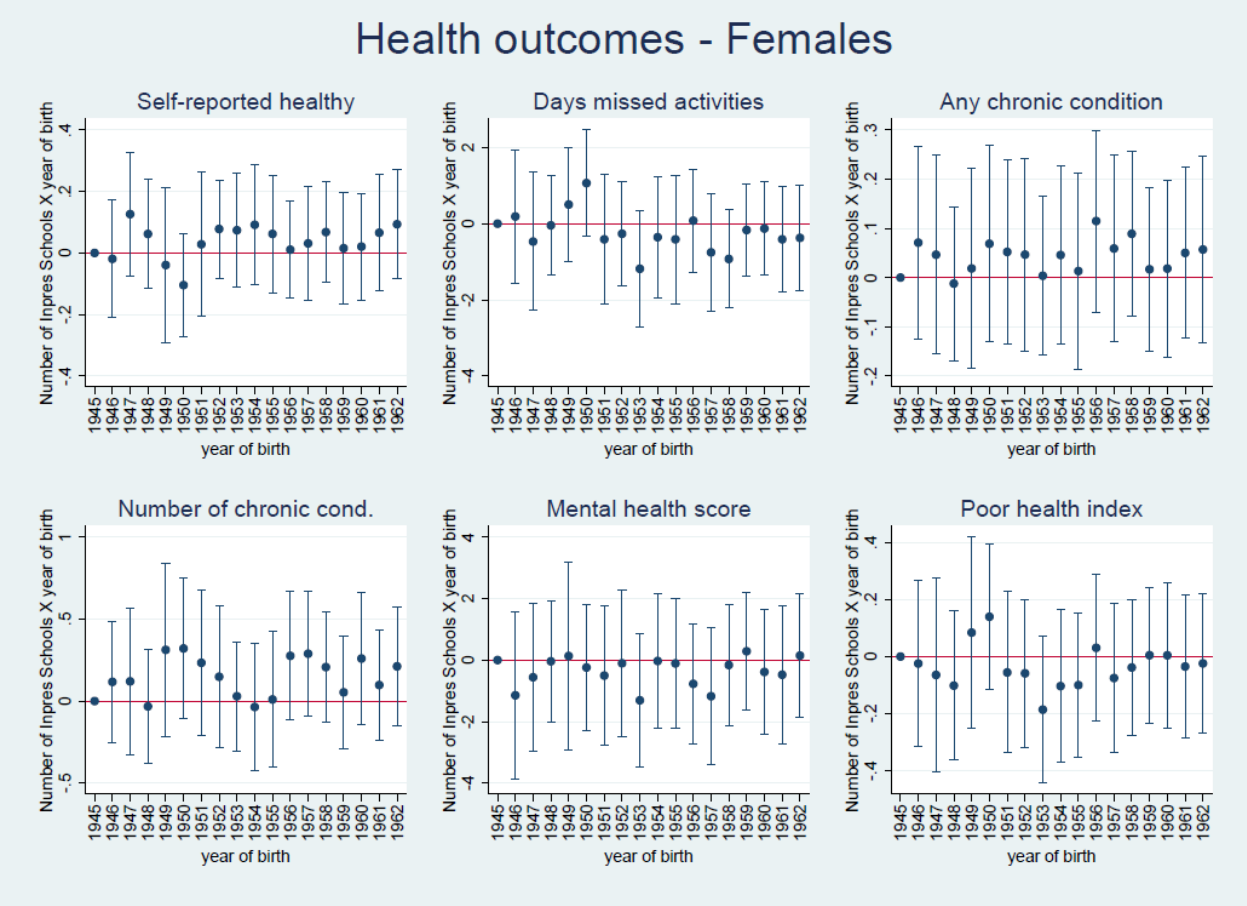
Notes: Coefficients from difference-in-differences model that interacts the number of INPRES schools and year of birth for cohorts born between 1945 and 1962. Year of birth 1945 is the omitted category. Bars indicate the 95% confidence intervals.

