Long-term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia*

Richard Akresh
University of Illinois at Urbana-Champaign

Daniel Halim
World Bank

Marieke Kleemans
University of Illinois at Urbana-Champaign

June 2019

Abstract

We study the long-term and intergenerational effects of additional education, focusing on Indonesia's 1970s school construction program, one of the largest ever conducted. Exploiting variation across birth cohorts and districts in the number of schools built suggests education benefits for men and women persist 43 years after the program. Exposed men are more likely to be formal workers, work outside agriculture, and migrate. Exposed women migrate more and have fewer children. Households have improved living standards and pay more government taxes. Education benefits are transmitted to their children, with larger intergenerational effects if mothers, instead of fathers, are exposed to school construction. Intergenerational effects appear larger for daughters and appear to be driven by improved marriage partner's characteristics, including more education and secure employment.

JEL codes: I2, J62, O15, O22, J13

---

* We thank Manuela Angelucci, Catia Batista, Emily Beam, Taryn Dinkelman, Esther Duflo, Willa Friedman, Sylvie Lambert, Nicholas Li, Leigh Linden, Karen Macours, Edward Miguel, Adam Osman, Dean Spears, Rebecca Thornton, Pedro Vicente, and Bruce Wydick, as well as seminar participants at the University of Texas at Austin, Paris School of Economics, Universidade Nova de Lisboa, NEUDC, University of Illinois at Urbana-Champaign, Development Day at the University of Chicago, University of Minnesota, Wellesley College, DePaul University, PACDEV, Economic Demography Workshop, IZA Workshop in Gender and Family Economics, MWIEDC, SOLE, SEHO, Université Libre de Bruxelles, Aix-Marseille School of Economics, IZA/WB/NJD Conference on Jobs and Development, and the RISE annual conference for many helpful discussions and suggestions. All errors remain our own.

Contact information: University of Illinois at Urbana-Champaign. Department of Economics, 1407 W. Gregory Drive, David Kinley Hall Room 214, Urbana, Illinois 61801. Email: akresh@illinois.edu; dhalim@worldbank.org; kleemans@illinois.edu

1. Introduction

The question of which adult outcomes are affected by increases in educational attainment and whether these effects persists into the next generation is of broad research interest and great policy importance. Governments in developing countries spend approximately one trillion dollars annually on education, and households spend hundreds of billions more on the education of their children (Glewwe and Muralidharan, 2016). While much of this government spending is motivated by the belief that increases in education will translate to higher economic development and growth, the causal effect of schooling on economic growth is not uncontested and is often difficult to estimate.[1] In recent years, researchers have made significant advancements using randomized experiments, but in addition to potential challenges of external validity, the vast majority of interventions are evaluated soon after the intervention ends.[2] The question remains whether effects persist over time and continue into the next generation, or whether they fade out as some recent studies have indicated.[3]

In this paper, we test this directly by studying the long-term and intergenerational impacts of one of the largest primary school construction programs ever completed. Between 1973 and

[1] An extensive literature in macroeconomics and growth has pointed to a high correlation between cross-country differences in per capita income and in education, but some have argued that these may reflect reverse causality of increased educational attainment in anticipation of high rates of economic growth (Bils and Klenow, 2000). Microeconomic analyses of the returns to schooling date back to Gorseline (1932) and Walsh (1935) and have long recognized that without (quasi) exogenous variation in educational attainment, the causal impact of education is hard to estimate because the choice of how much education to obtain is correlated with a large number of individual, household, and community characteristics .

[2] Reviewing 111 primary school interventions in developing countries, McEwan (2015) finds that only 10 percent had any evaluation take place more than one month after the intervention had ended. Notable exceptions include Baird, Hamory Hicks, Kremer, and Miguel (2016) who show positive labor market impacts 10 years after a deworming intervention in Kenya, and Gertler et al. (2014) showing higher earnings 20 years after an early child stimulation program in Jamaica. Evidence from the U.S. indicates preschool and kindergarten programs lead to improved adult outcomes (Garces, Thomas, and Currie, 2002; Heckman et al., 2010; Chetty et al., 2011) as do health interventions (Bhalotra and Venkataramani, 2018).

[3] Evans and Ngatia (2018) find that positive outcomes from a free school uniform program are gone eight years after the intervention. Blattman, Fiala, and Martinez (2018) find that grants to Ugandan youth show significant positive effects on earnings after four years but no effect by nine years. Andrabi et al. (2011) using data from Pakistan find that only one-fifth to one-half of student learning persists between grades. Jacob, Lefgren, and Sims (2010) find low persistence of teacher learning in the U.S. with three-quarters or more fading out within one year.

1979, the Indonesian government constructed over 61,000 primary schools, averaging two schools per 1,000 primary school age children. We use 2016 nationally representative Indonesian data to examine the effects on a wide range of outcomes, including education, employment, migration, marriage, fertility, living standards, and taxes. Following the seminal work by Duflo (2001) who studies the effects of this school construction program on men's education and earnings in 1995, we employ a difference-in-differences strategy, exploiting variation across districts in the number of schools built and across birth cohorts in their exposure to the schools.

In addition to studying the long-term effects of this program on an extensive range of outcomes, many of which have not been studied before or have only been studied in isolation, one of the main contributions of this paper is the focus on the program's intergenerational impacts. A large literature on the intergenerational transmission of human capital has mainly focused on developed countries and has documented considerable persistence in economic opportunities as a source of increased economic inequality.[4] Evidence on intergenerational human capital spillovers from developing countries is sparse and we know even less on whether government interventions, such as school construction, can improve educational opportunities for disadvantaged children and improve intergenerational mobility.[5] We are also able to extend the

[4] Black and Devereux (2011) review the large literature on the intergenerational transmission of human capital that measures the persistence between parents' and children's educational attainment, while Currie (2011) and Almond and Currie (2011) provide a review of the long-term effect on education of negative shocks while in utero or early childhood. In addition to the focus on the estimation of correlations between parent and child educational outcomes, recently there is an increased emphasis on estimating causal relationships. Researchers have used changes in school compulsory laws (see Chevalier, 2004 for U.K; Black, Devereux, and Salvanes, 2005 for Norway; Oreopoulos, Page, and Stevens, 2008 for U.S.), other educational policies (Currie and Moretti, 2003; Maurin and McNally, 2008), and environmental shocks (Black et al., forthcoming) to estimate these effects. There is however limited evidence from developing countries.

[5] Recent research, conducted contemporaneously with ours, examines the effect of the same Indonesia school construction program on children's national primary school examination scores (Mazumder, Rosales-Rueda, and Triyana, 2019). They use the Indonesian Family Life Survey for their analysis, and in Section 3 and Appendix B, we highlight the reasons we instead use the National Socioeconomic Survey (Susenas) for studying intergenerational human capital impacts. In addition to the work by Duflo (2001), research by Breierova and Duflo (2004), Martinez-Bravo (2017), Ashraf et al. (forthcoming), Rohner and Saia (2019), and Karachiwalla and Palloni (2019) measure

focus on working-age men in Duflo (2001) to study the impact of school construction on women and observe gender differences for both first and second-generation outcomes. This extension allows us to explore marriage market outcomes that appear to play a crucial role in the intergenerational transmission of human capital. Finally, for this school construction program that has large up-front costs and then benefits that are dispersed over time, we perform a detailed cost-benefit analysis to assess the welfare implications. In addition to calculating benefits based on improvements in living standards, we use tax data to study whether school construction pays for itself in terms of higher future government tax revenues.

Figure 1 provides an overview of our findings. Due to the data's richness and the large number of outcomes, we want to be careful not to overemphasize any single significant result and so we take two approaches. First, following Kling, Liebman, and Katz (2007), we create an index for each family of outcomes where we aggregate all individual outcomes in that family together. As described further in Section 3, we then estimate standardized effects from exposure to school construction on these outcome indexes (Banerjee et al., 2015). Second, since we examine multiple outcomes, we correct for the potential issue of simultaneous inference using multiple hypothesis testing. We calculate q-values using the Benjamini-Hochberg step-up method to control for the false discovery rate (Benjamini and Hochberg, 1995).

The consistent pattern seen in Figure 1 is that exposure to school construction improves most outcomes we are able to explore in the data.[6] School construction, not surprisingly, leads to improved educational outcomes. Duflo (2001) previously showed this for men, and we confirm it

various impacts of the school construction program in Indonesia. Wantchekon, Klasnja and Novta (2014) study human capital externalities after school construction in Benin.

[6] An increase of one additional school built per 1,000 children would increase these indexes for those exposed to the school construction by 0.02 to 0.07 standard deviations relative to the control group.

also improves women's education.[7] The education effects for women are concentrated in primary school only, while men also see significant increases in lower and upper secondary education. As adults, men who are exposed to the program are more likely to be employed, to work in the formal sector, to work in the non-agricultural sector, and to migrate. Women are more likely to migrate and have fewer children. Households in which either parent is exposed to the program have higher living standards, better housing, more assets, and pay more government taxes. While nutrition and health investments increase, we do not observe any improvements in health outcomes. School construction leads to improved marriage market outcomes with spouses being more educated, more likely to be literate, and more likely to have migrated.

Parents transmit these effects to the next generation, who have more education, with larger impacts in secondary and tertiary education. Mother's exposure to school construction has a larger intergenerational education effect than father's exposure. We perform a mediation analysis indicating that the intergenerational transmission of human capital appears to be driven by changes in parents' marriage outcomes, especially whether the spouse completed primary school, is literate, and works in the formal sector and outside of agriculture.

To quantify the policy implications, we conduct a cost-benefit analysis in which we create an accounting model to calculate the discounted costs of school construction and subsequent benefits for the government in terms of increased tax revenues and overall improved living standards for the Indonesian population. Across a range of different parameter estimates, we find that school construction leads to increased government tax revenues that directly offset

[7] This finding is consistent with Breierova and Duflo (2004) who find similar point estimates for female education. In addition to the work by Duflo focusing on Indonesia, studies evaluating school construction projects have been carried out in Mozambique (Handa, 2002), Pakistan (Alderman, Kim, and Orazem, 2003), Afghanistan (Burde and Linden, 2013) and Burkina Faso (Kazianga et al., 2013). These studies focus on improvements in enrollment rates, as opposed to later-life outcomes, and all confirm large increases in school enrollment.

construction costs in most cases within 40 years. Furthermore, accounting for improved living standards of the Indonesian population reveals high internal rates of return ranging from 13-21 percent and benefits surpassing costs within 17-30 years after the schools were built. These results provide strong support for the cost-effectiveness of school construction interventions.[8]

The rest of this paper is organized as follows. Section 2 describes the institutional context and school construction program in Indonesia. Section 3 describes the empirical identification strategy and the data. Section 4 presents the results examining the effects of exposure to school construction on a range of long-term outcomes and Section 5 discusses the intergenerational effects. Section 6 shows results of a number of robustness checks. Section 7 presents the cost-benefit analysis and Section 8 concludes.

2. Institutional Context

Indonesia is the fourth most populous country in the world and the seventh largest economy in terms of total GDP at purchasing power parity. The country has experienced over 40 years of high economic growth. Beginning from Soeharto's rise to power in 1967, Indonesia's Ministry of National Development Planning (*Bappenas*) outlined their plans for national development and the reduction of poverty in a series of Five-Year Development Plans (*Repelita*). One important part of these plans included the establishment of the "presidential instructions" (INPRES) program, which set up a system for distributing revenues from the central government to lower

[8] Most recent education research focuses on evaluating demand-side interventions. These include information campaigns (see Jensen, 2010 providing information to parents about the returns to schooling), cash transfers (see Fiszbein et al. 2009 for an overview; Behrman, Parker, Todd, 2011 for evidence on the five-year impacts of the Mexican conditional cash transfer program Progresa; and Parker and Vogl, 2017 for evidence on Progresa's longer-term impacts; Baird et al., 2011, Akresh, de Walque, and Kazianga, 2013, 2016, and Benhassine et al., 2015 for research that explores the role of conditionality in these cash transfer programs), scholarships (see Kremer, Miguel, and Thornton, 2009 examining the impact of merit-based scholarships), or other household level interventions (see Oster and Thornton, 2011 for evidence on providing female sanitary products to secondary school girls and Muralidharan and Prakash, 2017 for evidence from providing bicycles to families).

administrative levels. Starting with the oil boom in 1973, the central government emphasized the explicit goal of reducing regional disparities (Ravallion, 1988).

As part of this redistribution goal, the government began a nationwide school construction program, the Sekolah Dasar INPRES, which was one of the first and largest INPRES programs. Between 1973 and 1979, around 61,800 primary schools were constructed. Enrollment rates in 1972 before the start of school construction were 71 percent among primary school-age children. By 1978, enrollment rates among this age group reached 85 percent.[9] Prior to this program in 1973, capital expenditures in education were low and enrollment rates in the few years before school construction began were stagnant (World Bank, 1989).

School construction nearly doubled the stock of primary schools from a baseline of around 63,000 primary schools. On average, the program added over 200 schools per district or 2 schools for every 1,000 children of primary school age. There was large heterogeneity across districts in how many schools the government built as the government designed the school construction program to target districts in which enrollment was initially lower.[10] The government designed each school for 120 students, and they recruited teachers and paid their salaries for these newly constructed schools. During the same period, the government attempted to train new teachers, and the percentage of teachers who met the minimum qualification of having an upper secondary school degree did not change over this period (World Bank, 1989).

[9] World Bank Databank. 2018. “Adjusted Net Enrollment Rate, Primary (% of Primary School Age Children)” (Accessed on October 17, 2018: databank.worldbank.org)

[10] Appendix Figure A.1 presents a map of Indonesia indicating the geographical distribution of the number of schools constructed in each district.

3. Empirical Strategy and Data

3.1 Difference-in-differences

Following Duflo (2001), we estimate a difference-in-differences specification in which an individual's district of birth and date of birth jointly determine their exposure to the INPRES school construction program. Children in Indonesia typically attend primary school between age seven and twelve. INPRES school construction started during the 1973-1974 school year, so children who were born in or before 1962 were at least 12 years of age in 1974 and would not have benefited from the school construction.[11] Children younger than seven in 1974 would have been exposed to the full potential benefits of the newly constructed schools. Children who were of primary school age in 1974 might partially benefit from the new INPRES schools as some of them were induced to enroll, and their propensity to enroll likely decreased with the child's age.

Besides variation across birth cohorts, there is considerable variation across geographical districts in the intensity of the school construction program. This is because the number of schools constructed was linked to the districts' primary school enrollment rate in 1972 (prior to the school construction), and areas with low prior enrollment rates had more schools built.

Exploiting these two sources of variation (birth cohort and geographical), we estimate the effect of school construction in the following regression:

$$y_{ijt} = \alpha + \beta\, School_j \cdot Young_{it} + \left(\boldsymbol{X}_j \boldsymbol{B}_t'\right) \gamma_t + \mu_j + \delta_t + \varepsilon_{ijt} \qquad (1)$$

where y_{ijt} is the outcome of individual *i* born in district j in year *t*, $School_j$ measures the number of schools constructed by the INPRES program between 1973 and 1979 per 1,000 children in the individual's birth district j. We use an individual's birth district instead of current residence

[11] The 1993 Indonesian Family Life survey indicates that less than 3 percent of individuals born between 1950 and 1962 were still in primary school in 1974. As a further check, we use the 1976 Intercensal Survey and find that only 4.3 percent of individuals born between 1950 and 1962 were still in primary school in 1976.

because the latter may be endogenous to program placement if households move to access schools for their children.[12] $Young_{it}$ is an indicator variable for being born between 1968-1972 (ages 2-6 in 1974) and thus being young enough to benefit from the program. Individuals born between 1957 and 1962 (ages 12-17 in 1974) represent older cohorts who are not exposed to the program. Following Duflo (2001), we exclude individuals born between 1963 and 1967 (ages 7-11 in 1974) as they might have only partially benefited from school construction. We perform several robustness checks to confirm results are consistent across various definitions of exposed and unexposed cohorts.[13] μ_j are time-invariant district of birth fixed effects, δ_t are cohort of birth fixed effects, and $\boldsymbol{X}_j\boldsymbol{B}_t'$ controls for district-specific time-varying trends that might influence outcomes. Following Duflo (2001), we do this by interacting birth cohort indicators with district enrollment in 1971 and with the presence of water and sanitation programs in the district.[14] Note that we closely follow Duflo (2001) with the only exceptions that, unlike Duflo, we cluster our standard errors, and we do so at the district level, and that our data allows us to estimate the effects of school construction on both men and women. To allow for gender heterogeneity, we estimate Equation (1) separately for men and women.

[12] In the African context, child fostering, where the biological parents send their own child to live with another family, is quite common and often done to send the child to school (Akresh, 2004, 2009). In the Indonesian context, child fostering is much less common (Marazyan, 2012).