Figure B.3: Pre-trends raw regression: First generation's health outcomes

Panel A. Males

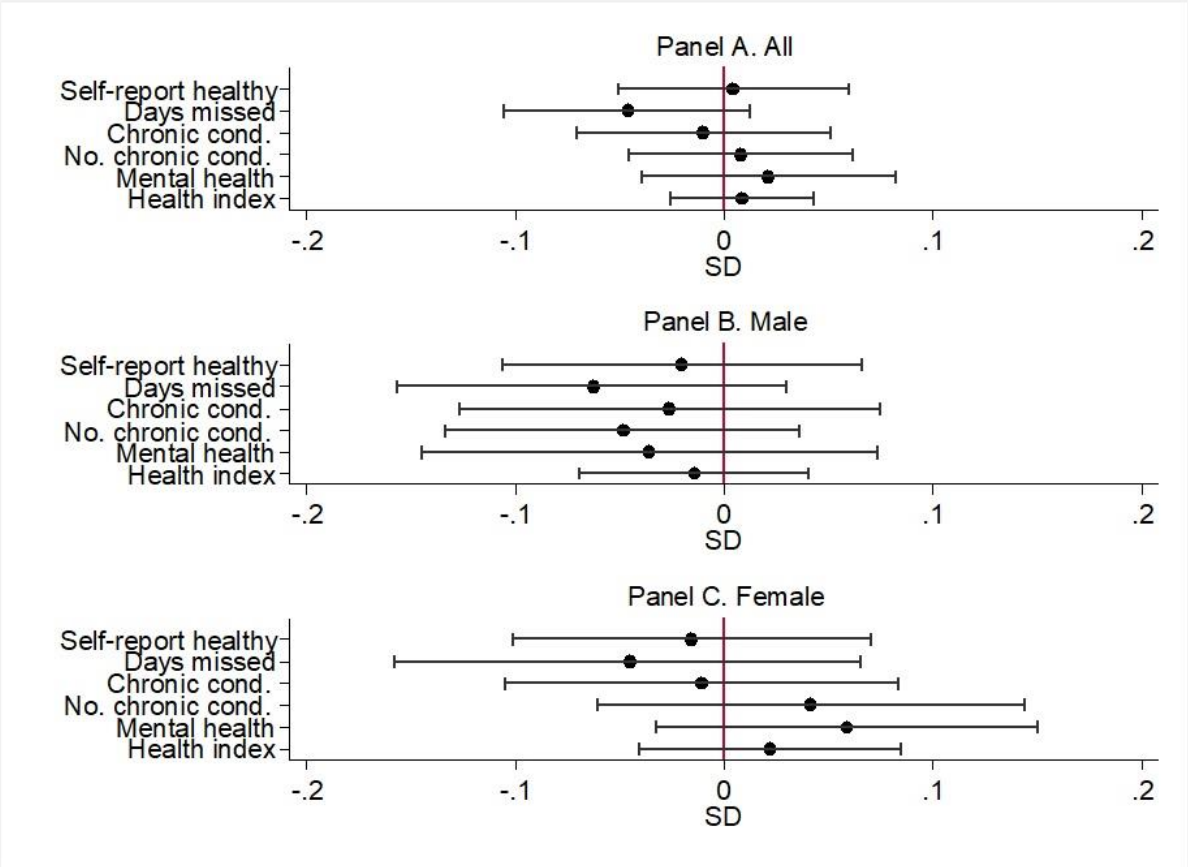


Panel B. Females



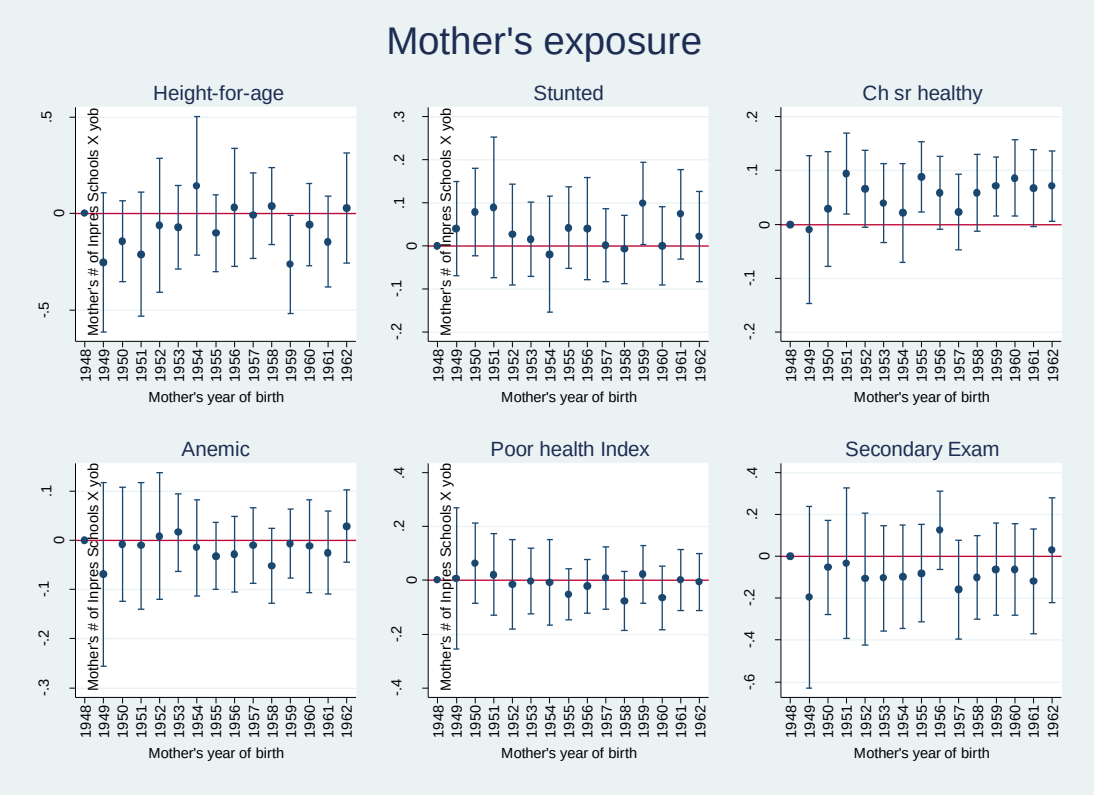
Notes: Coefficients from difference-in-differences model that interacts the number of INPRES schools and year of birth for cohorts born between 1945 and 1962. Bars indicate the 95% confidence intervals.

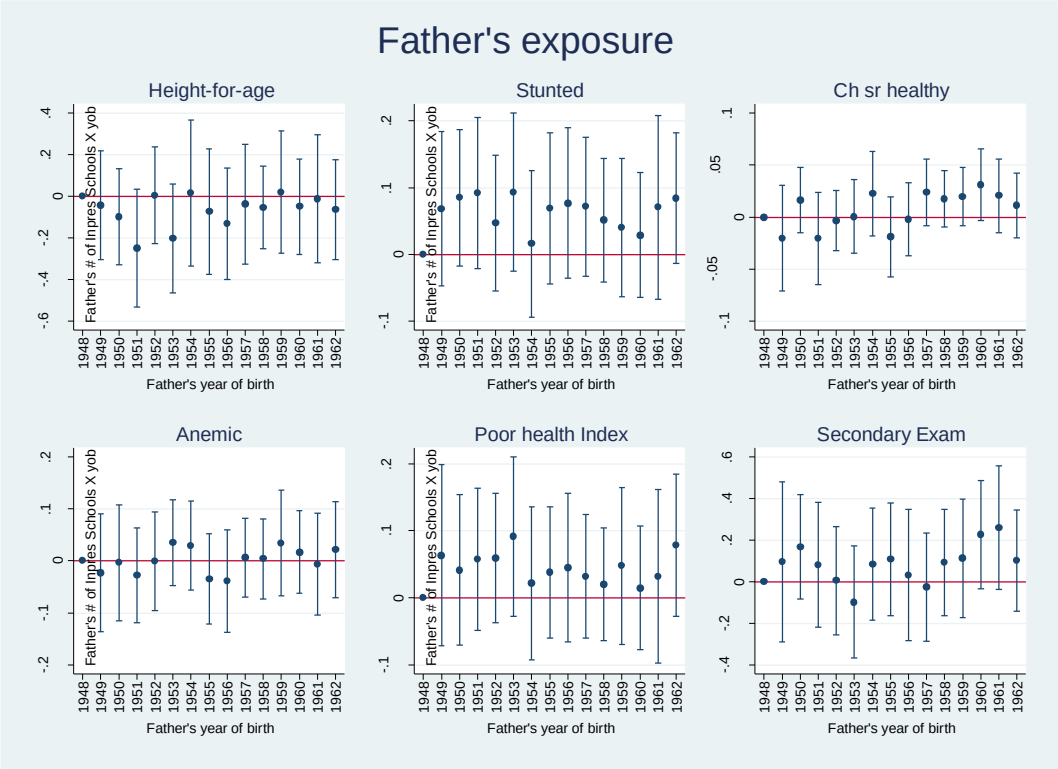
Figure B.4: Placebo regression: First generation



Notes: Sample includes individuals born between 1950 and 1962. "Placebo exposed group" for individuals born between 1957 and 1962. Coefficients reported in standard deviation units. Overall health index corresponds to a summary index from the multiple self-reported health measures analyzed: self-reported general health, days missed, if chronic conditions, number of conditions and mental health screening score. Health index has mean 0, SD 1 based on those born between 1950-1962 in low INPRES areas. Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. Covariates include: district, year of birth, month of birth, ethnicity (Javanese dummy), birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth. 95% confidence intervals are included.

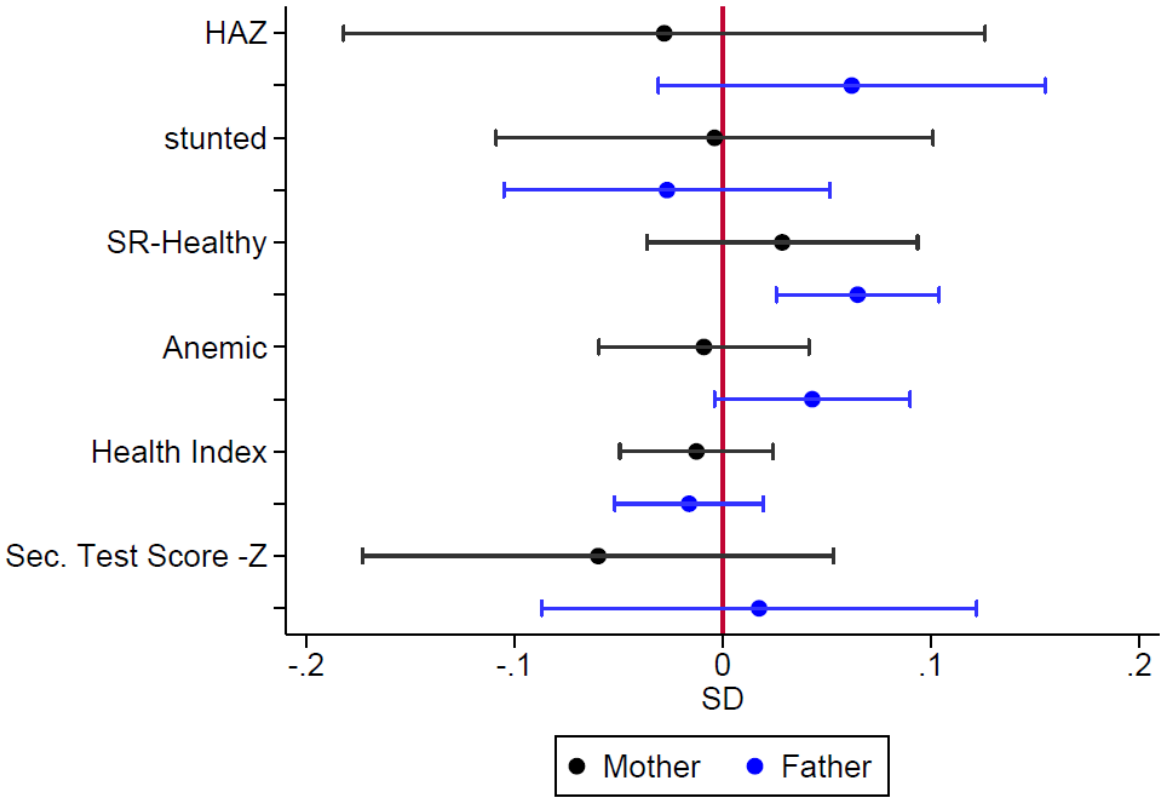
Figure B.5: Pre-trends raw regression: Second generation outcomes