[13] One of the reasons for Duflo (2001) to restrict the young cohort to those born before 1972 is so that these cohorts would have completed schooling and begun participating in the labor market by 1995, the survey year of the data she uses. The 1972 cohorts turned 23 in 1995, which is old enough to have completed tertiary education. With our 2016 data, cohorts born after 1972 would have also been exposed to the school construction and had sufficient time to complete school and join the labor market. In the robustness checks discussed in Section 6, we explore the robustness of the results to alternative cohort definitions. In particular, we show that results are robust to adding in younger cohorts (born 1973-1980), older cohorts (born 1950-1956), and partially exposed cohorts (born 1963-1967).

[14] We use district enrollment in 1971 because program intensity was tied to 1972 district enrollment and not controlling for pre-program enrollment might bias the results as there could be mean reversion even in the absence of the program. In addition, the oil boom, which provided financial resources for school construction, could have also provided resources for other government programs that were correlated with INPRES school placement. Water and sanitation programs were the second largest set of INPRES programs delivered by the central government.

Given the program targeted less developed areas, we want to show that our effects are not explained by general catching up (or mean reversion) from those areas, as this would violate the parallel trends assumption. To test this identification assumption, we estimate placebo regressions in which we compare the old cohort (ages 12-17 in 1974) with an even older cohort (ages 18-24 in 1974). We discuss results in Section 6 and show there are no differential time trends in outcomes prior to the school construction. Further, in Appendix Figure A.2, we estimate a regression where we interact the number of INPRES schools constructed in one's birth district with an indicator for age in 1974. We omit the age group 19-24 in 1974 from the regression so that we are comparing each age against this older cohort. For both men and women, we do not observe any differential trend effects for non-exposed ages.[15, 16]

We explore individual and household-level outcomes to measure impacts of school construction. For household level data, such as expenditures, we use the household head or spouse's birth cohort and district of birth and present results separately for men and women.[17] In Equation (1), *j* refers to the man or woman's birth district, while *t* refers to birth year.[18]

The duration between school construction that started in 1973 and data collection in 2016 allows us to study not only the long-term effects of exposure to the program but also the effects of school construction on the next generation's outcomes. Specifically, we estimate the impact

---

15 In our main specifications, we prefer to show results using the regression specified in Equation (1) over figures such as Appendix Figure A.2 because this allows us to pool together all ages that belong to the young cohort and all those that belong to the old cohort, given that within cohort, everyone has the same exposure status.
16 Appendix Figures A.3 and A.4 provide further evidence for the validity of the difference-in-differences specification. We present graphs for each index used in Figure 1 that highlight the parallel trends across birth cohorts in the high and low intensity districts.
17 Female household heads represent 13.8 percent of the sample and are included in the regressions for women. Results are robust to estimating the regressions separately for household heads and spouses, irrespective of gender. Note that in the household level regressions the analysis for men does not condition on the women's exposure to school construction. Likewise, for women, the analysis does not condition on the men's exposure.
18 This assignment is arguably the most natural way to define exposure for household-level outcomes as there could be multiple individuals living in a household and these individuals could be in the old, young, and intermediate birth cohorts. Robustness checks in Section 6 show that this assignment decision does not influence results.

on children's schooling and other child outcomes based on whether their mother or father (or both) are exposed to school construction. We estimate reduced-form relationships between second-generation outcomes and schools construction in the following regression:

$$y_{ijtca} = \alpha + \beta\ School_j \cdot Young_{it} + \left(\boldsymbol{X}_j \boldsymbol{B}'_t\right) \gamma_t + \mu_j + \delta_t + \theta_a + \varepsilon_{ijtca} \qquad (2)$$

where y_{ijtca} denotes the outcome of child c who is age a, born to a parent i who was born in district j in year t, $School_j$ is the number of schools constructed in the father's or mother's birth district, $Young_{it}$ indicates if the father or mother belongs to the young cohort, and θ_a is child c's age fixed effect.[19] Standard errors are clustered at the father or mother's birth district.[20]

3.2 Strategies to address the large number of outcomes

We adopt two strategies to address the large number of outcomes that we examine to avoid overemphasizing any single significant result. First, as mentioned in the introduction, we create indexes for each family of outcomes following Kling, Liebman, and Katz (2007). These indexes combine all of the outcomes in each family of outcomes. To construct the indexes, we define each outcome so that higher values correspond with better outcomes. Then we standardize each outcome into a Z-score by subtracting the mean and dividing by the standard deviation of the older cohort born in low intensity districts. We then average all of the Z-scores and standardize the average relative to the older cohort born in the low intensity districts.[21] We then estimate the effect of exposure to school construction on these standardized outcome indexes.

---

19 We include child age fixed effects because old cohort parents mechanically have older children on average than young cohort parents and older children have more time to complete additional schooling. Therefore, the marginal benefit to children's years of schooling is estimated across different households but among children of the same age.
20 As in most household surveys, Susenas 2016 identifies household relationships with respect to the household head. If a child is not the biological or adopted child of the household head and spouse, the child is recorded as 'other household member'. Therefore, we restrict our intergenerational analysis to children of the household head and spouse. With further assumptions, we can include grandchildren and parents of the household head in multi-generational households. These add less than 3,000 additional observations and results are robust to their inclusion.
21 Banerjee et al. (2015) also use this approach to evaluate the effect of poverty graduation programs in six countries on a range of outcomes. Ajayi and Ross (2017) modify this standardization approach to use with a difference-in-differences empirical strategy when there is not a randomly assigned control group.

Second, we correct for the potential issue of simultaneous inference using multiple hypothesis testing. Following Benjamini and Hochberg (1995), we use the concept of a false discovery rate (FDR) to allow inference when conducting many tests. Intuitively, FDR allows a researcher to tolerate a certain number of tests to be incorrectly discovered. An FDR adjusted q-value of 0.05 implies that 5 percent of significant tests result in false positives compared with an unadjusted p-value of 0.05 that implies 5 percent of all tests result in false positives. In all regression tables, we show standard errors based on unadjusted p-values and FDR adjusted q-values that address the multiple hypotheses being tested in a given family of outcomes.

3.3 Data

To measure the impact of this school construction program, we use Duflo's data of the Sekolah Dasar INPRES program that reports the number of schools constructed in each district between 1973 and 1979.[22] We combine the school construction data with the National Socioeconomic Survey conducted in 2016, henceforth Susenas 2016, administered by Indonesia's Central Statistics Bureau, Badan Pusat Statistik. Susenas 2016 is a nationally representative household survey covering all 34 provinces and 511 districts of Indonesia.[23] The data combines a large sample size of 291,414 households and 1,048,575 individuals with a wide range of variables on education, employment, migration, living standards, taxes, assets, nutrition, health, marriage and demographics, welfare program participation, and the next generation's educational outcomes.[24]

[22] We are grateful to Esther Duflo for sharing these data.
[23] The smallest geographical unit in the Susenas 2016 is the Indonesian '*kabupaten*', loosely translated as district.
[24] Susenas 2016 is particularly suited to study the effects of this program because it includes information on the individual's district of birth and the sample is sufficiently large to precisely estimate the observed relationships. Appendix B provides further rationale for the choice of data, in particular showing that the sample for the Indonesia Family Life Survey (IFLS) is not large enough to detect the effects of school construction. Furthermore, Bharati, Chin, and Jung (2018) use the most recent round of the IFLS to examine the combined impacts of rainfall at birth and school construction. They find strong effects of rainfall in the expected direction, but the IFLS is underpowered to estimate the main effect of school construction on education or the interaction effect of construction and rainfall.

We present summary statistics by families of outcomes in Tables 1 to 5 and Appendix Table A.1, which we discuss in the next section together with the impacts of school construction. For the birth cohorts that our analysis focuses on (1957-1962 and 1968-1972), households have on average four members and the sample is evenly split between men and women. Average years of schooling for individuals in these cohorts is 8.0 years for men and 7.1 years for women. 81 percent of men and 73 percent of women complete primary school. These individuals have lower rates of lower and upper secondary school completion (39 and 34 percent for men respectively and 31 and 26 percent respectively for women). Tertiary completion rates are 9.5 and 7.7 percent for men and women. These individuals are ages 44 to 48 (young cohort) and 54 to 59 (old cohort) at the time of the survey in 2016. Most men are working (95 percent), while women have lower labor force attachment (64 percent). Conditional on working, only 33 percent of men and 24 percent of women are in the formal sector. Just over half of men and women work in non-agricultural sectors and around one-quarter have migrated from their birth district.

4. First Generation Effects of School Construction

This section describes the impact of school construction on first generation outcomes. Following the estimation strategy previously outlined, the main explanatory variable is an interaction of the number of schools constructed per 1,000 children in a person's birth district with an indicator for being young enough to benefit from the program. As discussed in the introduction, Figure 1 reveals positive impacts of school construction across indexes of outcomes for first generation individuals exposed to the program and for their children. In the following sections, we discuss the family of outcomes that these indexes are based on in more detail.

4.1. Impact on educational attainment

Table 1 studies the relationship between school construction and educational attainment.[25] On average, the program increases years of education for men by 0.27 years and for women by 0.23 years. The analysis by Duflo (2001) is restricted to men, and the comparable point estimate in her study is slightly lower at 0.19 years.[26] Note that both estimates are modest in size given that the number of primary schools almost doubled. At the mean number of schools built per 1,000 children (1.98), our estimates imply an increase in years of schooling of 0.53 and 0.46 for men and women, respectively.

The next four rows examine completed levels of education and show considerable gender differences. For men, the program increases the likelihood of completing primary school by 2.6 percentage points. Even though the program only targeted primary schools, we see effects for men in lower and upper secondary education at 2.3 and 2.6 percentage points. These represent larger percentage increases than for primary school because average completion rates at these levels are lower. On the other hand, the results for women are concentrated in primary school only, which they are 4.1 percentage points more likely to complete, and we are able to reject the equality of this coefficient with the male effect. The effects on lower and upper secondary completion rates are considerably smaller and indistinguishable from zero. For both men and

---

[25] The survey records educational outcomes for household members aged five and older and is missing otherwise.

[26] If we restrict our sample to only wage earners, our point estimate for men is 0.18, while for this subgroup, Duflo's estimate is higher than ours at 0.26. The comparable estimate we obtain for female wage earners is higher at 0.28, perhaps because it is less common for women to be engaged in wage work so they are less representative of women in general. While it is hard to pin down the exact reason for differences between Duflo's and our estimates given that we use different surveys (Intercensal survey versus Susenas) in different years (1995 versus 2016), two natural reasons for differential attrition come to mind. First, there may be differential mortality. In particular, if higher educated people live longer and are therefore more likely to still be in our 2016 dataset, then an increase in education caused by the program would lead us to find larger estimates in 2016 than in 1995. Conversely, if higher educated people are more likely to migrate internationally outside our sample, and therefore not be present during the 2016 data collection that would lead us to find smaller estimates in 2016. While we cannot directly test for these opposing trends, one way to reconcile the differences between our estimates and those of Duflo is that differential mortality may play a large role for the general population, while differential migration could be more prevalent amongst wage earners.

women, school construction did not affect tertiary education completion rates. As shown in row 6, literacy rates are high on average at 95 percent for men and 91 percent for women, and the program raises these by 1.5 and 3.3 percentage points, respectively. The FDR q-values (shown in square brackets) that correct for multiple hypothesis testing across all of the outcomes in the education table reveal that the coefficients remain statistically significant.

In the last row, we create an index using all other rows combined, following Kling et al. (2007) as discussed previously. The point estimates correspond with those shown in Figure 1 and confirm broad increases in educational attainment for men and women. Building two additional schools in an individual's birth district increases the educational outcomes for those exposed to the school construction by approximately 0.13 standard deviations relative to the control group.

We explore the gender dynamics and patterns by grade in further detail in Figure 2 showing the impact of school construction on the likelihood of completing at least a certain number of years of education. For example, it shows that the program increased the likelihood of completing at least one year of school by 0.95 percentage points for men and 2.3 percentage points for women. For men and women, effects are significantly different from zero throughout all primary school years and show an increasing pattern by grade. Consistent with Table 1, effects for men continue throughout lower and upper secondary school and seem stable across grades. While positive, the secondary school effects for women are not distinguishable from zero, nor are the effects on tertiary education for either gender.[27]

[27] While we could use the school construction program as an instrument for years of education, we prefer to study later-life outcomes using OLS in order to capture broad impacts and because the exclusion restriction could be violated if the program caused community-level changes that affect long-term outcomes in ways other than through increased schooling. There is a strong first stage relationship with an F-statistic of 32.3 for men and 31.8 for women. That said, for scaling purposes, the coefficients on long-term outcomes could be multiplied by four to calculate the effect of an extra year of education, given that the program increased schooling by approximately 0.25 years.

4.2. Long-run labor market impacts

Having observed large increases in education in response to school construction, Table 2 studies subsequent labor market and migration outcomes.[28, 29] As shown in row 1, 95 percent of men are working and the program raises this by 0.6 percentage points. The effect for women is half as large and insignificant, but allows for an economically meaningful increase within its confidence bounds, especially considering a lower average employment rate of 64 percent. In row 2, we explore the intensive margin of employment, namely number of hours worked conditional on working. Estimates indicate increases of 0.26 hours for men and 0.16 for women, but neither are significant. In response to school construction, men move to jobs that are generally deemed more desirable: they are 1.1 percentage points more likely to work in the formal sector that tends to offer higher quality and more stable jobs (Kleemans and Magruder, 2018). Given an average formal sector employment rate of 33 percent for men, increasing the number of schools in an individual's birth district by the sample mean raises the likelihood of men being in the formal sector by almost 7 percent. Men furthermore move away from agricultural work, which they are 1.2 percentage points less likely to hold, compared to 44 percent on average, and shift towards service sectors. We do not find any evidence of occupational shifts for women.[30]

Most research on the relationship between education and migration generally focuses on the selection into migration in terms of educational attainment.[31] We extend this work by

---

[28] Employment outcomes are recorded for household members aged ten and older and are missing otherwise.

[29] Heckman, Humphries, and Veramendi (2018) provide a recent overview of the extensive literature examining the relationship between education and labor market outcomes. Wydick, Glewwe and Rutledge (2013) find that child sponsorship leads to increased probability and quality of employment in six developing countries. Duflo, Dupas, and Kremer (2017) focuses on Ghana and is one of the few randomized control trials in education that follows individuals over eight years and finds that secondary school scholarships improves labor market outcomes.

[30] The only occupation-related variable that shows up as statistically significant for women is whether they are self-employed in their own micro-enterprise, which almost a quarter of women are. They are 1.1 percentage points more likely to do so. Given the large number of outcomes variables, we decided not to report all subcategories separately, but instead combine them into the 'formal worker' variable, which is not statistically significant.

[31] Empirical evidence for Indonesia (Hicks et al., 2018) and for developing countries in general (Young, 2013) shows positive selection from rural to urban areas and negative selection from urban to rural.

studying the causal relationship of increases in education on the likelihood of migration. On average, 27 percent of men and 25 percent of women migrate away from their birth district. School construction increases migration rates by 0.7 and 0.8 percentage points respectively. At the mean level of school construction, this represents an economically meaningful increase of 5.1 and 6.5 percent for men and women, respectively. Row 7 indicates that the increase in migration is concentrated in shorter distance moves within—rather than between—provinces.

Correcting for multiple hypothesis testing across all of the outcomes in the work/migration table shows that the FDR q-values are somewhat larger but coefficients generally remain statistically significant. Finally, aggregating the seven outcomes into an index following Kling et al. (2007) shows a positive and significant impact for men with an increase of 0.076 standard deviations due to an increase of two additional schools built in the district.

4.3. Long-term impacts on living standards

Susenas 2016 collects detailed data on household expenditures, which we use as a proxy for living standards.[32] Table 3 shows the effects of exposure to school construction on five aggregated living standard measures and on total tax payments. Row 1 shows that total expenditure increases by 2.1 or 3.2 percent in households in which either males or females are exposed. The increase is larger for non-food than food expenditure (rows 2 and 3) and, as a result, the ratio of non-food to total increases (row 4). Households where the head or spouse is

---

[32] Susenas 2016 does not include information on income, unlike the 1995 Intercensal survey that Duflo (2001) used to measure the returns to education. After the 1995 round, the earnings question was discontinued so we do not have access to more recent income data. That said, in a developing country setting with substantial home production, informal employment, unemployment, and underemployment, expenditure data may suffer less from under-reporting and more closely resemble living standards (Rizky, Suryadarma, and Suryahadi, 2018). All expenditure values refer to average monthly expenditure measured in 10,000 Indonesian rupiah (IDR). In 2016, the exchange rate was 1 USD=13,308 IDR. We convert expenditure categories reported in weekly or annual amounts to monthly expenditure. In regression analyses, we apply an inverse hyperbolic sine transformation to the nominal values since expenditure data tends to be skewed and a log transformation would not be defined for zero expenditures. The inverse hyperbolic sine is approximately equal to log (2y) or log (2) + log (y), so in most cases it can be interpreted the same way as a standard logarithmic dependent variable.

exposed spend 16 to 19 percent more on education (row 5). Note that the data are collected in 2016, so increased educational expenditures likely benefit the children of those exposed and may be a source of educational spillovers to the next generation; we return to this issue in Section 5. All results remain statistically significant after correcting for multiple hypothesis testing. The last row combines the expenditure data from rows 1, 2, and 5 into a living standards index, showing an overall increase of 0.03 and 0.05 standard deviations for men and women, respectively, for each additional school built in one's birth district.[33], [34]

In addition to expenditures, we study tax payments, which are important for the cost-benefit analyses in Section 7 and allow us to study if the program pays for itself from increased taxes over time. Row 6 of Table 3 shows that total taxes increase by 7.8 percent in households where males are exposed to school construction and 12.3 percent if women are exposed. In Appendix Table A.1, Panel A (rows 2, 3, and 4), we analyze the three main sub-components of taxes, revealing increases in land and building taxes, taxes on motorized and non-motorized vehicles, and local community taxes.[35]

---

[33] The variables in rows 3 and 4 of Table 3 can be derived from those in rows 1 and 2, so to avoid double counting these are excluded from the living standards index. We later discuss the variable in row 6, which is included in the tax index.