Notes: “Ch sr healthy” corresponds to an indicator of child's self-reported health status as ‘very healthy’ or ‘healthy’. Coefficients from difference-in-differences model that interacts the number of INPRES schools in parent's birth place and parent's year of birth. Sample corresponds to children from parent's born between 1948-1962 (we exclude children from parents born in 1945-1947 because of very small sample sizes). Bars indicate the 95% confidence intervals.

Figure B.6: Placebo regression: Second generation - each parent separately



Notes: Sample includes the children of individuals born between 1950 and 1962. Placebo exposed group is defined as a dummy equal to one for children born to adults born between 1957 and 1962. Coefficients reported in standard deviation units. Overall health index corresponds to a summary index from the multiple health measures analyzed: height for age z-score, stunting, anemia, self-reported general health. Educational outcome is the z-score of the secondary school examination. The health index has mean 0, SD 1. Covariates include the following FE: parent year of birth × 1971 enrollment, parent year of birth × 1971 number of children, parent year of birth × water sanitation program, child's gender, birth order, year and month of birth dummies, urban, and ethnicity (Javanese dummy). Because some children may be observed multiple times, robust standard errors clustered two-way at parent's district of birth and individual level for health outcomes and clustered at parent's district of birth for test scores. 95% confidence intervals are included.

Table B.1: Summary of the Literature

Study	Research question/Outcomes	Data and sample	Methods	Results
First generation medium-term effects				
Duflo (2001)	Medium run effects on education and earnings for men.	Cross-sectional data. Sample of men born between 1950 and 1975 from the Indonesian intercensal survey (SUPAS)	Difference-in-differences exploiting geographic variation on the number of INPRES schools built in district of birth and cohort variation. 2SLS that instruments years of schooling using INPRES exposure.	Each new school constructed per 1000 children increased years of schooling by 0.12 to 0.2 and earnings by 1.5 to 2.7 percent for men.
Duflo and Breinova (2004)	Medium run effects of mother's and father's education on child health and fertility.	Cross-sectional data. Sample of men and women born between 1950 and 1975 from the Indonesian intercensal survey (SUPAS)	2SLS: to instrument for average education in a family, they combine cohort variation and program intensity in exposure to INPRES for each partner. To instrument for the difference in education between husband and wife, they add the interaction between whether the husband was old to benefit from the program (age >12), the age difference between husband and wife and the number of schools constructed in the district of birth.	- Higher mother and father education reduced child mortality. - Female education matters more than male education in delaying marriage and lower fertility (number of children by age 15 or 25). The 2SLS estimates of the impacts of education gaps were not significant.

Karachiwalla and Palloni (2019)	Long-term impact of education on participation in agriculture.	Cross-sectional data. Men born between 1950 and 1972 in SUPAS 1995.	Difference-in-Differences as in Duflo (2001) and 2SLS.	Treated men are more likely to be employed for a wage and work in non-agricultural sectors.
Ashraf et al. (2020)	How the impact of INPRES on women's education depend on bride price customs.	Cross-sectional data. Women born between 1950 and 1972 in SUPAS 1995.	Difference-in-Differences as in Duflo (2001) allowing for heterogeneity by bride price ethnic groups.	Exposure to INPRES increased the probability of primary school completion by 2.5 percentage points for bride price females but no significant effect for non-bride price females.
First generation general equilibrium effects and human capital externalities				
Duflo (2004)	Medium run general equilibrium effects on labor market. Human capital externalities.	Cross-sectional data. Repeated cross-section surveys from the annual Indonesian Labor Force Survey (SAKERNAS), from 1986 to 1999. Sample: males aged 20-60, excluding Jakarta.	2SLS: use interactions between survey year and the number of INPRES school per 1000 children at the district of residence as instruments for the fraction of educated workers in the district. Also, a two-sector model as framework to interpret these findings.	Between 1986 and 1999, an increase of 10 percentage points in the fraction of primary school graduates in the labor force decreased the wages of the older workers (both educated and uneducated) by 4% to 10%.
Martinez-Bravo (2017)	Effect on local governance and provision of public goods.	Cross-sectional data. Information on local leaders from the Village Potential Statistics (PODES) from 1986 to 2003.	Difference-in-differences by comparing villages before and after the first post-1992 election, when individuals exposed to the program could potentially become local leaders.	- Increased provision of public goods: number of doctors, primary health care centers, access to basic services. - Effects are stronger in villages exposed to a higher program intensity.

Zha (2019)	Impacts of a market-level shock to the education distribution on marital outcomes.	Cross-sectional data. Women born between 1953 and 1961 and between 1965 and 1970 in SUPAS 2005 and Census 2010.	Difference-in-Differences as in Duflo (2001) and 2SLS.	- In densely populated areas, INPRES did not affect primary school attainment rate. -INPRES decreased secondary school attainment rate for both men and women due to crowding out of teacher resources. - When average education of other women decreases, a woman marries earlier, holding their potential husbands' education distribution unchanged.
First Generation long-term effects				
Bharati et al. (2016)	Effect on time preference.	Longitudinal data. Individuals born between 1958-1962 and 1968-1972 from the Indonesian Family Life Survey (IFLS).	Difference-in-Differences as in Duflo (2001).	- Higher education among women increased their patience (as measured by the hypothetical time-preference question).
Bharati et al. (2019)	Program effect on education to mitigate early-life weather shocks.	Longitudinal data. Individuals born between 1957-1961 and 1969-1973 from the Indonesian Family Life Survey (IFLS).	Difference-in-differences exploiting variation in rainfall in early life, and geographic variation on the number of INPRES schools built in district of birth and cohort variation.	- Individuals who experienced a negative weather shock in early life and were exposed to the program can catch up in their educational attainment.

First Generation long-term effects and intergenerational effects				
Akresh et al. (2019)	Impacts on long-term and intergenerational effects.	Cross-sectional data. First generation individuals born between 1950 and 1972 and their co-resident offspring from the 2016 Susenas.	Difference-in-Differences as in Duflo (2001).	- First generation: Higher educational attainment for men and women. For men: more likely to be employed, work in the formal sector, work in the non-agricultural sector, and migrate. For women: more likely to migrate and decline in number of children (measured as children co-resident in the household). For households: increase in several measures of expenditures and assets, and higher taxes payments. - Second generation: Mother's exposure increases their children's educational attainment.
Mazumder et al (2019)	Effect on primary school completion and intergenerational effect on primary school performance.	Longitudinal data. First generation individuals born between 1950 and 1972 and their offspring from IFLS and IFLS-East regardless of co-residency.	Difference-in-Differences as in Duflo (2001).	- Higher rates of primary school completion for the first generation of men and women. - Children whose mothers were exposed to the program score higher on the national primary school examination.

Bazzi el al., (2020)	How private school markets (in particular Islamic schools at different levels of education) responded to INPRES mass primary school expansion. Does that process influence identity and beliefs and its transmission to the next generation?	- Mix survey and administrative data. Administrative data of school registries (longitudinal). School curriculum data. Repeated cross-sectional data from National Sample survey (Susenas 2012-2018) to capture individual school enrollment, identity, and religiosity. Cross-sectional data from 2010 population Census and 2008 survey of Islamic beliefs and practices. Data on legislature candidates.	- Individual responses: See Duflo (2001) - School level responses: Difference-in-Differences estimations at the type of school, district, and year of establishment level. Geographic variation: number of INPRES schools built in district. Time variation: post 1972.	- Primary enrollment shifted towards state schools, However, Islamic secondary schools supply responded and absorbed the higher demand for continued education. The Islamic sector also increased religious curriculum. - Individuals exposed report greater religiosity and transmit these religious values to the next generation.

Table B.2: Replication: Primary completion

Panel A. Expanded sample

	(1) All	(2) Male	(3) Female
Born between 1963-72	0.028**	0.025*	0.030*
× INPRES	(0.014)	(0.014)	(0.017)
No. of obs.	13,856	6,991	6,865
Dep. var. mean	0.68	0.74	0.62
R-squared	0.252	0.232	0.290

Panel B. Restricted sample

	(1) All	(2) Male	(3) Female
Born between 1968-72	0.044***	0.032*	0.052***
× INPRES	(0.014)	(0.017)	(0.018)
No. of obs.	7,650	3,869	3,781
Dep. var. mean	0.73	0.78	0.68
R-squared	0.256	0.271	0.290

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. Covariates include: district, year of birth, month of birth, ethnicity (Javanese dummy), birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors in parentheses clustered at the district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.3: First generation health outcomes with multiple hypothesis adjustment

Panel A. Expanded sample		
	(1)	(2)
	Original p-value	Adjusted p-value
Self-reported healthy	0.000	0.000
Number of days missed	0.013	0.026
Any chronic conditions	0.025	0.031
Number of conditions	0.015	0.026
Mental health	0.202	0.202

Panel B. Restricted sample		
	(1)	(2)
	Original p-value	Adjusted p-value
Self-reported healthy	0.009	0.047
Number of days missed	0.437	0.437
Any chronic conditions	0.396	0.437
Number of conditions	0.131	0.219
Mental health	0.078	0.196

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. Covariates include: district, year of birth, month of birth, ethnicity (Javanese dummy), birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.4: Additional health outcomes for the first generation

	Panel A: Expanded sample								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	High BP measured			Diagnosed hypertension			BMI		
	All	Male	Female	All	Male	Female	All	Male	Female
Born bet. 1963-1972	0.002	-0.011	0.012	-0.023**	-0.013	-0.042**	0.162*	0.205	0.086
X INPRES	(0.009)	(0.014)	(0.012)	(0.011)	(0.011)	(0.019)	(0.095)	(0.146)	(0.160)
No. of obs.	10480	5072	5408	10727	5265	5462	10046	5011	5035
Dep. var. mean	0.722	0.710	0.732	0.258	0.185	0.326	24.187	22.915	25.460
R-squared	0.07	0.09	0.11	0.09	0.10	0.11	0.16	0.12	0.11
Gender difference p-value		0.209			0.093			0.587	
	Panel B: Restricted sample								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	High BP measured			Diagnosed hypertension			BMI		
	All	Male	Female	All	Male	Female	All	Male	Female
Born bet. 1968-1972	-0.024**	-0.025	-0.015	-0.018	0.002	-0.038**	0.201*	0.236*	0.269
X INPRES	(0.012)	(0.018)	(0.017)	(0.012)	(0.012)	(0.019)	(0.120)	(0.143)	(0.219)
No. of obs.	5837	2807	3030	5939	2900	3039	5639	2781	2858
Dep. var. mean	0.698	0.687	0.708	0.240	0.152	0.318	24.440	23.172	25.631
R-squared	0.08	0.12	0.13	0.11	0.12	0.14	0.16	0.13	0.14
Gender difference p-value		0.688			0.031			0.897	

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. High blood pressure defined as systolic pressure greater than 130 or diastolic pressure greater than 80. Diagnosed hypertension based on self-reported diagnosis of chronic conditions. BMI defined as a person's weight in kilograms (kg) divided by his or her height in meters squared. District, year of birth, month of birth fixed effects included. Birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.5: Gender-specific first generation's health behavior

Panel A. Expanded sample					
	(1)	(2)	(3)	(4)	(5)
	Male				Female
	Ever smoked	Currently smoking	No. of daily cigarettes	Alcohol expenditure	Teen pregnancy
Born bet. 1963-1972	0.012	0.029**	-0.320	779.799*	-0.008
X INPRES	(0.013)	(0.013)	(0.380)	(408.435)	(0.014)
No. of obs.	5296	5296	4041	6006	6003
Dep. var. mean	0.851	0.632	12.439	268.183	0.25
R-squared	0.12	0.11	0.16	0.08	0.14
Panel B. Restricted sample					
	(1)	(2)	(3)	(4)	(5)
	Male				Female
	Ever smoked	Currently smoking	No. of daily cigarettes	Alcohol expenditure	Teen pregnancy
Born bet. 1968-1972	-0.006	0.022	-0.161	446.972	-0.015
X INPRES	(0.019)	(0.022)	(0.560)	(500.690)	(0.020)
No. of obs.	2930	2930	2242	3407	3310
Dep. var. mean	0.855	0.656	12.550	247.151	0.25
R-squared	0.15	0.14	0.18	0.06	0.16

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. High blood pressure defined as systolic pressure greater than 130 or diastolic pressure greater than 80. High BMI defined as BMI greater than 25. District, year of birth, month of birth fixed effects included. Birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.6: Intergenerational outcomes: health index - by gender

Panel A: Expanded Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mother only			Father only			Both parents		
	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters
Mother bet. 1963-72 X INPRES	-0.034***	-0.027*	-0.047***				-0.037**	-0.022	-0.052**
	(0.011)	(0.015)	(0.015)				(0.015)	(0.021)	(0.021)
Father bet. 1963-72 X INPRES				-0.017	-0.007	-0.029*	-0.003	-0.017	0.010
				(0.012)	(0.016)	(0.016)	(0.017)	(0.024)	(0.022)
No. of obs.	17919	9137	8772	17384	8855	8514	19042	9771	9264
R-squared	0.17	0.20	0.19	0.15	0.18	0.17	0.17	0.20	0.19
Panel B: Restricted sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mother only			Father only			Both parents		
	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters
Mother bet. 1963-72 X INPRES	-0.027*	-0.028	-0.030				-0.026	-0.018	-0.041
	(0.016)	(0.023)	(0.019)				(0.019)	(0.025)	(0.025)
Father bet. 1963-72 X INPRES				-0.019	-0.032	-0.019	-0.010	-0.008	-0.009
				(0.016)	(0.021)	(0.026)	(0.019)	(0.027)	(0.025)
No. of obs.	9911	5068	4829	9007	4627	4357	13707	7073	6617
R-squared	0.17	0.22	0.21	0.15	0.18	0.18	0.17	0.20	0.20