[34] As income is generally correlated with expenditure, we expect improvements in labor market outcomes to go hand-in-hand with expenditure increases. While this is true for men, we observe increased expenditure in the absence of labor market improvements for women. We explore this apparent puzzle in Appendix Table A.2. Note that labor market outcomes are observed for each individual while expenditure is measured at the household level, so household dynamics and spouse characteristics can interact with the direct effect of program exposure. Column 1 repeats the total expenditure regression and the other columns add spouse characteristics as mediating variables. Controlling for spouse characteristics changes the direct effect of program exposure on expenditure: the effect of men's exposure drops significantly when controlling for his spouse's education-related variables but not when controlling for the wife's labor market outcomes. This is not surprising, as we did not observe labor market effects for women. The effect for exposed women drops 30 to 50 percent not only when controlling for her husband's education but when controlling for his labor market outcomes. When controlling for all mediating variables in the last column, the direct effect falls by 60 percent for both men and women. We interpret this as suggestive evidence that household dynamics and marriage market outcomes play an important role; an issue we return to in Section 4.4.

[35] Panel B of Appendix Table A.1 explores effects on housing and assets. On average 43 percent of the sample lives in urban areas and even though exposure increases migration, school construction does not increase the likelihood of living in urban areas. They do appear to move to more valuable and larger housing. Row 2 shows an increase of 2.8 percent in the monthly rent equivalent payments if women are exposed, and a smaller and insignificant effect if males are exposed. Exposure increases floor area by 1.2-1.5 square meters (row 3) and utility usage by 5.1 to 8.5

Another component of living standards concerns people's health and nutrition. While there exists a strong correlation between education and health, research estimating a causal relationship finds mixed evidence.[36] Panels C, D and E of Appendix Table A.1 show three components of health impacts in response to school construction. Panel C shows overall calories increase by 1.8 percent for women exposed to the program while the effect for men is smaller and not significant.[37] Health investments also improve as shown in Panel D. Health expenditures increase by 5.5 and 7.1 percent for exposed women and men, but only the men's coefficient is significant. Using detailed data from various types of health investments, the health investment index shows broad increases for exposed men and women, with an improvement in health investments of 0.13 standard deviations when an additional two schools are built in the individual's birth district.[38] While there are increases in nutrition and health investments, Panel E reveals that overall we do not observe improvements in health. While we see increases in not reporting a health complaint in the last month (row 1) and the number of days uninterrupted by health complaints (row 2), neither are statistically significant. Considering severe health

---

percent (row 4). To approximate for household wealth, row 5 studies the impact of school construction on an asset index that is created using a principal component analysis of household ownership of the following durable goods: ownership of motorcycle, car, home phone, computer/laptop, television, gold/jewelry, refrigerator, water heater, LPG gas tube, boat, motorized boat, and air conditioner. The program leads to a 3-4 percent increase in the index if men or women are exposed. Aggregating all five housing and asset outcomes into an index in row 6 confirms broad increases for men and women.

[36] Lleras-Muney (2005) finds positive effects of education on mortality in the U.S., while Clark and Royer (2013), Malamud, Mitrut, and Pop-Eleches (2018), and Meghir, Palme, and Simeonova (2018) find no effects of education on mortality in the U.K., Romania, and Sweden, respectively. Baird at al. (2016) show that improved health from deworming treatment at primary schools in Kenya increases education, as well as later-life labor market outcomes.

[37] Patterns are similar for consumption of protein, fats, and carbohydrates with respective increases of 1.8, 2.3, and 1.7 percent when women are exposed to the program and smaller increases for men. We are unable to answer definitively (although Panel E examines self-reported health outcomes) whether these nutrition changes for women are health-improving as additional protein is likely beneficial for individuals in developing countries, but additional fats can indicate a worse diet.

[38] Preventative health investments that include medical check-ups, family planning, and immunizations increase by 19-24 percent if the woman or man is exposed to school construction. We see large increases in expenditures for family planning, including contraceptives and consultations, of 23 and 32 percent for women and men, respectively. On the curative side, households with either the man or woman exposed are 4.8-7.5 percent more likely to use a private instead of public hospital, which generally provide higher quality and more expensive health care. Health insurance expenditures rise 14 percent for exposed women.

complaints only, we observe a 0.5 percentage point decrease in reports from exposed men. The aggregated health index in row 4 shows an improvement in health outcomes for men exposed to school construction but is insignificant for exposed women.[39]

4.4. Long-run marriage and fertility effects

Evidence estimating the causal relationship between education and demographic outcomes has generally been mixed and nuanced, and this is true in our setting.[40] Table 4 explores marriage and fertility outcomes for those exposed to the program.[41] On average, women marry four years younger than men (row 1), but there is no effect of school construction exposure on the age of first marriage. Coefficients are small and statistically insignificant. In rows 2 to 9, we explore if program exposure changed the types of spouses that men and women marry. Rows 2, 3, and 4 reveal that both men and women marry partners with higher levels of education and literacy. Results remain statistically significant after correcting for multiple hypothesis testing. The results are perhaps unsurprising given that average education levels improved in districts with additional schools, but important nonetheless because education increases for exposed individuals as well as their spouses. In Appendix Table A.3, we examine further if this education increase is only a level effect (everyone has more education) or whether there is a distributional effect (husbands

---

39 We also explore the effects of exposure to school construction on the first generation's utilization of government welfare programs. Susenas 2016 collects data on four national programs that aim to reduce poverty and inequality. Ex-ante it is unclear if increased take-up of welfare programs reflects higher needs due to increased poverty or whether it is indicative of increased awareness of existing programs. As shown in Panel F of Appendix Table A.1, we see few changes to any of the individual government programs or to the welfare program index in the last row.

40 Osili and Long (2008) find evidence of education reducing fertility in Nigeria. On the other hand, McCrary and Royer (2011) find only a small fertility effect but a larger effect on the quality of the marriage partner in the U.S. In the Kenyan context, education subsidies reduce women's likelihood of teenage marriage and pregnancy (Duflo, Dupas, and Kremer, 2015). Looking at a wider age range, Geruso and Royer (2018) find increased education lowers teen fertility and increases the education of the spouse but has no impact on total completed fertility. Research by Breierova and Duflo (2004) examining this same Indonesian school construction program find that women exposed to the program have lower fertility, consistent with our results discussed below.

41 Ashraf et al. (forthcoming) also examine the INPRES school construction program and the relationship between education and marriage. They find that marriage market customs, in particular the practice of bride price, strongly influences how much education women receive. School construction has large positive effects on education for women but only among ethnic groups that practice bride price.

marry women from higher in the education distribution). There is weak evidence that spouses come from higher in the education distribution, although the results do not survive multiple hypothesis testing (row 2). We see stronger evidence that women exposed to school construction reduce the education gap (in levels and percentiles) with their husbands (rows 3 and 4). This highlights the increased average education in the household, which may play an important role for the intergenerational transmission of education, which we explore in Section 5. As shown in rows 5, 6, and 7, we do not observe any changes in the labor market outcomes of spouses, but we do see that men and women are 0.7-0.8 percentage points more likely to have a spouse who migrated (row 8).[42] We do not observe changes in spouse's health (row 9).

We do not have complete birth histories for each woman that would allow us to measure the relationship between education and total fertility. However, we can measure impacts on the number of children aged 0-14 living in the household at the time of the survey. Exposure to school construction reduces the number of children for women. The combined marriage and fertility index (row 10) shows broad improvements for both genders: two additional schools built in a district lead to increases of 0.13 and 0.10 standard deviations for exposed men and women.

5. Second-generation Effects of School Construction

Having observed large long-term effects of Indonesia's school construction program on a wide range of outcomes, including education, employment, migration, living standards, and marriage, we now investigate whether the effects extend to the next generation and affect the children of those exposed to the program. To guide our analyses, we develop a simple conceptual framework building on Behrman and Rosenzweig (2002) who study and compare the effects of mother's and

[42] We define migration relative to the individual's birth district, and it may have occurred anytime between birth and 2016, so spouses may have migrated together.

father's schooling on their children's schooling. In its most general form, consider a linear reduced-form equation of the schooling of child i in family j:

$$S_{ij}^{c} = \delta_1 S_j^f + \delta_2 S_j^m + \gamma_1 X_j^f + \gamma_2 X_j^m + \varepsilon_{ij}^{c} \tag{3}$$

where superscript c denotes that individual i is a child in family j; S_j^f is schooling of the father, and S_j^m is schooling of the mother. X_j^f refers to father characteristics other than schooling that may affect S_{ij}^c, including labor market conditions, migration status, and health outcomes. X_j^m refers to such characteristics of the mother and ε^c is a child specific term.

As Behrman and Rosenzweig (2002) discuss, Equation 3 is a general and reduced-form model of household resource allocations that is consistent with many models, both dynamic and static. It emphasizes the interrelationship of parent characteristics affecting child schooling. In particular, the characteristics of the parents will be correlated with each other due to nonrandom matching in the marriage market. Dropping subscripts i and j for convenience, this is illustrated in the following assortative matching relationships:

$$S^f = f(S^m, X^m)$$

$$X^f = g(S^m, X^m)$$

Combined with Equation 3, these relationships highlight that any empirical association between parent and child schooling need not reflect a direct relationship only, but may also manifest itself through positive assortative matching.[43] Indeed, Table 4 and Appendix Table A.3 (discussed in the previous section) show that individuals exposed to school construction match with higher educated spouses, both in absolute terms and relative to others on the marriage market.[44]

[43] Behrman and Rosenzweig (2002) caution that studying intergenerational schooling externalities is also challenging due to ability bias, which points to parent ability as an omitted variable determining both their own schooling and their child's ability and through that, the child's schooling. Studying effects in response to school construction helps eliminate this concern and obtain causal estimates of second-generation effects.
[44] On average, the correlation in years of education between spouses is high at 0.63.

We explore these interrelationships in the remainder of this section. In Section 5.1 and 5.2, we study the impact of parental exposure to school construction on their children's schooling. In Section 5.3, we perform a mediation analysis using the rich data we have available to see how controlling for the additional terms of Equation 3 changes the effect of parental exposure on their children's' schooling. We perform this analysis first for a person's own characteristics, other than schooling, to see if the effect of parental exposure on child education manifests itself through characteristics such as labor market outcomes, living standards, and nutrition. Then, we repeat this exercise by controlling for the spouse's schooling and characteristics. Following the conceptual framework we laid out above, this helps shed light on the mechanisms through which parental exposure affects child's schooling.

5.1. Second-generation effects on education and wellbeing

As explained in Section 3, we measure second-generation impacts using the same difference-in-differences framework as first generation effects. The main explanatory variable is an interaction of school construction intensity in a parent's birth district with an indicator for being young enough to benefit from the program. We consider outcomes for all children living in the parent's household and include age fixed effects to ensure comparisons are across same-aged children.

Table 5 shows the effect of parental exposure to school construction on their children's educational attainment.[45] Row 1 confirms that program effects persist into the next generation.[46] Children with fathers exposed to the program obtain an additional 0.10 years of school, while children with exposed mothers obtain 0.17 years more. We can reject the equality of these

[45] Even though school construction may affect both father's and mother's exposure, they enter each regression separately in Table 5. We return to this issue in Section 5.2 and in Table 6 where we control for exposure of both parents simultaneously.

[46] Related research explores the production function for children's human capital (Attanasio, Meghir, and Nix, 2017) as well as focuses on how parents or teachers respond to inequalities across children (Akresh et al., 2012, Pop-Eleches and Urquiola, 2013).

coefficients. The magnitudes decrease compared to the first generation results of 0.27 and 0.23 years for men and women, but are still economically meaningful. In the next section, we explore potential channels through which the effects persist into the next generation.

Unlike the first generation education results, we observe no effects on children for primary school completion rates (row 2) because by 2016 primary school has become nearly universal.[47] There are large effects on completing lower and upper secondary for children whose parents are exposed to school construction, with the effect for exposed mothers being statistically larger than for exposed fathers. In addition, unlike the first generation education results, increases in educational attainment now extend to tertiary education completion rates. Children with exposed mothers are 0.8 percentage points more likely to complete tertiary education, compared to a 0.4 percentage point increase for children with exposed fathers. An increase of the mean number of schools in a mother's birth district leads to a 25 percent increase in the likelihood her child completes tertiary education relative to average tertiary education levels.

To account for the fact that some second-generation children may still be attending school, we study the effects on age-for-grade (row 6), defined as an indicator variable for whether the child is on track to complete the appropriate grades on time.[48] Results confirm that having parents exposed to school construction increases the likelihood of being on track by 1.1 percentage points if the father is exposed and 1.8 percentage points if the mother is exposed.

All of the education results for mothers remain statistically significant after correcting for multiple hypothesis testing, while lower secondary and tertiary completion rates for children

[47] UNICEF statistics (accessed November 19, 2018: unicef.org/infobycountry/indonesia_statistics) indicate primary school net enrollment rates in Indonesia from 2008-2012 are 98 and 100 percent for boys and girls, respectively.
[48] The indicator variable is zero for those who did not start school by age 7 or had to repeat one or more grades before completing upper secondary education, which is compulsory in Indonesia. The indicator variable is one for those who have already completed upper secondary school or are on track to do so in a timely manner.

when the father is exposed are no longer statistically significant. We aggregate the six outcomes into a second-generation education index, and it shows broad increases for children with parents exposed to school construction. An increase of two additional schools built in the father or mother's birth district increases their children's educational attainment by 0.04 or 0.07 standard deviations, respectively, relative to parents who are not exposed to the program.[49]

5.2. Heterogeneity of second-generation results by gender and grade

We next explore two dimensions of heterogeneity in the second-generation education results. First, we examine if school construction had different second-generation effects at different grade levels. Second, we examine, within a household, if paternal or maternal exposure to school construction had differential impacts on their children and if those impacts differed by whether the child is a son or daughter.

In Figure 3, we estimate the likelihood of a second-generation child completing at least a certain number of years of school. We explore the effects depending on whether the father or mother is exposed to school construction and whether their child is a son or daughter. Results highlight that effects are small and indistinguishable from zero during primary school. Consistent with Table 5, for all other grades, exposure to school construction by mothers has a larger effect than fathers on their children's education. For grades in lower secondary, upper secondary, and tertiary, we observe effects that are significantly different from zero for daughters when either their mother or father is exposed to school construction. Effect sizes for second-generation

[49] Having observed educational increases for children with parents exposed to the program, Appendix Table A.1 Panel G explores effects on children's employment and self-reported health. Rows 1 and 2 examine the number of days and hours not engaged in work. We see reductions in days and hours worked if the father is exposed but no effect if the mother is exposed. Results are not statistically significant after correcting for multiple hypothesis testing. Rows 3, 4, and 5 study second-generation health effects. Mother's exposure does not affect children's health. Children whose fathers were exposed report worse health outcomes. We cannot determine if these children are less healthy or if their better-educated parents have an understanding of health that makes them more likely to report their child as ill. Aggregating these employment and health indicators into a second-generation wellbeing index shows no effect for children whose fathers are exposed and a positive impact if the child's mother is exposed.

daughters are approximately of the same magnitude as those of the first generation's men exposed to the program (see Figure 2 for this comparison). Effect sizes are largest for daughters when the mother is exposed to the program and lowest for sons when the father is exposed. When examining each grade separately, we cannot statistically distinguish the results by gender of the parent or child, but in Table 6, we investigate this issue in more detail.