Notes: Overall poor health index corresponds to a summary index from the following health measures for the second generation: being stunted, anemic and self-reported poor health. Sample corresponds to children born to first generation INPRES individuals. Expanded sample includes children born to adults born between 1950-1972. Restricted sample includes children born to adults born between 1957-1962 or 1968-1972. For column 1 and 2, covariates include the following FE: parent year of birth × 1971 enrollment, parent year of birth × 1971 number of children, parent year of birth × water sanitation program, child's gender, birth order, year, and month of birth dummies, urban, and ethnicity (Javanese dummy). Also, we add child age (at survey) and wave indicators. Robust standard errors in parentheses clustered at the parent's district of birth. For column 3 (both parents), the sample corresponds to children either born mothers or fathers in the first generation sample. These models include mother's and father's exposure and the full set of covariates for the mother, while for the father we include: province of birth, two-year bins for year of birth fixed effects and interactions between the father's year of birth (in two-year bins) and the father's district-level covariates. In this estimation, standard errors are clustered two-way at the mother and father district of birth. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table B.7: Intergenerational outcomes: test scores - by gender

Panel A: Expanded Sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mother only			Father only			Both parents		
	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters
Mother bet. 1963-72 X INPRES	0.083***	0.091	0.063				0.110***	0.142**	0.073
	(0.031)	(0.056)	(0.044)				(0.039)	(0.066)	(0.056)
Father bet. 1963-72 X INPRES				0.047	0.025	0.070	0.006	-0.021	0.005
				(0.033)	(0.051)	(0.071)	(0.043)	(0.066)	(0.068)
No. of obs.	6819	3419	3400	5744	2846	2898	6604	3280	3285
R-squared	0.11	0.16	0.16	0.11	0.17	0.17	0.14	0.21	0.19
Panel B: Restricted sample									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Mother only			Father only			Both parents		
	All	Sons	Daughters	All	Sons	Daughters	All	Sons	Daughters
Mother bet. 1963-72 X INPRES	0.105*	0.042	0.185**				0.112**	0.032	0.178**
	(0.059)	(0.105)	(0.073)				(0.057)	(0.095)	(0.082)
Father bet. 1963-72 X INPRES				0.103	0.020	0.139	0.041	-0.008	-0.021
				(0.065)	(0.100)	(0.106)	(0.056)	(0.080)	(0.080)
No. of obs.	3512	1727	1785	2639	1294	1345	4361	2148	2184
R-squared	0.14	0.23	0.23	0.17	0.27	0.30	0.17	0.27	0.26

Notes: Overall poor health index corresponds to a summary index from the following health measures for the second generation: being stunted, anemic and self-reported poor health. Sample corresponds to children born to first generation INPRES individuals. Expanded sample includes children born to adults born between 1950-1972. Restricted sample includes children born to adults born between 1957-1962 or 1968-1972. For column 1 and 2, covariates include the following FE: parent year of birth × 1971 enrollment, parent year of birth × 1971 number of children, parent year of birth × water sanitation program, child's gender, birth order, year, and month of birth dummies, urban, and ethnicity (Javanese dummy). Robust standard errors in parentheses clustered at the parent's district of birth. For column 3 (both parents), the sample corresponds to children either born mothers or fathers in the first generation sample. These models include mother's and father's exposure and the full set of covariates for the mother, while for the father we include: province of birth, two-year bins for year of birth fixed effects and interactions between the father's year of birth (in two-year bins) and the father's district-level covariates. In this estimation, standard errors are clustered two-way at the mother and father district of birth. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table B.8: Intergenerational outcomes with multiple hypothesis adjustment

Panel A. Expanded sample				
	(1)	(2)	(3)	(4)
	Maternal exposure		Paternal exposure	
	Original p-value	Adjusted p-value	Original p-value	Adjusted p-value
Height-for-age	0.067	0.084	0.583	0.583
Stunted	0.065	0.084	0.421	0.527
Anemia	0.014	0.036	0.373	0.527
Self-reported healthy	0.348	0.348	0.334	0.527
Secondary test score	0.008	0.036	0.162	0.527
Panel B. Restricted sample				
	(1)	(2)	(3)	(4)
	Maternal exposure		Paternal exposure	
	Original p-value	Adjusted p-value	Original p-value	Adjusted p-value
Height-for-age	0.336	0.444	0.257	0.428
Stunted	0.285	0.444	0.667	0.667
Anemia	0.355	0.444	0.474	0.593
Self-reported healthy	0.459	0.459	0.116	0.291
Secondary test score	0.076	0.378	0.115	0.291

Notes: Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. Covariates include: district, year of birth, month of birth, ethnicity (Javanese dummy), birth-year interacted with: the number of school-aged children in the district in 1971 (before the start of the program), the enrollment rate of the district in 1971 and the exposure of the district to another INPRES program: a water and sanitation program. Robust standard errors clustered at district of birth. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.9: Intergenerational outcomes: average poor health index

Panel A. Expanded sample			
	(1)	(2)	(3)
	Mother only	Father only	Both parents
Mother bet. 1963-72 X INPRES	-0.034***		-0.026
	(0.013)		(0.017)
Father bet. 1963-72 X INPRES		-0.018	-0.008
		(0.013)	(0.018)
No. of obs.	10712	10805	11703
R-squared	0.19	0.17	0.19
Panel B. Restricted sample			
	(1)	(2)	(3)
	Mother only	Father only	Both parents
Mother bet. 1963-72 X INPRES	-0.021		-0.016
	(0.020)		(0.021)
Father bet. 1963-72 X INPRES		-0.020	-0.008
		(0.019)	(0.020)
No. of obs.	6084	5686	8531
R-squared	0.20	0.17	0.20

Notes: Sample includes the first generation's children who are between ages 8 and 18. Sample corresponds to children born to first generation INPRES individuals. Expanded sample includes children born to adults born between 1950-1972. Restricted sample includes children born to adults born between 1957-1962 or 1968-1972. Covariates include the following FE: parent year of birth and district of birth fixed effects, parent year of birth × 1971 enrollment, parent year of birth × 1971 number of children, parent year of birth × water sanitation program, child's gender, birth order, year and month of birth dummies, ethnicity (Javanese dummy), examination year dummies. Average regressions weighted by the number of observations per child. Robust standard errors in parentheses clustered at the parent's district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.10: Intergenerational outcomes: Complete secondary school

Panel A. Expanded sample			
	(1)	(2)	(3)
	Mother only	Father only	Both parents
Mother bet. 1963-72 X INPRES	-0.007		-0.012
	(0.010)		(0.014)
Father bet. 1963-72 X INPRES		0.015	0.021
		(0.013)	(0.015)
No. of obs.	13446	10764	12466
Dep. var. mean	0.76	0.78	0.76
R-squared	0.23	0.24	0.25
Panel B. Restricted sample			
	(1)	(2)	(3)
	Mother only	Father only	Both parents
Mother bet. 1963-72 X INPRES	0.005		-0.011
	(0.014)		(0.017)
Father bet. 1963-72 X INPRES		0.018	0.024
		(0.021)	(0.017)
No. of obs.	6724	4725	8040
Dep. var. mean	0.77	0.78	0.77
R-squared	0.26	0.27	0.27

Notes: Sample corresponds to children born to first generation INPRES individuals. Expanded sample includes children born to adults born between 1950-1972. Restricted sample includes children born to adults born between 1957-1962 or 1968-1972. Covariates include the following FE: parent year of birth and district of birth fixed effects, parent year of birth × 1971 enrollment, parent year of birth × 1971 number of children, parent year of birth × water sanitation program, child's gender, birth order, year and month of birth dummies, ethnicity (Javanese dummy). Robust standard errors in parentheses clustered at the parent's district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.11: Robustness: Alternative exposure based on schools built per year

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First generation			Second generation			
	Health index			Health index		Test score	
	All	Male	Female	Maternal exposure	Paternal exposure	Maternal exposure	Paternal exposure
Born bet. 1963-1972 × INPRES	-0.032**	-0.027	-0.035*	-0.012	-0.014	0.116***	0.104**
	(0.013)	(0.019)	(0.021)	(0.013)	(0.014)	(0.038)	(0.045)
No. of obs.	9836	4785	5051	17851	17306	6802	5718
R-squared	0.10	0.12	0.11	0.17	0.15	0.11	0.11

Notes: The alternative exposure variable is based on the number of schools built per year for each cohort. Expanded sample includes children born to adults born between 1950-1972. Restricted sample includes children born to adults born between 1957-1962 or 1968-1972. See Table 1, Table 3, and Table 5 for covariates. Robust standard errors clustered at district of birth for cols. 1-3. Robust standard errors in parentheses clustered at the parent's district of birth for cols. 4-5. Robust standard errors in parentheses two-way clustered at the parent's district of birth and individual level for cols. 6-7. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.12: Robustness: Alternative cohorts

Panel A. Expanded sample							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First generation			Second generation			
	Health index			Health index		Test score	
	All	Male	Female	Maternal	Paternal	Maternal	Paternal
Born bet. 1963-1975	-0.041***	-0.021	-0.058***	-0.024**	-0.013	0.102***	0.041
× INPRES	(0.013)	(0.021)	(0.019)	(0.012)	(0.011)	(0.029)	(0.034)
No. of obs.	9891	4817	5074	20341	19145	7338	5981
R-squared	0.10	0.12	0.11	0.16	0.14	0.11	0.11
Panel B. Restricted sample							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First generation			Second generation			
	Health index			Health index		Test score	
	All	Male	Female	Maternal	Paternal	Maternal	Paternal
Born bet. 1968-1975	-0.038**	-0.010	-0.061**	-0.014	-0.012	0.133***	0.086
× INPRES	(0.016)	(0.032)	(0.026)	(0.015)	(0.014)	(0.050)	(0.061)
No. of obs.	5537	2668	2869	12334	10766	4031	2876
R-squared	0.12	0.15	0.14	0.16	0.13	0.13	0.15

Notes: Expanded sample includes those born between 1950-1975. Restricted sample includes those born between 1957-1962 or 1968-1975. See Table 1 , Table 3, and Table 5 for covariates. Robust standard errors clustered at district of birth for cols. 1-3. Robust standard errors in parentheses clustered at the parent's district of birth for cols. 4-5. Robust standard errors in parentheses two-way clustered at the parent's district of birth and individual level for cols. 6-7. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.13: Potential mechanism: Household resources

Panel A. Expanded sample						
	(1)	(2)	(3)	(4)	(5)	(6)
		Food			Non-food	
	All	Male	Female	All	Male	Female
Born bet. 1963-1972	0.040***	0.038*	0.041	0.053**	0.050	0.064
× INPRES	(0.015)	(0.022)	(0.027)	(0.026)	(0.034)	(0.043)
No. of obs.	11941	6007	5934	11925	5998	5927
Dep. var. mean	12.430	12.481	12.378	11.522	11.577	11.467
R-squared	0.14	0.15	0.17	0.16	0.16	0.20
Panel B. Restricted sample						
	(1)	(2)	(3)	(4)	(5)	(6)
		Food			Non-food	
	All	Male	Female	All	Male	Female
Born bet. 1968-1972	0.040*	0.052	0.031	0.069**	0.062	0.105**
× INPRES	(0.023)	(0.032)	(0.034)	(0.034)	(0.048)	(0.053)
No. of obs.	6752	3408	3344	6742	3403	3339
Dep. var. mean	12.458	12.499	12.417	11.584	11.622	11.545
R-squared	0.14	0.15	0.19	0.16	0.17	0.22

Notes: Log per capita expenditure in 2012-14 based on weekly or monthly per capita food and non-food expenditure in 2012 *Rupiah*). Expanded sample includes those born between 1950-1972. Restricted sample includes those born between 1957-1962 or 1968-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table B.14: Potential mechanism: Housing quality (Expanded sample)

	(1) Drinking water	(2) Bad Toilet	(3) Bad Occupancy	(4) Bad Floor	(5) Bad Roof	(6) Bad Wall
Panel A. All						
Born between 1963-1972	0.010	-0.006	-0.015*	-0.006	0.003	-0.012
× INPRES	(0.009)	(0.007)	(0.008)	(0.006)	(0.003)	(0.008)
No. of obs.	11835	11849	11934	12132	12121	12022
Dep. var. mean	0.437	0.086	0.092	0.163	0.033	0.234
R-squared	0.21	0.13	0.11	0.39	0.27	0.33
Panel B. Men						
Born between 1963-1972	0.010	0.001	-0.026**	-0.010	0.000	-0.009
× INPRES	(0.013)	(0.009)	(0.012)	(0.008)	(0.003)	(0.010)
No. of obs.	5887	5891	5944	6042	6036	5986
Dep. var. mean	0.435	0.087	0.094	0.164	0.031	0.235
R-squared	0.24	0.15	0.14	0.43	0.27	0.36
Panel C. Women						
Born between 1963-1972	0.013	-0.011	-0.002	-0.002	0.004	-0.019*
× INPRES	(0.012)	(0.009)	(0.008)	(0.009)	(0.005)	(0.011)
No. of obs.	5948	5958	5990	6090	6085	6036
Dep. var. mean	0.439	0.086	0.089	0.162	0.035	0.233
R-squared	0.22	0.15	0.12	0.38	0.31	0.34

Notes: Poor toilet is captured by not having access to a toilet (including shared or public toilet). Poor floor includes board or lumber, bamboo, or dirt floor. Poor roof includes leaves or wood. Poor wall includes lumber or board and bamboo or mat. High occupancy per room is defined as more than two persons per room in the house (based on household size). Expanded sample includes those born between 1950-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table B.15: Potential mechanism: Housing quality (Restricted sample)

	(1) Drinking water	(2) Bad Toilet	(3) Bad Occupancy	(4) Bad Floor	(5) Bad Roof	(6) Bad Wall
Panel A. All						
Born between 1968-1972	0.016	-0.003	-0.012	-0.006	0.002	-0.020**
× INPRES	(0.013)	(0.008)	(0.010)	(0.008)	(0.004)	(0.009)
No. of obs.	6696	6698	6760	6867	6861	6795
Dep. var. mean	0.449	0.085	0.092	0.161	0.035	0.232
R-squared	0.22	0.14	0.11	0.39	0.27	0.34
Panel B. Men						
Born between 1968-1972	0.006	-0.002	-0.037***	-0.015	0.005	-0.026**
× INPRES	(0.016)	(0.010)	(0.013)	(0.010)	(0.005)	(0.011)
No. of obs.	3319	3319	3363	3408	3405	3371
Dep. var. mean	0.448	0.087	0.096	0.160	0.032	0.231
R-squared	0.25	0.17	0.15	0.43	0.27	0.37
Panel C. Women						
Born between 1968-1972	0.043**	-0.006	0.013	0.000	-0.001	-0.020
× INPRES	(0.020)	(0.012)	(0.014)	(0.012)	(0.007)	(0.014)
No. of obs.	3377	3379	3397	3459	3456	3424
Dep. var. mean	0.450	0.082	0.088	0.162	0.037	0.234
R-squared	0.24	0.17	0.14	0.40	0.33	0.35

Notes: See Table B.14 for notes. Restricted sample includes those born between 1957-1962 or 1968-1972. Robust standard errors clustered at district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table B.16: Robustness: Non-movers only