Table 6 examines the intergenerational education effects when controlling for the partner's exposure to school construction, and we explore if the effects differ by child gender. We face several challenges with our identification strategy that focuses on young (born 1968-1972) and old (born 1957-1962) cohorts. For a household to be included in the regression, we need both the father and mother to be in these specific birth cohorts.[50] Given this selected sample, we observe that for these households the impact of mother's exposure to school construction has a much larger effect on the child than the father's exposure. An additional school built in the mother's birth district raises her child's education by 0.16 years; there is no effect if the father is exposed. In column 2, we address this selection issue by expanding the range of birth cohorts included in the regressions, now including all individuals born between 1950 and 1980. All birth cohorts born after 1968 could be exposed to the school construction that began in 1973. Both parents no longer need to be part of the young and old cohorts as previously defined. The sample size expands to 246,466 second-generation children with parents in this extended cohort range. Results are consistent, with mother's exposure to school construction

[50] With this additional sample restriction, there are only 44,105 second-generation children in the regression, a loss of almost two-thirds of the sample from the 120,838 children in the regressions in Table 5.

increasing her child's education more than father's exposure. We can reject the equality of coefficients in both the restricted (column 1) and extended birth cohort samples (column 2).[51]

We explore if parental exposure has different effects for sons and daughters. Panel A includes all children, Panel B sons, and Panel C daughters. Results for sons and daughters show consistently larger education effects for the second-generation child if the mother is exposed to school construction, and in all cases, mother exposure is statistically significant. Focusing on the extended cohort sample (column 2), we can reject the equality of the mother and father exposure coefficients in the case of daughters but not sons. The benefit to daughters is three times larger if their mother rather than their father is exposed to school construction.

5.3. Channels for intergenerational persistence of education

In line with the conceptual framework we laid out at the start of this section, we perform mediation analyses to understand the mechanisms linking first and second-generation outcomes. Results are shown in Tables 7 and 8. Column 1 in each table repeats the effects of parent exposure on children's years of schooling. Subsequent Table 7 columns add indexes from Figure 1 as control variables that may be mediators through which parental exposure manifests itself. Subsequent Table 8 columns add characteristics about the spouse as potential mediating factors. Naturally, the additional regressors are endogenous to the program and we are not able to control for all relevant characteristics, but given our rich data, we, nonetheless, believe a mediation analysis is insightful.

Column 2 in Table 7 shows that adding the father's work and migration index leads to an 18 percent reduction of the effect of father's exposure, but the effect remains large at 0.08

---

[51] We examine if there is an additional benefit when the father and mother are both exposed to school construction but do not find a significantly larger effect in those households (results not shown), highlighting that the intergenerational education effect is primarily driven by mother's exposure.

additional years of school. Controlling for the mother's work and migration index barely affects the coefficient on mother's exposure. This is not surprising as we find few labor market effects for women in response to school construction. A 23 percent drop occurs for women when we control for their living standards index (column 3), in line with Section 4.3 showing large increases in expenditure for women exposed to school construction. Despite this drop, children still get an additional 0.13 years of education even if we hold expenditure constant. Columns 4 and 5 control for the index of taxes and housing/assets. Point estimates remain similar as when we control for living standards. In columns 6, 7, and 8, we include controls for the parent's nutrition, health investments, and health outcomes. We do not find support for these factors as potential mechanisms mediating the effect of parent's exposure to school construction.

In column 9, we control for the individual's marriage index, which includes spouse characteristics as well as age of first marriage and household size. For fathers, this leads to a 42 percent drop in the direct effect of exposure on their children's education. This is consistent with our conceptual framework and the previous section's finding that mothers, whose characteristics are now partly controlled for, have a large impact on the intergenerational transmission of education. On the contrary, when we control for the marriage market index for exposed mothers, the drop is only 19 percent, and the direct effect of exposure on children's schooling remains.

In the last column, we include all indexes as control variables. The exposure coefficients drop to 0.055 for fathers exposed and 0.111 for mothers exposed. Holding constant many of the variables that are affected by school construction explains 43 and 34 percent of the effect on second-generation's years of schooling and the rest appears to be a more direct effect. We do not observe many channels in the data through which these effects could manifest themselves, for example through increased encouragement to attend school or study. We cannot distinguish between these channels, but based on this table, we conclude there remains a direct effect from parents to their children that is not explained by changes in own characteristics.

As highlighted in our conceptual framework, as well as by the marriage index's role as a mediating variable in Table 7, marriage plays a key role in explaining the intergenerational transfer of education. We explore this further in Table 8 by controlling for the spouse's schooling and other characteristics. The effect of father's exposure on second-generation's years of schooling drops significantly if we control for the mother's education. Controlling for whether she completes primary school reduces the estimate by half and makes it statistically insignificant. Controlling for the wife's years of schooling or literacy leads to reductions of 29 and 49 percent, respectively. This reaffirms the importance of mother's education for her children's education, as we saw in Section 5.2 and Table 6. When the mother is exposed, controlling for the father's education variables leads to a smaller but still sizable drop of 25–32 percent of the main effect.

Spouse's labor market outcomes explain part of the effect of exposure on children's years of schooling, especially whether the spouse has a formal sector job or works in a non-agricultural sector. The coefficient for father's exposure falls by 28–30 percent when controlling for the mother's sectoral choice, and the mother's coefficient drops by 23–25 percent. Controlling for spouse's migration status or health does not significantly alter the program exposure coefficients.

Lastly, we control for all spouse characteristics (column 10), and the main effect for fathers falls to 0.020–an 80 percent drop–and is not statistically significant. This suggests that for fathers most of the effect of exposure works through his spouse's characteristics. For mothers, the coefficient drops by almost half, but remains statistically significant and economically meaningful at 0.091 additional years of schooling. Overall, these analyses provide suggestive evidence that especially for men, marriage market outcomes and spouse characteristics are an important channel through which exposure to school construction increases children's schooling.

5.4 Selection of second-generation individuals

Two data issues are relevant for the second-generation analysis and the selection of individuals in the regressions. First, there is a tradeoff between the selection of who remains in a household, and is therefore in the survey, and what age they finish higher levels of school. Focusing only on younger children ages 0-15 reduces the selection bias as few of them leave the household by that age, but they are not old enough to complete higher levels of schooling, which are important to study since primary school is nearly universal in 2016. As we include older children, they have time to complete higher schooling levels but a larger percentage of them have left the household. In the second-generation analyses thus far, we include all children who still live with their parents, regardless of age. We, of course, include child age fixed effects. We illustrate the rationale for not imposing age restrictions in Appendix Figure A.5. In the top panel, we show the exposure coefficient for second-generation children's years of schooling if we limit the analysis to individuals under a certain age, and on the x-axis, we vary the upper bound of included ages. Given that lower levels of education are universal in 2016, we find no effect if we restrict to children age 0-15. As we include older ages, we show the sample size increase shown in the bottom panel, and effect sizes increase as individuals have time to complete their education.

Second, Susenas 2016, like most household surveys, only includes information on individuals currently residing in a given household but not on family members living elsewhere. For second-generation children who are no longer living with their parents, perhaps because they started a new household, we cannot link them to their biological parents, and we do not know if the parents of these children are exposed to the program. For this reason, we based all second-generation analyses thus far on those still living with their parents. We next explore robustness of the results under various assumptions about the children who have left the household.

To do this, we conduct three bounding analyses. First, we estimate extreme bounds in which we assume all non-co-resident children have parents who are or are not exposed (Manski, 1990). The intuition behind this is to re-assign individuals living apart from their parents back into the sample.[52] In Appendix Table A.4, we compare our baseline estimate for the second-generation's years of schooling (Table 5, row 1) to the first bounding strategy. Including all individuals in the survey under age 40 in the second-generation regression increases the sample from 120,838 and 105,523 in the father and mother's regressions to 644,675 observations. In regressions measuring the effect of father's exposure on his child's years of schooling, results are no longer statistically significant with these extreme bounds in which we assume all non-co-resident children are born to non-exposed fathers (column 2) and then all non-co-resident children are born to exposed fathers (column 3). The education effects for second-generation children whose mothers are exposed to school construction remain statistically significant despite these extreme assumptions. Assuming all non-co-resident children are born to mothers who are not exposed yields an effect of 0.05 more years of school, while assuming all non-co-resident children are born to mothers who are exposed yields an effect of 0.03.[53]

[52] For children living away from their parents, we must assume the parent's birth district and birth year to determine the parent's exposure. As we have no other information, we assume the parent's and child's birth district are the same. To test this assumption's robustness, we estimate Table 5 using a child's instead of parent's birth district, and results are consistent. In the main regressions, we include birth year fixed effects, but it is harder to predict parent's birth year given only a child's age, so we now include an indicator variable for whether the parent is in the young cohort. Estimating Table 5 replacing birth year dummies with a young cohort dummy yield consistent results. Further, to minimize the probability of including individuals who are unlikely to be children of a parent in our young or old cohorts, we impose an age restriction of 40 that would imply parents in the old cohort were at least 14-19 years old at the time of birth. We maintain the exposure status of children who still live with their parents.

[53] The effects using these extreme bounds are smaller than our Table 5 estimates. This is to be expected if parental exposure to school construction leads to an increase in their children's years of schooling. The reason for this is that the children we add to our sample are a combination of children whose parents are and are not exposed. In the case we assume all parents are exposed, some of these children actually had non-exposed parents and thus no increased educational attainment due to their parent's exposure, but we incorrectly assign them to the group of exposed

Aside from the extreme assumption that parents of non-co-resident children are either all exposed or all not exposed, we also likely add too many individuals to the regression. The second bounding exercise attempts to address these issues. The bounding regressions for second-generation children should only include children born to parents in the old (1957-1962) or young (1968-1972) cohorts. To improve our bounds, we use the Indonesia Family Life Survey (IFLS) that does well in tracking individuals over time and matching parents to children who moved away.[54] We use these data to obtain the fraction of children at each age born to old and young cohort parents among all children no longer living with their parents. We use these IFLS-based fractions to randomly assign at each age non-co-resident children in the Susenas data to either old or young cohort parents and exclude the rest from the regression.[55] We simulate this randomization assignment procedure 1,000 times and estimate the second-generation years of schooling regression. Appendix Figure A.6 shows the distribution of coefficients from these 1,000 repetitions for father and mother's exposure to school construction. Effect sizes for father's exposure on their children's years of schooling range from 0.011 to 0.047 at the 5th and 95th percentiles with a median coefficient of 0.028. Effect sizes for mother's exposure are larger, ranging between 0.018 and 0.065 with a median coefficient of 0.043.[56]

Third, we repeat the second-generation analysis directly using the IFLS. The last three columns in Appendix Table A.4 show estimates for all children, only those living with their

---

parents, which biases the estimates downwards. Similarly, if we assume no parents are exposed, some of these children actually had exposed parents so increased educational attainment, but we incorrectly assign them to the group of unexposed parents, which again leads to a downward bias of the estimates.

54 87.8 percent of individuals surveyed in the 1993 first wave were tracked or deceased by the last wave in 2014/5. We match 91 percent of children in the last wave's household roster to their co-resident or non-co-resident parents.

55 Results are consistent if we use the fraction of non-co-resident children at each age and gender who are born to old and young cohort parents and use these age-gender based fractions to draw random samples in the Susenas data.

56 Coefficients for father's exposure are statistically significant at the 10, 5, and 1 percent levels in 47, 27, and 5 percent of the regressions. Coefficients for mother's exposure are statistically significant more often. In 63, 42, and 13 percent of the regressions, the coefficients are statistically significant at the 10, 5, and 1 percent levels.

parents, and those who moved away. Those living with their parents provides the closest comparison to our Susenas sample. Across all three samples, we find no statistically significant effect for fathers exposed to school construction. The estimated effect of mother's exposure is 0.539 in the sample of children still living with their parents compared to 0.300 in the sample of all children. This suggests an effect only 56 percent as large if we are unable to include non-co-resident children in the analysis. Scaling down our estimate of mother's exposure in column 1 by this magnitude yields an estimated effect of 0.094 additional years of schooling, which is still economically meaningful and substantially higher than the coefficients in Appendix Figure A.6.

6. Threats to Identification and Robustness Checks

6.1 Possible general equilibrium effects

The analysis presented so far exploits variation across districts and cohorts to identify partial equilibrium school construction effects. This raises the concern that general equilibrium effects might undo direct program effects (Heckman, Lochner, and Taber, 1998). If school construction increases the young cohort's education in high intensity districts, this could affect non-exposed individuals (either young cohorts in low intensity districts or older cohorts). Depending on how these general equilibrium effects work and whether they have a negative or positive effect depends on the substitutability or complementarity between old and young cohorts.

School construction leads to more educated young workers. If those workers are substitutes for the older cohorts in the labor market, then this increase could drive down wages for the older cohorts who are competing with them for jobs in those locations. If that happens, then the effects we observe in our specification for improved living standards for the young relative to the old cohort would be over-estimating the true effect of the program. Duflo (2004) provides evidence that these general equilibrium effects might have occurred in Indonesia, although the magnitudes of the bias appear to be rather small. Focusing on the instrumental variables specification she estimates, she finds that an increase of 10 percentage points in the share of primary school graduates leads to a 2.9-3.8 percent decrease in wages for the old cohorts. Given we observe an increase of only 2.6 percentage points in the likelihood of completing primary school for men, the subsequent old cohort wage decreases would be less than 1 percent. Adjusting our estimates by that magnitude does not significantly alter our results. In addition, if we adjust by this magnitude the cost-benefit calculations discussed in Section 7, it would not affect our overall interpretation of the benefits of school construction.[57]

Alternatively, if the young cohorts are complements for the old cohorts (for instance, because they start more businesses and hire older cohort individuals or because they spend more money on goods and services produced by the older cohorts), then the older cohorts benefit by having more educated younger cohorts in their location. If these effects dominate the general equilibrium impact, we would underestimate the true effect.

Unfortunately, the data do not allow us to distinguish between these competing stories of complementarity and substitutability among old and young cohorts. Furthermore, the evidence on general equilibrium effects from developed countries (Angrist, 1995; Crepon et al., 2013; Bianchi, 2018) is unlikely to be helpful in understanding the developing country, Indonesian context over the past four decades. In our case, we can show there are no systematic trend breaks when comparing the old cohort with an even older cohort. While this is certainly not definitive, the results for household expenditures in Appendix Figure A.7 highlight that there is not a

[57] Recent research on large-scale government education investments in India finds that the general equilibrium effects could be larger, with these effects working to depress the returns to education by 32 percent (Khanna, 2018). However, the analysis of the Indian policy highlights that skilled workers are worse off while unskilled workers benefit. In the Indonesian context, this evidence about unskilled workers benefiting implies older cohorts who were more likely to be lower educated and unskilled would have benefited from these general equilibrium effects, which would lead us to underestimates the true effect.

differential trend when comparing the old cohort (ages 12-17 in 1974) with an even older cohort (ages 18-24 in 1974). If general equilibrium effects negatively impact older cohorts, we would expect to see the oldest cohorts be the worst off but that is not what we observe.[58]

6.2 Robustness checks

In this section, we present a set of specification checks highlighting the robustness of the main results. Our identification assumption is that the change in outcomes across birth cohorts in districts that built many schools would have been the same in the absence of the program as the change across birth cohorts in districts that built fewer schools. However, educational patterns between birth cohorts could vary systematically across districts because of issues such as mean reversion. To test this assumption, we estimate placebo regressions comparing old cohorts (ages 12-17 in 1974) and even older cohorts (ages 18-24 in 1974). If the assumption is correct, then any change in outcomes between cohorts in these groups, both of whom were not exposed to the program, should not differ across districts. Appendix Figure A.8 presents the results from estimating placebo regressions for each of the indexes for every family of outcomes (similar to Figure 1). Across the first generation outcomes for females and males, the placebo regressions show no statistically significant effects, which is suggestive evidence that the main difference-in-differences results are not driven by a failure of the identification assumption. We do observe a statistically significant effect in the placebo regressions for second-generation education. This implies there may have been a time trend across districts that could have influenced educational outcomes for second-generation children. However, if children whose parents born 1957-1962

[58] General equilibrium effects may be prevalent for other outcome variables as well. For example, part of the observed increase in taxes may be due to higher tax rates resulting from increased preference for redistribution, or greater demand for public goods by higher-educated people. Martinez-Bravo (2017) shows that the INPRES school construction program led to increases in the provision of public goods and suggests this is driven by increased education of village heads. We observe overall increases in tax payments but our data does not allow us to disentangle between these interpretations.

are experiencing more education compared with children whose parents are born 1950-1956, then that likely means we are underestimating the true effect.

In the main results, we follow Duflo (2001) by using a conservative definition of school exposure. Individuals in the young cohort (born 1968-1972) benefit from full exposure to the program, while those in the old cohort (born 1957-1962) did not benefit from school construction at all. However, there are other birth cohorts, partly exposed and not exposed, that could be included in the analysis. Appendix Table A.5 examines how the regression for first generation years of schooling in Table 1 changes with alternative birth cohort definitions. Column 1 repeats the original results and columns 2-5 add additional birth cohorts. In column 2, we add older cohorts born between 1950 and 1956 who were not exposed to school construction. Column 3 adds individuals born between 1963 and 1967 who would have been primary school aged in 1974 (ages 7-11) when the schools were built. To be conservative, we assume none of these cohorts was exposed, although in Appendix Figure A.2, it appears children ages 7-8 probably did benefit from the program. Column 4 includes children born between 1973 and 1980. They were born during and just after the schools were built, and so they would have received full exposure to the program.[59] Finally, in column 5, we include all individuals born between 1950 and 1980. Results using the different sample definitions are consistent, showing that exposure to school construction increased years of schooling for men and women. In Appendix Figure A.9, we use the extended cohort definition (all individuals born between 1950 and 1980) and re-estimate the effect of school construction on indexes for families of outcomes (as in Figure 1). Results are consistent, showing large positive benefits for men and women exposed to school construction.