Panel A. Expanded sample							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First generation			Second generation			
	Health index			Health index		Test score	
	All	Male	Female	Maternal	Paternal	Maternal	Paternal
Born bet. 1963-1972	-0.039*	0.008	-0.074***	-0.037**	-0.030**	0.100***	0.032
× INPRES	(0.023)	(0.033)	(0.028)	(0.014)	(0.015)	(0.038)	(0.036)
No. of obs.	4779	2271	2508	12738	11568	4731	3790
R-squared	0.12	0.15	0.14	0.18	0.16	0.12	0.14
Panel B. Restricted sample							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First generation			Second generation			
	Health index			Health index		Test score	
	All	Male	Female	Maternal	Paternal	Maternal	Paternal
Born bet. 1968-1972	-0.026	0.019	-0.052	-0.044**	-0.032*	0.142*	0.114
× INPRES	(0.025)	(0.047)	(0.032)	(0.017)	(0.016)	(0.076)	(0.071)
No. of obs.	2662	1275	1387	6987	5941	2476	1778
R-squared	0.13	0.16	0.18	0.18	0.16	0.17	0.20

Notes: Expanded sample includes those born between 1950-1975. Restricted sample includes those born between 1957-1962 or 1968-1975. See Table 1, Table 3, and Table 5 for covariates. Robust standard errors clustered at district of birth for cols. 1-3. Robust standard errors in parentheses clustered at the parent's district of birth for cols. 4-5. Robust standard errors in parentheses two-way clustered at the parent's district of birth and individual level for cols. 6-7. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.17: Potential mechanism: Neighborhood quality (Expanded sample)

	(1) Elementary school	(2) Middle school	(3) High school	(4) Piped water	(5) Private providers	(6) Clinic	(7) Midwives	(8) Rice subsidy	(9) National health insurance	(10) District health insurance
Panel A. All										
Born between 1963-1972	0.001	0.028	0.044	0.010	0.004	0.010	0.028	-0.000	0.001	0.004
× INPRES	(0.090)	(0.059)	(0.060)	(0.011)	(0.006)	(0.025)	(0.025)	(0.005)	(0.006)	(0.006)
No. of obs.	9329	9329	9329	9140	9329	9329	9267	9063	8538	8394
Dep. var. mean	5.466	4.451	4.735	0.342	0.863	2.110	1.205	0.326	0.328	0.261
R-squared	0.55	0.49	0.48	0.43	0.51	0.42	0.44	0.47	0.42	0.42
Panel B. Men										
Born between 1963-1972	0.114	0.134	0.123	0.022	-0.003	0.037	0.047	0.005	0.008	0.011
× INPRES	(0.129)	(0.105)	(0.104)	(0.019)	(0.008)	(0.031)	(0.032)	(0.009)	(0.011)	(0.011)
No. of obs.	4526	4526	4526	4430	4526	4526	4495	4399	4156	4099
Dep. var. mean	5.405	4.397	4.679	0.331	0.853	2.099	1.213	0.331	0.331	0.260
R-squared	0.57	0.50	0.50	0.43	0.53	0.44	0.47	0.48	0.43	0.42
Panel C. Women										
Born between 1963-1972	-0.035	-0.017	0.018	-0.004	0.009	0.001	0.004	-0.008	-0.003	-0.003
× INPRES	(0.125)	(0.068)	(0.083)	(0.013)	(0.009)	(0.036)	(0.029)	(0.007)	(0.006)	(0.008)
No. of obs.	4776	4776	4776	4683	4776	4776	4744	4637	4353	4269
Dep. var. mean	5.519	4.507	4.795	0.352	0.874	2.120	1.200	0.322	0.326	0.263
R-squared	0.56	0.52	0.50	0.47	0.52	0.44	0.46	0.49	0.44	0.46

Notes: Poor toilet is captured by not having access to a toilet (including shared or public toilet). Poor floor includes board or lumber, bamboo, or dirt floor. Poor roof includes leaves or wood. Poor wall includes lumber or board and bamboo or mat. High occupancy per room is defined as more than two persons per room in the house (based on household size). Expanded sample includes those born between 1950-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table B.18: Potential mechanism: Neighborhood quality (Restricted sample)

	(1) Elementary school	(2) Middle school	(3) High school	(4) Piped water	(5) Private providers	(6) Clinic	(7) Midwives	(8) Rice subsidy	(9) National health insurance	(10) District health insurance
Panel A. All										
Born between 1963-1972	-0.065	-0.012	0.034	0.017	0.008	0.061*	0.048	0.003	-0.007	-0.007
× INPRES	(0.110)	(0.072)	(0.093)	(0.012)	(0.010)	(0.035)	(0.031)	(0.008)	(0.006)	(0.009)
No. of obs.	5118	5118	5118	5015	5118	5118	5077	4959	4672	4599
Dep. var. mean	5.514	4.421	4.670	0.354	0.861	2.092	1.216	0.323	0.329	0.266
R-squared	0.57	0.52	0.49	0.46	0.50	0.44	0.47	0.46	0.42	0.42
Panel B. Men										
Born between 1963-1972	0.117	-0.024	0.110	0.018	-0.002	0.060	0.031	0.014	0.002	0.009
× INPRES	(0.162)	(0.104)	(0.106)	(0.018)	(0.009)	(0.040)	(0.043)	(0.011)	(0.011)	(0.014)
No. of obs.	2464	2464	2464	2411	2464	2464	2442	2387	2251	2221
Dep. var. mean	5.419	4.339	4.635	0.338	0.852	2.064	1.226	0.323	0.329	0.272
R-squared	0.60	0.51	0.52	0.47	0.52	0.44	0.49	0.49	0.45	0.45
Panel C. Women										
Born between 1963-1972	-0.117	0.033	0.047	0.022	0.017	0.067	0.067*	-0.013	-0.017**	-0.019
× INPRES	(0.160)	(0.115)	(0.140)	(0.015)	(0.013)	(0.056)	(0.040)	(0.011)	(0.008)	(0.016)
No. of obs.	2617	2617	2617	2568	2617	2617	2598	2534	2383	2341
Dep. var. mean	5.596	4.493	4.702	0.365	0.869	2.117	1.211	0.324	0.328	0.259
R-squared	0.59	0.57	0.50	0.49	0.52	0.48	0.50	0.48	0.43	0.46

Notes: Poor toilet is captured by not having access to a toilet (including shared or public toilet). Poor floor includes board or lumber, bamboo, or dirt floor. Poor roof includes leaves or wood. Poor wall includes lumber or board and bamboo or mat. High occupancy per room is defined as more than two persons per room in the house (based on household size). Expanded sample includes those born between 1950-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table B.19: Potential mechanisms and their contributions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Association between	Expanded sample			Restricted sample		
	channel and outcome	INPRES effect on channel	INPRES effect on outcome	Contribution (%)	INPRES effect on channel	INPRES effect on outcome	Contribution (%)
Panel A: First generation							
Outcome: Poor health index							
Log per cap expenditures	0.016	0.046	0.04	1.840%	0.054	0.033	2.618%
Neighborhood quality:							
Health resources	0.018	0.024	0.04	1.080%	0.055	0.033	3.000%
Assortative mating for women	0.025	0.006	0.04	0.382%	0.032	0.033	2.468%

Notes: Expanded sample includes those born between 1950-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth.
Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table B.19: Potential mechanisms and their contributions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Association between	Expanded sample			Restricted sample		
	channel and outcome	INPRES effect on channel	INPRES effect on outcome	Contribution (%)	INPRES effect on channel	INPRES effect on outcome	Contribution (%)
Panel B: Second generation							
Outcome: Poor health index							
Number of children	0.019	-0.048	-0.034	2.723%	-0.052	-0.027	3.714%
Log per cap expenditures	-0.045	0.046	-0.034	6.131%	0.054	-0.027	9.063%
Neighborhood quality:							
Health resources	-0.036	0.024	-0.034	2.568%	0.055	-0.027	7.410%
Mother's education	-0.071	0.03	-0.034	6.248%	0.052	-0.027	13.637%
Outcome: Height for age z-score							
Number of children	-0.072	-0.048	0.056	6.172%	-0.052	0.044	8.510%
Log per cap expenditures	0.179	0.046	0.056	14.727%	0.054	0.044	22.003%
Neighborhood quality:							
Health resources	0.079	0.024	0.056	3.370%	0.055	0.044	9.828%
Mother's education	0.163	0.03	0.056	8.744%	0.052	0.044	19.289%
Outcome: Test scores							
Number of children	-0.040	-0.048	0.083	2.285%	-0.052	0.105	1.956%
Log per cap expenditures	0.117	0.046	0.083	6.484%	0.054	0.105	6.017%
Mother's education	0.177	0.03	0.083	6.409%	0.052	0.105	8.782%

Notes: Expanded sample includes those born between 1950-1972. See Table 1 for covariates. Robust standard errors clustered at district of birth.
Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.