[59] These Indonesian primary schools were initially expected to last for 20 years so the last cohort that could have gained the full six years of primary school education and completed primary schooling by 1993 would have been born in 1980. Note that with these 1973-1980 cohorts, parents could potentially have moved to give their children access to these schools, although results are consistent with the earlier ones.

We estimate all regressions using expenditure data with an inverse hyperbolic sine transformation (IHS). While this is typical to analyze expenditures, in Appendix Table A.6, we show robustness checks using alternative transformations. Columns 1-4 show total expenditures, columns 5-8 education expenditures. Columns 1 and 5 repeat the results from rows 1 and 5 of Table 3 using an IHS transformation. Column 2 uses the log of nominal expenditures; column 3 uses nominal expenditures. Results are similar in both cases. In column 4, we estimate household per capita instead of total expenditure (using an IHS transformation) to capture potential changes in household structure that could be correlated with exposure. Effects are smaller, but results show that male and female exposure still increases household expenditures. Results for education expenditures using a log transformation (column 6) or nominal values (column 7) leads to different results due to the large number of zeroes for education expenditures (over 20,000 observations are dropped) and because the education data are heavily skewed.

Finally, we re-estimate the results measuring the effect of school construction on outcome indexes (as in Figure 1) using alternative control variables. Appendix Figure A.10 shows results excluding the interaction of birth year dummies and water and sanitation programs. The magnitudes and levels of statistical significance are consistent in this case.

7. Rate of Return and Fiscal Impacts of School Construction

Regression results highlight the various beneficial impacts for individuals exposed to school construction and for the intergenerational transmission of those benefits. We conduct a cost-benefit analysis to evaluate whether school construction was cost efficient for the Indonesian government.[60] Most analyses compare a program's costs and overall welfare benefits of that program for the affected population. We do this as well, but what is unique in our case is our

[60] Appendix C discusses in more detail the assumptions made in our cost-benefit analysis and the specific parameters we include in the model.

ability also to use detailed data on tax revenues collected by the government to measure if these increases in government taxes collected offset the government's costs of building the schools.

We start by measuring the costs of school construction. Total costs include the initial investment to build the schools and train the teachers plus recurring commitments to pay teacher salaries each year. The key point for the costs is that there were large and upfront costs at the beginning of the program in 1973 and then subsequent smaller, but annual costs every year for teacher salaries. School construction cost approximately 782 million 2016 US dollars, around 1.5 percent of the 1973 Indonesian GDP (Duflo, 2001). Schools were expected to recruit three teachers and accommodate 120 students. Using estimates by Daroesman (1971), training three teachers across 61,800 schools would have cost the government 11.7 million in 2016 dollars.[61]

We focus on two main benefits. The first are taxes paid directly to the government. We have data on taxes each household paid and information on household expenditures that we use to estimate the ten percent Value-Added-Tax (VAT) the government would have collected on those purchases. The second benefit is improvements in the first generation's overall living standards. As shown previously, second-generation individuals also receive more education due to school construction, which could lead to further increases in future taxes and improved living standards. We do not include these second-generation effects in most of our cost-benefit analysis, so our estimates of the benefits and the internal rate of return are conservative. The key issue for the benefits is that the government or individual earns the benefits each year and they accrue over many years, but these benefits do not start until long after the schools are built.[62]

[61] On the cost side, we adjust various parameters in our model to see how the cost-benefit calculations respond. These parameters include the discount rate, the number of years the school is expected to last, the number of teachers and students per school, whether there is real growth in a teacher's salary, and the level of recurrent school administrative costs in addition to teachers' salaries.

[62] On the benefit side, we adjust various parameters to test the sensitivity of our results. These include the discount rate, the number of years the school is expected to last, the age individuals start paying taxes, an individual's life expectancy, the Indonesian economy's growth rate, and the lifetime curvature in mean taxes paid at each age.

We develop a cost-benefit model to include these costs and benefits in the specific years they would have been realized and then trace out the arc of when discounted benefits and costs offset. Table 9 summarizes results and highlights how different assumptions about parameters influence the level of costs and benefits, impact the year when benefits outweigh costs, and affect the internal rate of return (IRR). Column 1 is a less conservative approach.[63] Using these parameters yields a total cost (school building, teacher training, and recurrent teacher salaries) of 2.55 billion in 2016 USD and a total tax benefit (direct taxes paid plus VAT taxes collected) of 9.00 billion in 2016 USD. This gives a net benefit of 6.56 billion, a 1998 breakeven year, and an IRR of 10.5 percent. Moving beyond government taxes and focusing on the program's impact on improving living standards raises net benefits to 59.24 billion with an IRR of 20.7 percent.

From this baseline, we modify parameters and trace how those changes impact costs and benefits. Column 2 includes real salary growth for teachers and costs are higher and net benefits smaller. Column 3 adjusts for the lifetime curvature in an individual's tax payments that peak around ages 40-50. Tax and living standards benefits are smaller. Column 4 adjusts for real GDP per capita growth of 3.25 percent. Taxes and living standards are measured in 2016, but those had real growth prior to 2016 and this adjustment reduces net benefits. Column 5 represents what we believe is a reasonable baseline case. We maintain previous parameter values but increase the mean number of students per school from 120 to 180 (30 per grade), which is closer to what happened when these schools were built. Tax benefits are higher than costs with an IRR of 8.1 percent, while living standards are substantially larger than costs with an IRR of 16.8 percent.

[63] In this column, we assume a 5 percent discount rate, no real growth in teachers' salary, no adjustment for the lifetime curvature of an individual's earnings (and subsequent taxes), no real growth in GDP per capita, 120 students per classroom, schools last 20 years, individuals start paying taxes after age 18, school administration costs 1.25 times teachers' salaries, 3 teachers per school, and an individual's life expectancy is 60.

Column 6 assumes schools last 40 years. Original government plans called for schools to last 20 years, but since most schools still operate now, this is a reasonable assumption. Benefits increase as there are more cohorts exposed to the program, but there are more years to pay teacher salaries so costs also increases. Net benefits are higher, but the IRR only rises slightly because of the timing of when extra costs are incurred. Column 7 increases the age individuals start paying taxes; column 8 raises the recurrent cost multiplier from 1.25 to 1.5. Both changes have minor impacts on net benefits. Column 9 raises the number of teachers per school from three to six, and this raises costs. Lastly, column 10 incorporate intergenerational benefits—when the second generation starts paying taxes and spending their own household expenses. Tax and living standards benefits substantially outweigh costs.[64]

Appendix Figure A.11 graphs the discounted net tax and living standards benefits over time and highlights the breakeven years for when benefits outweigh school construction costs. Using the parameters from columns 5 and 10 in Table 9, we show two highly realistic scenarios that the government would have faced. Net tax benefits are considerably higher in Scenario 10 when we include intergenerational benefits (17.18 compared to 5.42 billion), but the breakeven year in that scenario is later because that scenario includes more teachers, higher recurrent costs, a later starting age for paying taxes, and a longer school lifetime. The net benefit to improved living standards is also much higher when the next generation's benefits are included (198.62 billion) and the improvement to the population's welfare offsets the program costs by 2003.

Across a range of different parameter estimates, school construction leads to increased government tax revenues that offset school construction costs in most cases within 55 years. We

[64] To have a situation in which tax benefits do not outweigh costs using the column 10 parameters, it is necessary to adjust parameters so that recurrent costs must be greater than 1.9, the number of students must be less than 145, or the discount rate must be larger than 5.7 percent. However, still net living standards benefits remain positive until the following more drastic parameter adjustment of increasing the discount rate to larger than 12 percent is made.

observe even larger net benefits when we include the population's improved living standards with net benefits ranging from 40 to 199 billion USD. Internal rates of return range from 13-21 percent and benefits outweigh costs within 17-30 years after the schools are built.

8. Conclusion

This paper studies the long-term and intergenerational effects of one of the largest school construction programs in history. We use a difference-in-differences estimation strategy exploiting variation across birth cohorts and districts in the number of schools built. We combine this with nationally representative data from Indonesia that contain information on a wide range of outcomes related to education, employment, migration, living standards, taxes, and marriage outcomes. We find that men and women exposed to the program attain more education, with men's education effects continuing beyond primary school. As adults, men exposed to school construction are more likely to be formal workers and work in a non-agricultural sector. Both men and women exposed to the program are more likely to have migrated from their birth district, although evidence points to increases in local migration within the province. Households in which either parent is exposed to school construction have higher living standards, more assets, and pay more government taxes. Exposure to school construction substantially alters marriage market outcomes with spouses being more educated and more likely to have migrated.

These benefits are transmitted to the next generation. Children with fathers or mothers who are exposed to the school construction program obtain more education. We observe significant effects at all levels of schooling beyond primary school, but we see the largest impacts in tertiary education with effect sizes indicating a 20 to 25 percent increase in the likelihood of the second-generation child completing university. These second-generation effects are significantly larger if the mother, as opposed to the father, is exposed to the program, with additional benefits accruing to daughters. We perform a detailed mediation analysis to explore the mechanisms that drive the intergenerational transmission of schooling. Marriage market outcomes appear to a play a crucial role, particularly whether the spouse has completed primary school, is literate, works in the formal sector, or works outside of agriculture.

Our cost benefit analysis highlights that under all reasonable assumptions school construction pays for itself in terms of additional expected government tax revenues, not to mention the additional benefits of improved living standards. Furthermore, given the observed intergenerational transmission of education, the likely long-run benefits are vast. To gain additional insight into the intergenerational transmission of education, we perform an exploratory analysis calculating the intergenerational elasticity (IGE) of education between children and parents. In policy circles, especially in developing countries, there is considerable debate about the optimal amount of intergenerational mobility. High mobility suggests equality of opportunity whereby a parent's outcome does not mechanically determine the child's, but it is critical to understand the mechanisms driving that intergenerational correlation (see Black and Devereux, 2011 and Mazumder, 2015 for a discussion of this literature). Low mobility that is due to differential access to schooling suggests that public policy can play a role in equalizing opportunities. Comparing the IGE across high and low program intensity areas and between young and old cohorts in our Indonesian data, we find there is an increase in mobility for children whose parents are exposed to school construction, highlighting the realized benefits of this government education policy. The broader societal impacts and changes in intergenerational transmission of human capital warrant further research.

References

Ajayi, Kehinde F. and Phillip H. Ross. 2017. "The Effects of Education on Financial Outcomes: Evidence from Kenya." Unpublished manuscript.

Akresh, Richard, Damien de Walque, and Harounan Kazianga. 2013. "Cash Transfers, Parental Investments, and Child Welfare: Evidence from a Randomized Evaluation of the Role of Conditionality" World Bank Policy Research Working Paper 6340.

Akresh, Richard, Damien de Walque, and Harounan Kazianga. 2016. "Evidence from a Randomized Evaluation of the Household Welfare Impacts of Conditional and Unconditional Cash Transfers Given to Mothers or Fathers" World Bank Policy Research Working Paper 7730.

Akresh, Richard, Emilie Bagby, Damien de Walque, and Harounan Kazianga. 2012. "Child Labor, Schooling, and Child Ability." World Bank Policy Research Working Paper 5965.

Akresh, Richard. 2004. "School Enrollment Impacts of Non-traditional Household Structure." IZA Discussion Paper 1379.

Akresh, Richard. 2009. "Flexibility of Household Structure: Child Fostering Decisions in Burkina Faso." *Journal of Human Resources*, 44(4): 976-997.

Alderman, Harold, Jooseop Kim, Peter F. Orazem. 2003. "Design, Evaluation, and Sustainability of Private Schools for the Poor: The Pakistan Urban and Rural Fellowship School Experiments." *Economics of Education Review*, 22(3): 265-274.

Almond, Douglas and Janet Currie. 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*, Vol. 4B, ed. Orley Ashenfelter and David Card, 1315-1486. New York: Elsevier.

Andrabi, Tahir, Jishnu Das, Asim Khwaja, and Tristan Zajonc. 2011. "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics." *American Economic Journal: Applied Economics*, 3(3): 29-54.

Angrist, Josh. 1995. "The Economic Returns to Schooling in the West Bank and Gaza Strip." *American Economic Review*, 85(5): 1065-1087.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. Forthcoming. "Bride Price and Female Education." *Journal of Political Economy*, forthcoming.

Attanasio, Orazio, Costas Meghir, and Emily Nix. 2017. "Human Capital Development and Parental Investment in India." NBER Working Paper 21740.

Baird, Sarah, Craig McIntosh, Berk Ozler. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics*, 26(4): 1709-1753.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-Run Impacts of a Child Health Investment." *Quarterly Journal of Economics*, 131(4): 1637-1680.

Banerjee, Abhijit, Esther Duflo, Nathanael Golderg, Dean Karlan, Robert Osei, William Pariente, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science*, 348(6236): 1260799.

Behrman, Jere and Mark Rosenzweig. 2002. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review*, 92(1): 323-334.

Behrman, Jere, Susan Parker, Petra Todd. 2011. "Do Conditional Cash Transfers for Schooling Generate Lasting Benefits? A Five-Year Followup of PROGRESA/Oportunidades." *Journal of Human Resources*, 46(1): 93-122.

Benhassine, Najy, Florencia Devoto, Esther Duflo, Pascaline Dupas, Victor Pouliquen. 2015. "Turning a Shove into a Nudge? A 'Labeled Cash Transfer' For Education." *American Economic Journal: Economic Policy*, 7(3): 86-125.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society, Series B*, 57(1): 289-300.

Bhalotra, Sonia and Atheendar Venkataramani. 2018. "Shadows of the Captain of the Men of Death: Health Innovation, Human Capital Investment, and Institutions." Unpublished manuscript.

Bharati, Tushar, Seungwoo Chin, and Dawoon Jung. 2018. "Recovery from an Early Life Shock Through Improved Access to Schools: Evidence from Indonesia." Unpublished manuscript.

Bianchi, Nicola. 2018. "The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in STEM Majors." Unpublished manuscript.

Bils, Mark and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review*, 90(5): 1160-1183.

Black, Sandra E. and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, editors, Volume 4, Elsevier: 1487-1541.

Black, Sandra E., Aline Butikofer, Paul J. Devereux, and Kjell G. Salvanes. Forthcoming. "This is Only a Test? Long-run Impacts of Prenatal Exposure to Radioactive Fallout." *Review of Economics and Statistics*.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review*, 95(1): 437-449.

Blattman, Chris, Nathan Fiala, and Sebastian Martinez. 2018. "The Long Term Impacts of Grants on Poverty: 9-Year Evidence From Uganda's Youth Opportunities Program." NBER working paper 24999.

Breierova, Lucia and Esther Duflo. 2004. "The Impact of Education on Fertility and Child Mortality: Do Fathers Really Matter Less Than Mothers?" NBER Working Paper 10513.

Burde, Dana and Leigh Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics*, 5(3): 27-40.

Cameron, A. Colin, Jonah Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business and Economic Statistics*, 29(2): 238-249.

Chetty, Raj, John Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *Quarterly Journal of Economics*, 126(4): 1593-1660.

Chevalier, Arnaud. 2004. "Parental Education and Child's Education: A Natural Experiment." IZA Discussion Paper No. 1153.

Clark, Damon and Heather Royer. 2013. "The Effect of Education on Adult Mortality and Health: Evidence from Britain." *American Economic Review*, 103(6): 2087-2120.

Crepon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. 2013. "Do Labor Market Policies Have Displacement Effects: Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics*, 128(2): 531-580.

Currie, Janet and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Quarterly Journal of Economics*, 118(4): 1495-1532.

Currie, Janet. 2011. "Inequality at Birth: Some Causes and Consequences." American Economic Review, 101(3): 1-22.

Daroesman, Ruth. 1971. "Finance of Education." *Bulletin of Indonesian Economic Studies*, 7(3): 61-95.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2015. "Education, HIV, and Early Fertility: Experimental Evidence from Kenya." *American Economic Review*, 105(9): 2757-2797.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2017. "The Impact of Free Secondary Education: Experimental Evidence from Ghana." Unpublished manuscript.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795-813.

Duflo, Esther. 2004. "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics*, 74(1): 163-197.

Evans, David and Muthoni Ngatia. 2018. "School Costs, Short-Run Participation, and Long-Run Outcomes: Evidence from Kenya." World Bank Policy Research Working Paper 8421.

Fiszbein, Ariel and Norbert Schady. 2009. "Conditional Cash Transfers: Reducing Present and Future Poverty." The World Bank, Washington, DC.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." *American Economic Review*, 92(4): 999-1012.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan M. Chang, Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science*, 344(6187): 998-1001.

Geruso, Michael and Heather Royer. 2018. "The Impact of Education on Family Formation: Quasi-Experimental Evidence from the UK." NBER Working Paper 24332.

Glewwe, Paul and Karthik Muralidharan. 2016. "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications." *Handbook of the Economics of Education*, Volume 5. Elsevier Publisher.

Gorseline, Donald. 1932. "The Effect of Schooling upon Income. Bloomington." University of Indiana Press.