C Cost-Benefit Analysis

Cost

We follow Duflo (2001) in our cost estimation. The cost of the school construction and the number of schools built each year came from Duflo (2001). We assume the schools were built between 1973 and 1977 and were operational for 20 years, so the school's last year of operation is 1997. Each school was designed with 3 classrooms and 3 teachers. Teacher's salary was USD 360 (in 1990 USD) in 1974 and USD 2467 (in 1990 USD) in 1995, and we assume linear growth between 1974 and 1995. We assume each teacher would require training, and that would cost a third of the salary. The maintenance cost is assumed to be 25% of the wage bill.

Cohort size

We estimate the number of first generation individuals exposed to the program, starting from those born in 1963. Assuming children start school at 6, the last cohort to benefit is born in 1989. We use the 1971, 1980, and 1990 Census to obtain the population of the cohort born between 1963 and 1989. We then estimate the number of students enrolled based on the enrollment rates between 1970 and 1995 from the World Development Indicators.

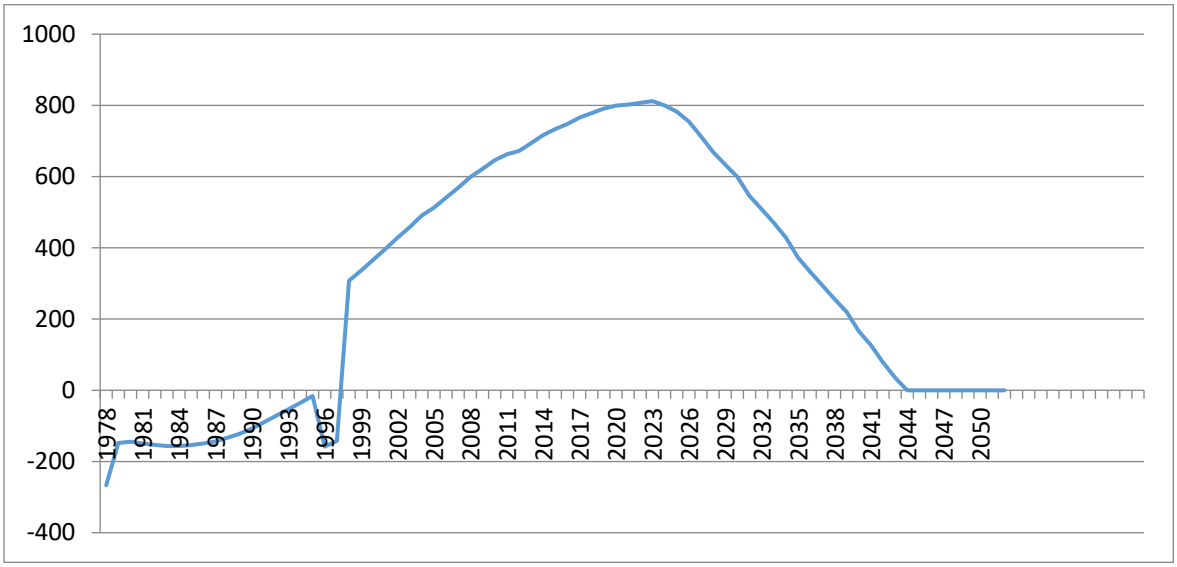
The program sought to attain a teacher student ratio of 40 students per teacher. Each school would typically hold a morning and afternoon session, with 3 classes in each session, so each school served 240 students.[2] To obtain the INPRES coverage for each year, we use the number of INPRES schools in each year divided by the number of primary school-aged children (6-12 year olds) enrolled in primary school. The INPRES exposure for each cohort is given by the average INPRES coverage for the cohort's primary schooling (6 years).

First generation benefits

We follow the literature and assume that returns to primary education is 20% (Psacharopoulos and Patrinos*, 2004). To calculate the base earnings on which to apply the return to primary schooling, we assume the Indonesian population will be in the labor force between the ages of 16 and 55, which is the official age of retirement up to the early 2000s. We calculate the mean earnings of individuals aged 16 to 55 in IFLS-1. We then use the CPI to deflate earnings to 1990 USD. The estimated lifetime earnings are about USD 57000 (in 1990 USD).

[2] Conversations with Bappenas and former Bappenas officials.

Figure B.7: Cost Benefit Analysis



Notes: Cost benefit in USD million. We assume the benefits are derived solely from the earnings gain of the first generation individuals. Cohorts born in 1963 to 1989 benefit from the program. Benefits accrue from 1979, the first year that the first cohort entered the labor market, and end in 2052, when the last cohort served by the program would retire.

We include the gains from health based on the relationship between poor health and mortality at older ages. We follow the literature and assume that self-reported poor health is associated with a 2.73 odds ratio among those 50 and above in Indonesia (Frankenberg and Jones, 2004). We then combine this with mean earnings between the ages of 50 and 55 (in 1990 USD) and estimated survival probability for those ages from Statistics Indonesia.[3]

Second generation benefits

To obtain the cohort size in the second generation, we assume each first generation individual has 1.2 children at age 22.[4] We assume second generation individuals have 20% higher lifetime earnings compared to the first generation individuals and the second generation would be in the labor market between the ages of 16 and 55. The effect of INPRES on the second generation's height is 0.056 standard deviations. With about 6 centimeters standard deviation in height, this would correspond to about 0.366 centimeters height increase. With an 8% gain in earnings resulting from the height premium (Sohn, 2015), the program effect would then translate to a 0.26% gain in lifetime earnings for the second generation. For the second generation gain in education, we use literacy as a proxy for gains in test score. Following the literature, we assume a one standard deviation increase in literacy would increase earnings by 8.5% (Perez-Alvarez, 2017).

[3] Pengembangan Model Life Table Indonesia (2011). Last accessed: July 15, 2019.

[4] Indonesia's total fertility rate in 2000 is 2.4 per woman, so we assume a fertility rate of 1.2 for the first generation individuals.

Table C.1: Internal Rate of Return Estimates

		Internal Rate of Return
First generation	Earnings returns to Primary Completion	7.91%
	+ earnings gains from better health and lower mortality	8.75%
Second generation	1st gen. gains + returns to height	14.53%
	1st gen. gains + returns to test scores	21.48%
	1st gen. gains + returns to height and test scores if independent	24.76%

Notes: First generation earnings based on returns to primary completion. First generation health gains based on earnings gains between 50 and 55 from program reduction in poor health that is associated with mortality improvement. Test score gains based on earnings gain from improved secondary test score. Health gains based on the height premium. Test score and health are assumed to be independent.

The program effect would then translate to a 0.68% gain in lifetime earnings for the second generation.

Scenarios

We present calculations based on several scenarios (Table C.1). First, we assume the gains for the first generation came from earnings only. Next, we assume the gains for the first generation came from earnings and mortality gains. We then add the gains from the second generation's test score alone, the second generation's health alone, and finally, we assume the gains from health and test scores are independent and combine the gains.