Handa, Sudhanshu. 2002. "Raising Primary School Enrollment in Developing Countries: The Relative Important of Supply and Demand." *Journal of Development Economics*, 69(1): 103-128.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "General Equilibrium Effects: A Study of Tuition Policy." *American Economic Review*, 88(2): 381-386.

Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the Highscope Perry Preschool Program." *Journal of Public Economics*, 94(1-2): 114-128.

Heckman, James, John Eric Humphries, and Gregory Veramendi. 2018. "Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking." *Journal of Political Economy*, 126(S1): S197-S246.

Hicks, Joan, Marieke Kleemans, Nicholas Li, Edward Miguel. 2017. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata." NBER Working Paper 23253.

Jacob, Brian, Lars Lefgren, and David Sims. 2010. "The Persistence of Teacher-Induced Learning." *Journal of Human Resources*, 45(4): 915-943.

Jensen, Robert. 2010. "The (Perceived) Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*, 25(2): 515-548.

Karachiwalla, Naureen and Giordano Palloni. 2019. "Human Capital and Structural Transformation." *IFPRI Discussion Paper* 01836.

Kazianga, Harounan, Dan Levy, Leigh Linden, and Matt Sloan. 2013. "The Effects of 'Girl-Friendly' Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso." *American Economic Journal: Applied Economics*, 5(3): 41-62.

Khanna, Gaurav. 2018. "Large-scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India." Unpublished manuscript.

Kleemans, Marieke and Jeremy Magruder. 2018. “Labour Market Responses to Immigration: Evidence from Internal Migration Driven by Weather Shocks.” *Economic Journal*, 128(613): 2032-2065.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica*, 75(1): 83-119.

Kremer, Michael, Edward Miguel, and Rebecca Thornton. 2009. “Incentives to Learn.” *Review of Economics and Statistics*, 91(3): 437-456.

Lleras-Muney, Adriana. 2005. “The Relationship Between Education and Adult Mortality in the U.S.” *Review of Economic Studies*, 72(1): 189-221.

Malamud, Ofer, Andreea Mitrut, and Cristian Pop-Eleches. 2018. “The Effect of Education on Mortality and Health: Evidence from a Schooling Expansion in Romania.” NBER Working Paper 24341.

Manski, Charles. 1990. “Nonparametric Bounds on Treatment Effects.” *American Economic Review*, 80(2): 319-323.

Marazyan, Karine. 2011. “Effects of a Sibship Extension to Foster Children on Children’s School Enrollment: A Sibling Rivalry Analysis for Indonesia.” *Journal of Development Studies*, 47(12): 497-518.

Martinez-Bravo, Monica. 2017. “The Local Political Economy Effects of School Construction in Indonesia.” *American Economic Journal: Applied Economics*, 9(2): 256-289.

Maurin, Eric and Sandra McNally. 2008. “Vive la Revolution! Long-Term Educational Returns of 1968 to the Angry Students.” *Journal of Labor Economics*, 26(1): 1-33.

Mazumder, Bhashkar. 2015. “Estimating the Intergenerational Elasticity and Rank Association in the US: Overcoming the Current Limitations of Tax Data.” *Research in Labor Economics*, 43: 83-129.

Mazumder, Bhashkar, Maria Rosales-Rueda, and Margaret Triyana. 2019. “Intergenerational Human Capital Spillovers: Indonesia’s School Construction and Its Effects on the Next Generation.” *AEA Papers and Proceedings*, 109: 243-249.

McCrary, Justin and Heather Royer. 2011. “The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth.” *American Economic Review*, 101(1): 158-195.

McEwan, Patrick. 2015. “Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments.” *Review of Educational Research*, 85(3): 353-394.

Meghir, Costas, Marten Palme, and Emilia Simeonova. 2018. “Education and Mortality: Evidence from a Social Experiment.” *American Economic Journal: Applied Economics*, 10(2): 234-256.

Muralidharan, Karthik and Nishith Prakash. 2017. "Cycling to School: Increasing Secondary School Enrollment for Girls in India." *American Economic Journal: Applied Economics,* 9(3): 321-350.

Oreopoulos, Philip, Marianne Page, and Ann Huff Stevens. 2008. "The Intergenerational Effects of Worker Displacement." *Journal of Labor Economics*, 26(3): 455-483.

Osili, Una Okonkwo and Bridget Terry Long. 2008. "Does Female Schooling Reduce Fertility? Evidence from Nigeria." *Journal of Development Economics*, 87(1): 57-75.

Oster, Emily and Rebecca Thornton. 2011. "Menstruation, Sanitary Products, and School Attendance: Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics,* 3(1): 91-100.

Parker, Susan and Tom Vogl. 2018. "Do Conditional Cash Transfers Improve Economic Outcomes in the Next Generation? Evidence from Mexico." NBER Working Paper 24303.

Pop-Eleches, Cristian and Miguel Urquiola. 2013. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review*, 103(4): 1289-1324.

Ravallion, Martin. 1988. "INPRES and Inequality: A Distributional Perspective on the Centre's Regional Disbursements" *Bulletin of Indonesian Economic Studies*, 24(3): 53–71.

Rizky, Mayang, Daniel Suryadarma, and Asep Suryahadi. 2018. "Effect of Growing-Up Poor on Labour Market Outcome: Evidence from Indonesia." Unpublished manuscript.

Rohner and Saia. 2019. "Education and Conflict Evidence from a Policy Experiment in Indonesia." *CEPR Discussion Paper* No. DP13509.

Thomas, Duncan, Firman Witoelar, Elizabeth Frankenberg, Bondan Sikoki, John Strauss, Cecep Sumantri, and Wayan Suriastini. 2012. "Cutting the Costs of Attrition: Results from the Indonesia Family Life Survey." *Journal of Development Economics,* 98(1): 108-123.

Walsh, J. R. 1935. "Capital Concept Applied to Man." *Quarterly Journal of Economics*, 49(2): 255–285.

Wantchekon, Leonard, Marko Klašnja, and Natalija Novta. 2014. "Education and Human Capital Externalities: Evidence from Colonial Benin." *Quarterly Journal of Economics*, 130(2): 703–757.

World Bank. 1989. "Indonesia Basic Education Study." World Bank Report 7841-IND. Washington, D.C. Population and Human Resources Operations Division.

Wydick, Bruce, Paul Glewwe and Laine Rutledge. 2013. "Does International Child Sponsorship Work? A Six-Country Study of Impacts on Adult Life Outcomes." *Journal of Political Economy,* 121(2): 393-436.

Young, Alwyn. 2013. "Inequality, the Urban-Rural Gap, and Migration." *Quarterly Journal of Economics*, 128(4): 1727-1785.

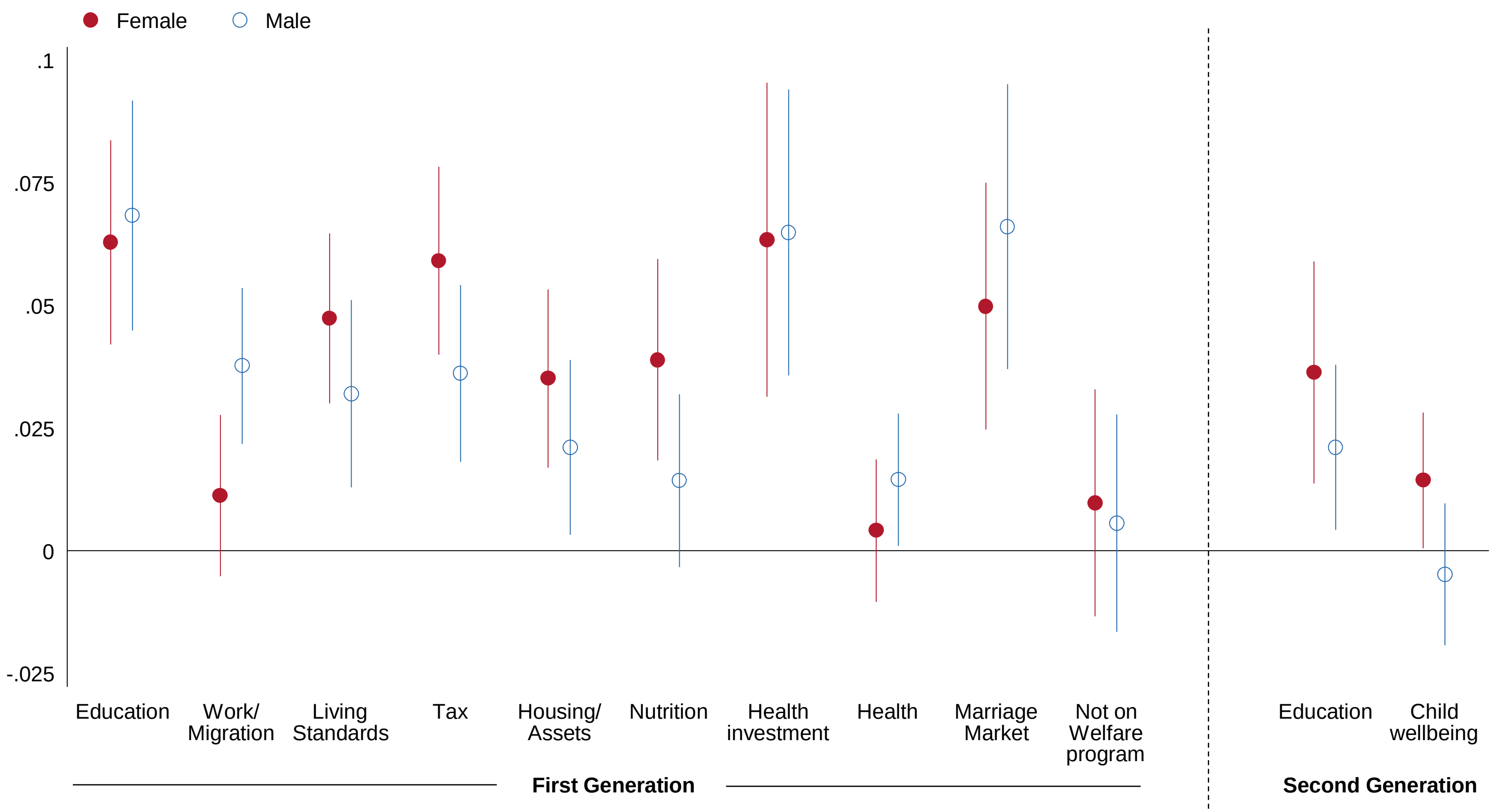


Figure 1. Effect of school construction on indexes of long-run outcomes

Notes: Following Kling, Liebman, and Katz (2007), we define indexes for families of outcomes by defining a Z-score for each outcome relative to the control group (defined in this case as the old cohort in low program intensity districts). Then, we average the Z-scores across all outcomes in the same family to get an index, such as “Marriage Market”. Following Banerjee et al. (2015) to get standardized treatment effects, we then standardize the Kling indexes relative to the mean and standard deviation of the control group. Each dot in the figure represents the coefficient of the interaction of the number of INPRES schools built between 1973 and 1979 in one’s birth district and a dummy for being born between 1968 and 1972. Solid lines represent 95% confidence intervals. The individual outcomes making up the index for each family are listed in Tables 1-5 and Appendix Table A.1.

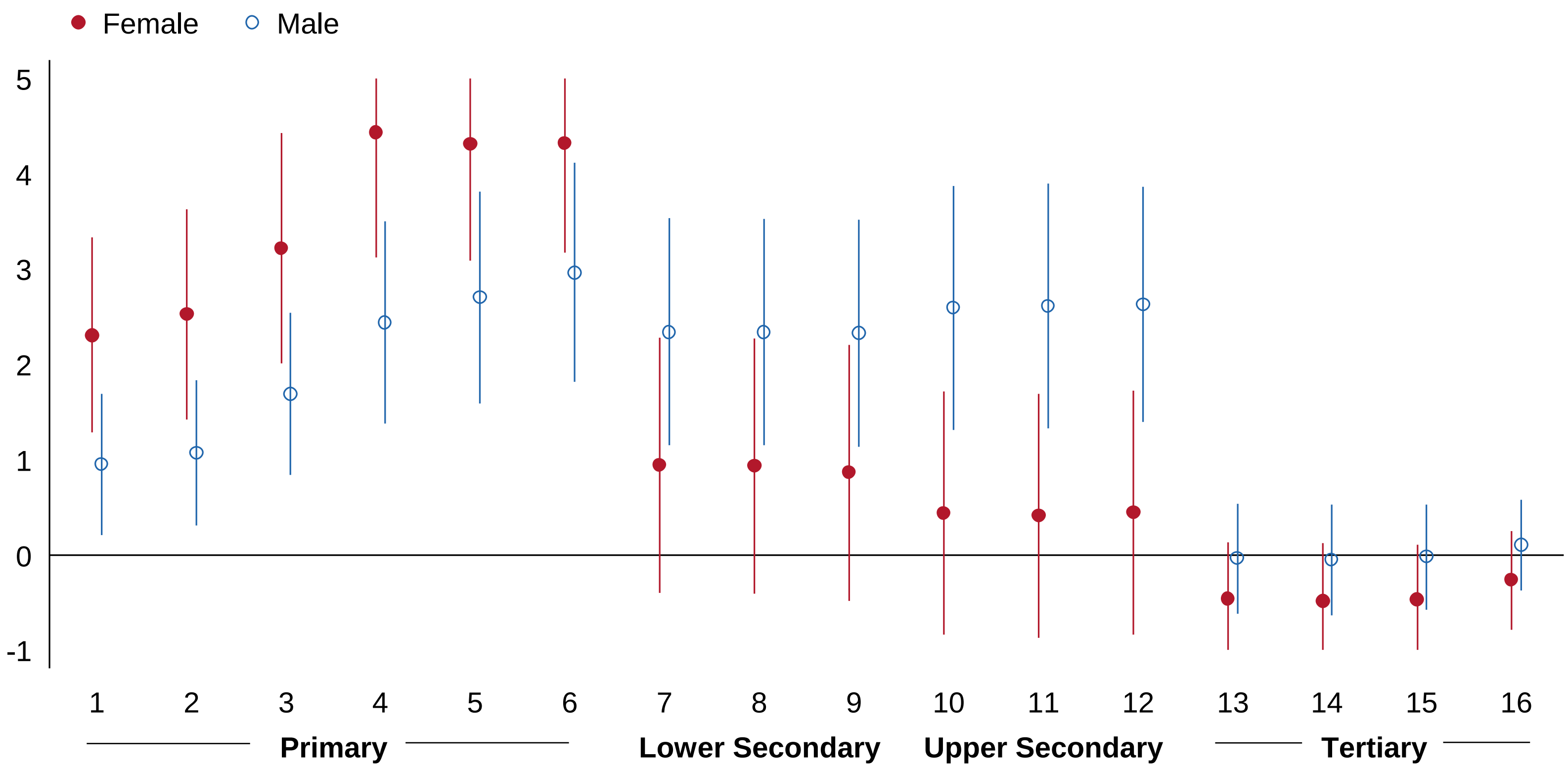


Figure 2. Effect of school construction on the probability of first-generation individual attending at least *n*-years of schooling

Notes: Effect size measures the impact of one additional school constructed per 1,000 children on the probability of completing at least *n*-years of schooling in percentage points. We show estimated regression coefficients and their respective 95% confidence intervals.

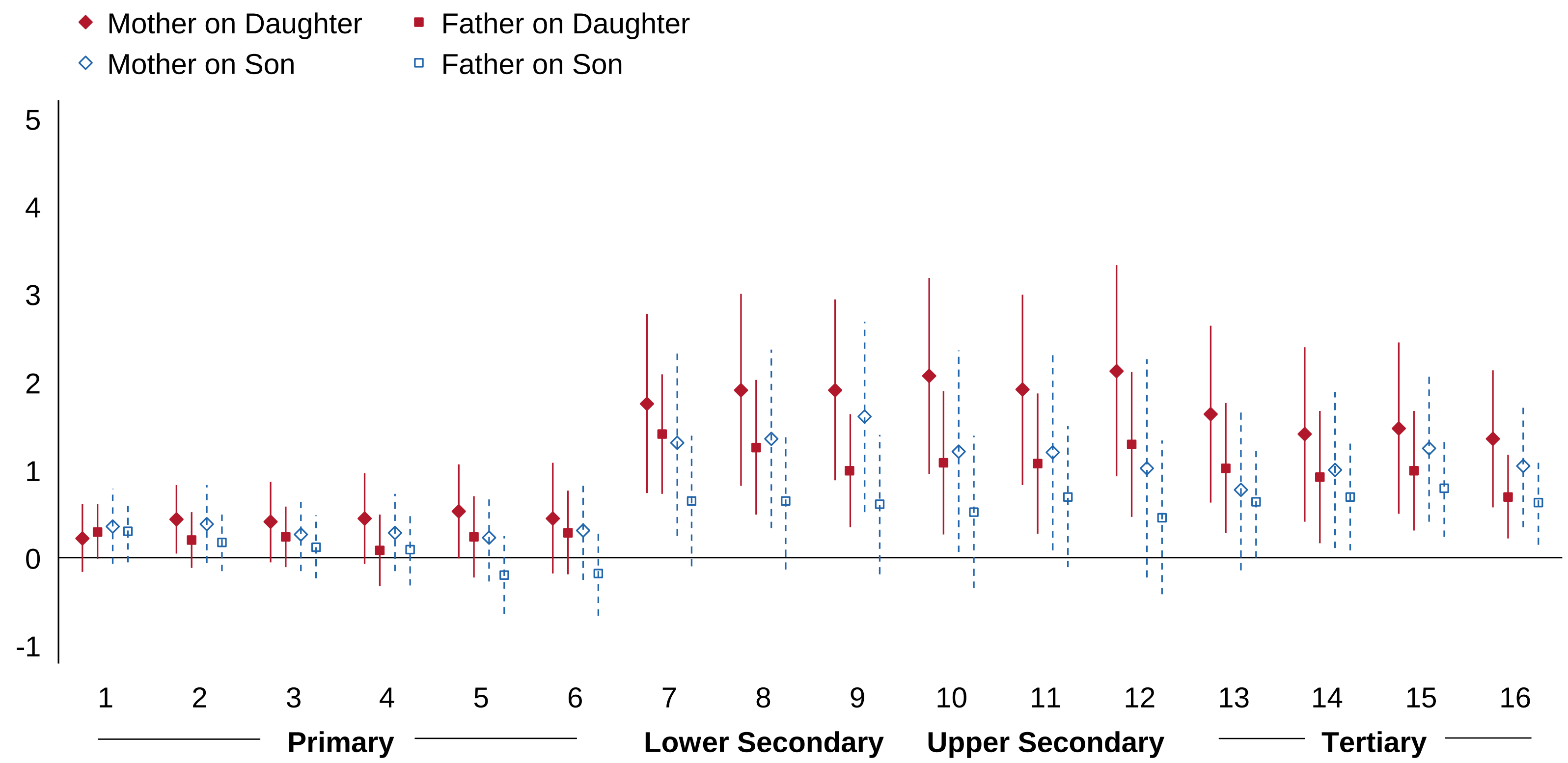


Figure 3. Effect of school construction on the probability of second-generation individual attending at least *n*-years of schooling

Notes: Effect size measures the impact of one additional school constructed per 1,000 children in the mother's or father's birth district on the probability of a second-generation individual (daughter or son) attending at least *n*-years of schooling in percentage points. Each dot represents a coefficient in a separate regression. We show estimated regression coefficients and their respective 95% confidence intervals.

Table 1. Effect of school construction on first generation's education

		Mean / SD		Effect of Program Exposure on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Years of schooling	Based on highest education level and grade attended. Standard durations of study are assumed; grade retentions are not counted	8.022 (4.230)	7.105 (4.215)	0.268*** (0.047) [0.000]	0.234*** (0.042) [0.000]
Completed Primary	Indicator defined as 1 if highest diploma completed is higher than or equal to Primary	0.813 (0.390)	0.727 (0.446)	0.026*** (0.006) [0.000]	0.041*** (0.006) [0.000]
Completed Lower Secondary	Indicator defined as 1 if highest diploma completed is higher than or equal to Lower Secondary	0.385 (0.487)	0.312 (0.463)	0.023*** (0.006) [0.000]	0.008 (0.007) [0.422]
Completed Upper Secondary	Indicator defined as 1 if highest diploma completed is higher than or equal to Upper Secondary	0.338 (0.473)	0.261 (0.439)	0.026*** (0.006) [0.000]	0.005 (0.006) [0.422]
Completed Tertiary	Indicator defined as 1 if highest diploma completed is higher than or equal to Tertiary	0.095 (0.293)	0.077 (0.267)	-0.001 (0.003) [0.741]	-0.003 (0.003) [0.422]
Literate	Literacy is a binary outcome and is self-reported	0.953 (0.212)	0.909 (0.287)	0.015*** (0.004) [0.001]	0.033*** (0.006) [0.000]
Education index	Aggregates all 6 outcomes and standardizes it to the mean of the old cohort in low-program districts. Effects are interpreted as standard deviation changes from the mean.			0.068*** (0.012)	0.063*** (0.011)

Notes: Effects of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over all 6 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. There are 72,367 and 71,423 observations in the men's and women's regressions, respectively.

Table 2. Effect of school construction on first generation's work and migration

		Mean / SD		Effect of Program Exposure on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Work	Indicator defined as 1 if individual worked in the past week or has an occupation but was temporarily absent from work in the past week	0.948 (0.223)	0.638 (0.481)	0.006** (0.003) [0.080]	0.003 (0. 005) [0.953]
Work hours	Hours worked in the past week conditional on working, i.e. missing for non-working individuals	40.981 (17.115)	36.227 (18.792)	0.258 (0.158) [0.101]	0.157 (0.208) [0.953]
Formal worker	Indicator defined as 1 if individual reported working as an employee as opposed to being self-employed, family/unpaid work or freelance work, conditional on working	0.327 (0.469)	0.236 (0.425)	0.011*** (0.004) [0.032]	-0.005 (0.005) [0.953]
Non-agriculture sector	Indicator defined as 1 for working in a sector outside of agriculture; conditional on working	0.560 (0.496)	0.547 (0.498)	0.012*** (0.005) [0.032]	0.002 (0.005) [0.953]
Service sector	Indicator for working in trade, hotel, restaurant, transportation; warehousing, information, communication; finance and insurance, and service sectors, conditional on working	0.364 (0.481)	0.459 (0.498)	0.010*** (0.004) [0.032]	-0.000 (0.006) [0.953]
Migrant	Indicator defined as 1 if the current district of residence is not the same as the individual's birth district	0.273 (0.445)	0.245 (0.430)	0.007** (0.003) [0.085]	0.008** (0.003) [0.166]
Local migration	Indicator defined as 1 if migration occurred within the individual's birth province	0.108 (0.310)	0.106 (0.307)	0.005* (0.003) [0.101]	0.005** (0.003) [0.229]
Work/Migration index	Aggregates all 7 outcomes and standardizes it to the mean of the old cohort in low-program districts. Effects are interpreted as standard deviation changes from the mean.			0.038*** (0.008)	0.011 (0.008)

Notes: Effect of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over all 7 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. There are 72,367 observations for men and 68,574 unconditional on working. There are 71,423 observations for women and 45,560 conditional on working.

Table 3. Effect of school construction on first generation's living standards

		Mean / SD		Effect of Program Exposure on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Total (Rp10k)	Household's average monthly expenditure; means are reported in 10,000 Indonesian Rupiah (IDR) increments. Total expenditures are made up of food and non-food expenditures.	391.65 (352.49)	375.62 (343.82)	0.021*** (0.007) [0.010]	0.032*** (0.007) [0.000]
Food (Rp10k)		194.44 (120.45)	184.22 (121.11)	0.014** (0.007) [0.036]	0.028*** (0.007) [0.000]
Non-food (Rp10k)		197.21 (271.88)	191.39 (261.11)	0.027*** (0.008) [0.004]	0.039*** (0.008) [0.000]
Non-food/Total	Share of non-food over total expenditures.	44.59 (13.38)	45.14 (13.75)	0.287*** (0.110) [0.024]	0.237*** (0.102) [0.021]
Education (Rp10k)	Education expenditures fall under non-food expenditures and include admission, tuition, extracurricular fees, textbooks, stationery, and tutoring.	13.97 (33.17)	12.20 (30.35)	0.160** (0.064) [0.024]	0.193** (0.076) [0.011]
Tax (Rp10k)	Self-reported total tax expenditures include taxes on land and building, vehicle, levies and retributions.	4.75 (11.43)	4.55 (10.74)	0.078*** (0.017) [0.000]	0.123*** (0.019) [0.000]
Living standards index	Aggregates total, food, and education expenditures and excludes non-food expenditure and non-food/total ratio to avoid collinearity in the regression. Tax expenditure belongs to the tax index and is excluded from living standards index. Standardizes it to the mean of the old cohort in low-program districts. Effects are interpreted as standard deviation changes from the mean.			0.032*** (0.010)	0.047*** (0.009)

Notes: Effect of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over the first 5 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. FDR q-values for tax expenditures are computed over 4 tax outcomes detailed in Appendix Table A1. There are 68,687 and 66,249 observations in the men's and women's regressions, respectively. We apply an inverse hyperbolic sine transformation to all monetary values. Estimates can be interpreted as percentage changes.

Table 4. Effect of school construction on first generation's marriage market

		Mean / SD		Effect of Program Exposure on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Age of first marriage	Age of first marriage for ever-married household members	25.219 (5.022)	20.888 (4.788)	0.058 (0.053) [0.867]	0.050 (0.059) [0.745]
Spouse's characteristics:					
Years of schooling	Based on highest education level and grade attended. Standard durations of study are assumed; grade retentions are not counted	7.635 (4.081)	7.426 (4.192)	0.180*** (0.046) [0.001]	0.116*** (0.043) [0.057]
Completed Primary	Indicator defined as 1 if highest diploma completed is higher than or equal to Primary	0.797 (0.402)	0.773 (0.419)	0.038*** (0.006) [0.000]	0.025*** (0.005) [0.000]
Literate	Literacy is a binary outcome and is self-reported	0.939 (0.239)	0.944 (0.230)	0.027*** (0.006) [0.000]	0.016*** (0.004) [0.001]
Work	Indicator defined as 1 if individual worked in the past week or has an occupation but was temporarily absent from work in the past week	0.614 (0.487)	0.918 (0.275)	0.003 (0.005) [0.867]	-0.001 (0.004) [0.745]
Formal worker	Indicator defined as 1 if individual reported working as an employee as opposed to being self-employed, family/unpaid work or freelance work, conditional on working	0.239 (0.427)	0.293 (0.455)	0.003 (0.005) [0.867]	0.005 (0.005) [0.745]
Non-agriculture sector	Indicator defined as 1 for working in a sector outside of agriculture; conditional on working	0.555 (0.497)	0.517 (0.500)	0.006 (0.005) [0.867]	0.005 (0.007) [0.745]
Migrant	Indicator defined as 1 if the current district of residence is not the same as the individual's birth district	0.260 (0.439)	0.277 (0.448)	0.007* (0.004) [0.387]	0.008** (0.003) [0.164]

No health complaint	Self-reported indicator defined as 1 if did not experience a health complaint in the past month	0.694 (0.461)	0.646 (0.478)	-0.001 (0.004) [0.867]	0.006 (0.005) [0.745]
Own characteristics:					
Children 0-14	Number of children aged 0-14 living in the household	0.910 (1.059)	0.559 (0.868)	-0.012 (0.017) [0.867]	-0.035** (0.016) [0.559]
Marriage market index	Aggregates all 10 outcomes and standardizes it to the mean of the old cohort in low-program districts. Effects are interpreted as standard deviation changes from the mean. For the index, we reverse the sign for children 0-14 to indicate a positive outcome.			0.066*** (0.015)	0.050*** (0.013)

Notes: Effect of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over all 10 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. Spouse's characteristics are defined for household heads and spouses in the sample. The spouse's years of schooling, completed primary, literacy, work, no health complaint regressions have 64,422 and 55,468 observations for men and women, respectively, because it is set to missing if the spouse does not currently live in the household (divorced, widowed). The spouse's formal worker and non-agriculture sector regressions have 39,550 and 50,892 observations because it is set to missing if the spouse does not currently live in the household and/or does not work. The age of first marriage regression has 70,571 and 69,623 observations because it is set to missing if the individual is never married. The children 0-14 regression has 68,687 and 66,249 observations.

Table 5. Effect of school construction on second-generation's education

		Mean / SD		Effect of Program Exposure by:	
Outcome	**Description**	**Father**	**Mother**	**Fathers on Children**	**Mothers on Children**
Years of schooling	Child's years of school based on highest education level and grade attended. Standard durations of study are assumed; grade retentions are not counted	7.967 (4.340)	8.854 (4.278)	0.097*** (0.032) [0.014]	0.169*** (0.045) [0.001]
Completed Primary	Indicator defined as 1 if child's highest diploma completed is higher than or equal to Primary	0.637 (0.481)	0.728 (0.445)	0.000 (0.002) [0.928]	0.001 (0.003) [0.796]
Completed Lower Secondary	Indicator defined as 1 if child's highest diploma completed is higher than or equal to Lower Secondary	0.413 (0.492)	0.504 (0.500)	0.006* (0.003) [0.171]	0.015*** (0.005) [0.006]
Completed Upper Secondary	Indicator defined as 1 if child's highest diploma completed is higher than or equal to Upper Secondary	0.217 (0.412)	0.300 (0.458)	0.009** (0.004) [0.061]	0.014*** (0.005) [0.013]
Completed Tertiary	Indicator defined as 1 if child's highest diploma completed is higher than or equal to Tertiary	0.041 (0.198)	0.064 (0.245)	0.004* (0.002) [0.171]	0.008** (0.003) [0.044]
Age-for-grade	Indicator for child starting primary school by age 7 and never repeating school up to Upper Secondary	0.835 (0.371)	0.789 (0.408)	0.011*** (0.004) [0.030]	0.018*** (0.005) [0.002]
Second-generation education index	Aggregates all 6 outcomes and standardizes it to the mean of the old cohort in low-program districts. Effects are interpreted as standard deviation changes from the mean.			0.021** (0.009)	0.036*** (0.012)

Notes: Effect of program exposure are the regression coefficients of father or mother's young cohort dummy interacted with the number of schools constructed in father or mother's district of birth. All regressions control for parent's district of birth and cohort of birth fixed effects, child age fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at parent's district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over all 6 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. The survey restricts questions on educational attainment to individuals aged 5 and older. There are 120,838 and 105,523 observations in the father's and mother's regressions, respectively.

Table 6. Effect of school construction on second-generation's education, by parent and child gender

	Years of schooling	
	(1)	(2)
Parents born between:	1957-1962 and 1968-1972	1950-1980
Panel A: Sons and Daughters		
Father exposed	0.001	0.044**
	(0.038)	(0.021)
Mother exposed	0.160***	0.118***
	(0.059)	(0.035)
Father = Mother (p-value)	0.046	0.050
Mean	8.674	7.827
Observations	44,105	246,466
Panel B: Sons Only		
Father exposed	-0.038	0.042
	(0.049)	(0.026)
Mother exposed	0.139**	0.094**
	(0.069)	(0.040)
Father = Mother (p-value)	0.076	0.267
Mean	8.575	7.787
Observations	24,366	133,896
Panel C: Daughters Only		
Father exposed	0.036	0.046**
	(0.051)	(0.023)
Mother exposed	0.188***	0.140***
	(0.072)	(0.038)
Father = Mother (p-value)	0.134	0.026
Mean	8.796	7.875
Observations	19,739	112,570

Notes: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Robust standard errors in parentheses, clustered at the father and mother's birth district level using the multiway clustering method of Cameron, Gelbach, and Miller (2011). Father exposed indicates an interaction of the number of INPRES primary schools constructed in the father's birth district and an indicator that the father is in the young cohort. Mother exposed is defined similarly. Father = Mother indicates the p-value testing the equality of coefficients of father exposed and mother exposed within each panel. The sample in Panel A consists of both sons and daughters, Panel B sons only, and Panel C daughters only.

Table 7. Mediators of the effect of school construction on second-generation's years of schooling

	Dependent Variable: Second-generation's years of schooling									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mediator:	None	Work/ Migration	Living Standards	Tax	Housing/ Asset	Nutrition	Health investment	Health	Marriage	All
Panel A: Father										
Schools constructed	0.097***	0.080**	0.082***	0.084***	0.082***	0.095***	0.086***	0.096***	0.056*	0.055**
* Young cohort	(0.032)	(0.031)	(0.031)	(0.030)	(0.028)	(0.032)	(0.031)	(0.032)	(0.028)	(0.026)
Mediator		0.386***	0.652***	0.539***	0.772***	0.112***	0.165***	0.045***	0.616***	
		(0.015)	(0.017)	(0.017)	(0.019)	(0.018)	(0.013)	(0.011)	(0.014)	
Observations	120,838	120,838	120,838	120,838	120,838	120,838	120,838	120,838	120,838	120,838
Mean	7.967	7.967	7.967	7.967	7.967	7.967	7.967	7.967	7.967	7.967
Panel B: Mother										
Schools constructed	0.169***	0.166***	0.130***	0.138***	0.133***	0.163***	0.156***	0.168***	0.137***	0.111***
* Young cohort	(0.045)	(0.045)	(0.044)	(0.041)	(0.039)	(0.045)	(0.043)	(0.045)	(0.043)	(0.039)
Mediator		0.353***	0.852***	0.721***	1.037***	0.183***	0.189***	0.062***	0.602***	
		(0.014)	(0.019)	(0.022)	(0.019)	(0.020)	(0.014)	(0.015)	(0.023)	
Observations	105,523	105,523	105,523	105,523	105,523	105,523	105,523	105,523	105,523	105,523
Mean	8.854	8.854	8.854	8.854	8.854	8.854	8.854	8.854	8.854	8.854

Notes: Each column shows a regression of the years of schooling for a second-generation child on parent's exposure to the school construction program and includes a potential mediator variable. These mediator variables (as indicated by the column heading) are the indexes reported in Figure 1. Regressions are as in Table 5. Effect of program exposure are the regression coefficients of father or mother's young cohort dummy interacted with the number of schools constructed in father or mother's district of birth. All regressions control for parent's district of birth and cohort of birth fixed effects, child age fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by region of birth interacted with birth year dummies. Robust standard errors clustered at parent's district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values.

Table 8. Spouse's characteristics as mediators of the effect of school construction on second-generation's years of schooling

	Dependent Variable: Second-generation's years of schooling									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mediator:	None	Years of Schooling	Completed Primary	Literate	Work	Formal worker	Non-agriculture sector	Migrant	No health complaint	All
Panel A: Father										
Schools constructed * Young cohort	0.097***	0.069**	0.044	0.049*	0.093***	0.068*	0.070**	0.090***	0.092***	0.020
	(0.032)	(0.030)	(0.028)	(0.027)	(0.032)	(0.036)	(0.034)	(0.032)	(0.032)	(0.032)
Mediator		0.146***	1.381***	1.798***	-0.054**	0.627***	0.940***	0.330***	-0.037*	
		(0.004)	(0.042)	(0.087)	(0.024)	(0.031)	(0.030)	(0.027)	(0.020)	
Observations	120,838	116,550	116,550	116,550	116,550	70,861	70,861	116,550	116,550	70,861
Mean	7.967	7.942	7.942	7.942	7.942	7.799	7.799	7.942	7.942	7.799
Panel B: Mother										
Schools constructed * Young cohort	0.169***	0.126***	0.115***	0.125***	0.153***	0.130***	0.126***	0.152***	0.155***	0.091**
	(0.045)	(0.045)	(0.043)	(0.044)	(0.046)	(0.047)	(0.044)	(0.046)	(0.046)	(0.043)
Mediator		0.176***	1.567***	1.906***	-0.456***	0.677***	1.007***	0.452***	-0.092***	
		(0.004)	(0.040)	(0.079)	(0.063)	(0.035)	(0.034)	(0.032)	(0.027)	
Observations	105,523	91,384	91,384	91,384	91,384	85,036	85,036	91,384	91,384	85,036
Mean	8.854	8.780	8.780	8.780	8.780	8.659	8.659	8.780	8.780	8.659

Notes: Each column shows a regression of the years of schooling for a second-generation child on parent's exposure to the school construction program and includes a potential mediator variable. These mediator variables (as indicated by the column heading) are spouse's characteristics from Table 4. Regressions are as in Table 5. Effect of program exposure are the regression coefficients of father or mother's young cohort dummy interacted with the number of schools constructed in father or mother's district of birth. All regressions control for parent's district of birth and cohort of birth fixed effects, child age fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at parent's district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values.

Table 9. Cost-benefit analysis of school construction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Parameters										
Discount rate (%)	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Teachers salary growth (Y/N)	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lifetime curvature (Y/N)	N	N	Y	Y	Y	Y	Y	Y	Y	Y
GDP/capita growth (%)	0	0	0	3.25	3.25	3.25	3.25	3.25	3.25	3.25
Students/schools	120	120	120	120	180	180	180	180	180	180
School lifetime (years)	20	20	20	20	20	40	40	40	40	40
Start paying taxes after age:	18	18	18	18	18	18	22	22	22	22
Recurrent costs/salaries multiplier	1.25	1.25	1.25	1.25	1.25	1.25	1.25	1.5	1.5	1.5
Teachers/schools	3	3	3	3	3	3	3	3	6	6
Intergenerational benefits (Y/N)	N	N	N	N	N	N	N	N	N	Y
Costs										
Schools construction	0.78									
Teachers training	0.12									
Teachers' salaries	1.65	2.95	2.95	2.95	2.95	5.08	5.08	6.10	12.19	12.19
Benefits										
Paid by cohorts born in	1968-1980					1968-2000	1968-2000			
Collected between years	1987-2040					1987-2060	1991-2060			1991-2089
<u>Tax receipts</u>	9.00	9.00	7.32	6.11	9.16	19.87	18.14	18.14	18.14	30.12
Net Benefit (Benefits - Costs)	6.56	5.26	3.58	2.37	5.42	14.00	12.27	11.25	5.15	17.18
Breakeven year	1998	2001	2007	2017	2009	2013	2016	2018	2031	2028
<u>Living standards</u>	61.69	61.69	53.18	43.64	65.46	142.00	128.34	128.34	128.34	211.23
Net Benefit (Benefits - Costs)	59.24	57.95	49.44	39.90	61.72	136.12	122.47	121.45	115.36	198.62
Breakeven year	1990	1991	1992	1995	1994	1994	1998	1999	2003	2003
Internal Rate of Return (%)										
Tax receipts	10.48	8.87	7.68	6.64	8.10	9.11	8.53	8.05	6.05	7.22
Living standards	20.68	19.38	17.69	14.83	16.84	17.57	15.77	15.26	13.08	13.54

Note: All values are in billions of US dollars in 2016. Assumptions on number of students and teachers per school, recurrent costs/salaries multiplier, and school lifetime follow Duflo (2001). Schools construction costs are obtained from Duflo (2001), teachers training from Daroesman (1972), and teachers' salaries from various sources (see Appendix C for more details). Benefits are paid by cohorts that could attend the full 6 years of primary education until their death. Life expectancy is assumed to be 60. Breakeven year is the first year when the present discounted value of benefits exceeds that of costs. Tax receipts consist of direct tax expenditures plus 10% VAT on total expenditures. Living standards is proxied with total household expenditures. Internal rate of return is the discount rate that equates the present discounted value of benefits and costs.

Online Appendix for

Long-term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia

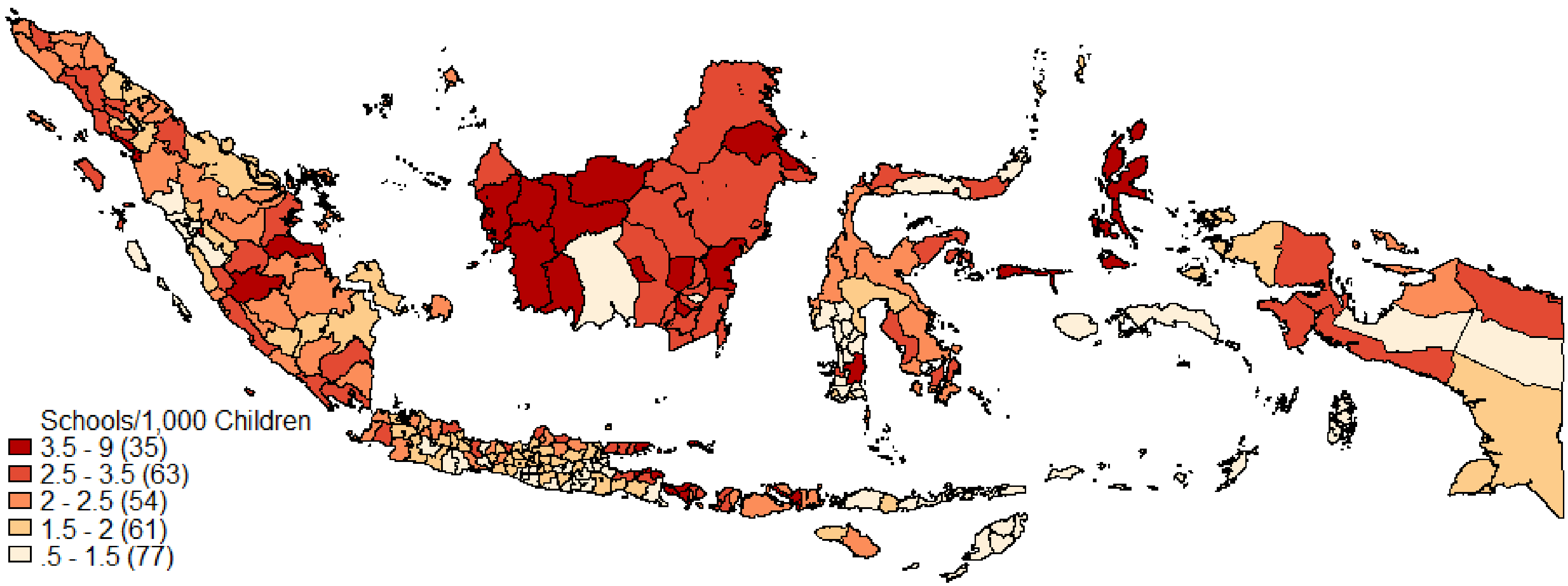
Richard Akresh
University of Illinois at Urbana-Champaign

Daniel Halim
World Bank

Marieke Kleemans
University of Illinois at Urbana-Champaign

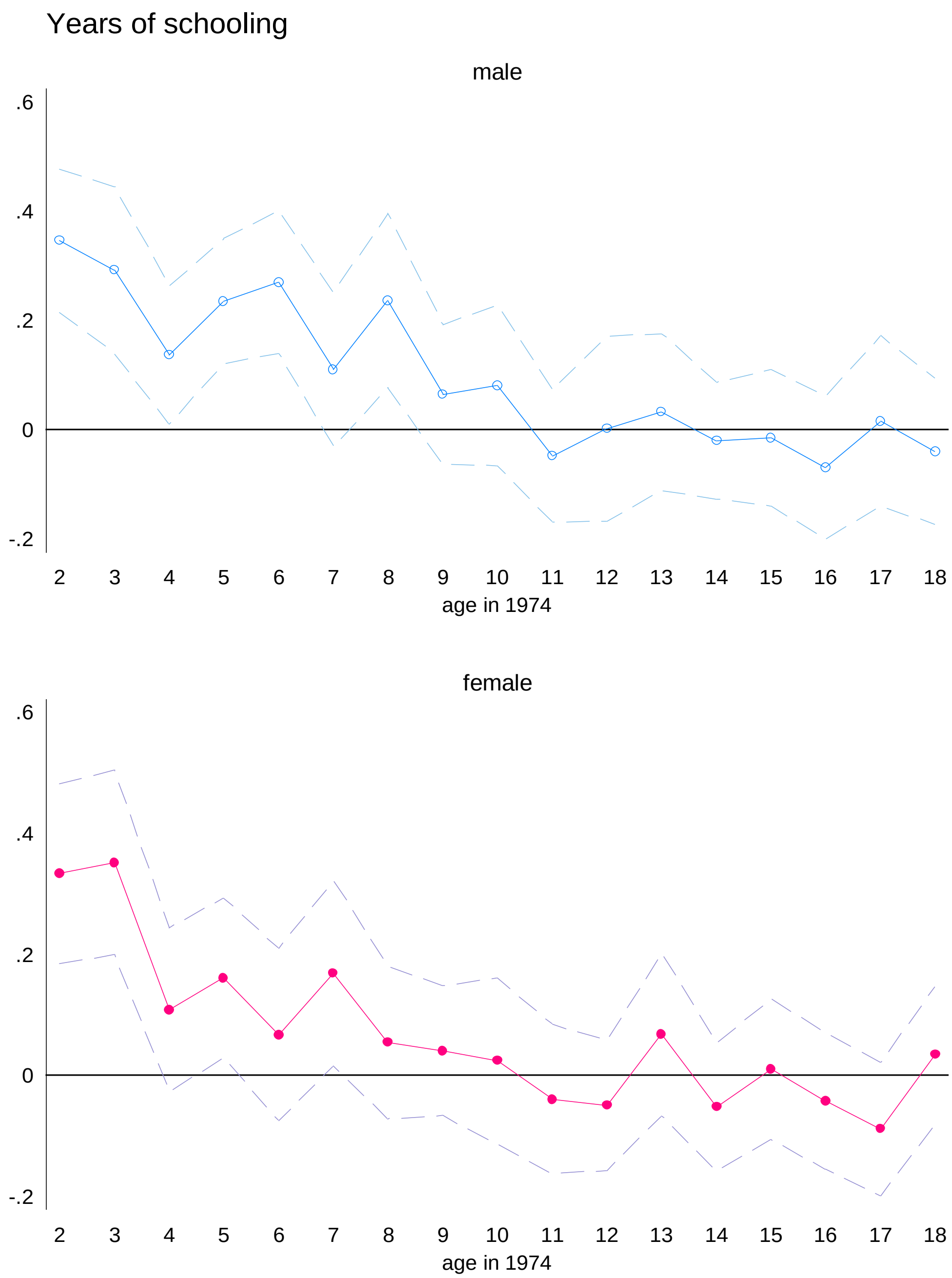
June 2019

A. Online Appendix Figures and Tables



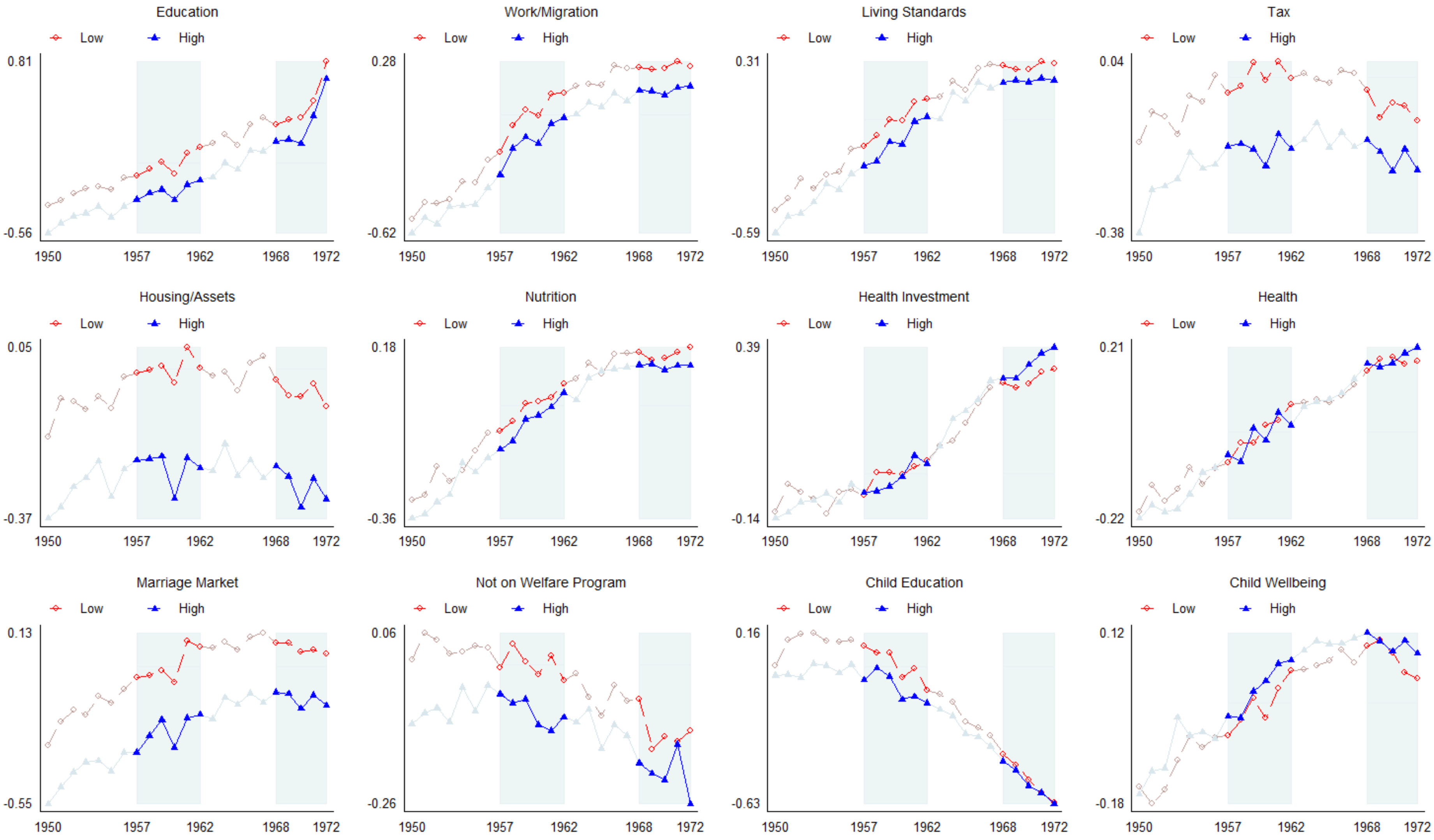
Appendix Figure A.1. Spatial distribution of schools constructed per 1,000 children between 1973 and 1979

Notes: Number of schools constructed between 1973 and 1979 and children's population in 1971 are obtained from Duflo (2001) and the Indonesian 1971 Census. The legend indicates the range and distribution of schools constructed across the Indonesian archipelago. The numbers in parentheses refer to the number of districts that fall in that range. The total number of districts, 290, reflects their existence in 1993. Districts often split over time; by March 2016, there were 511 districts. In our analyses, we maintain the 1993 district boundaries to allow matching with Duflo (2001)'s school construction data.



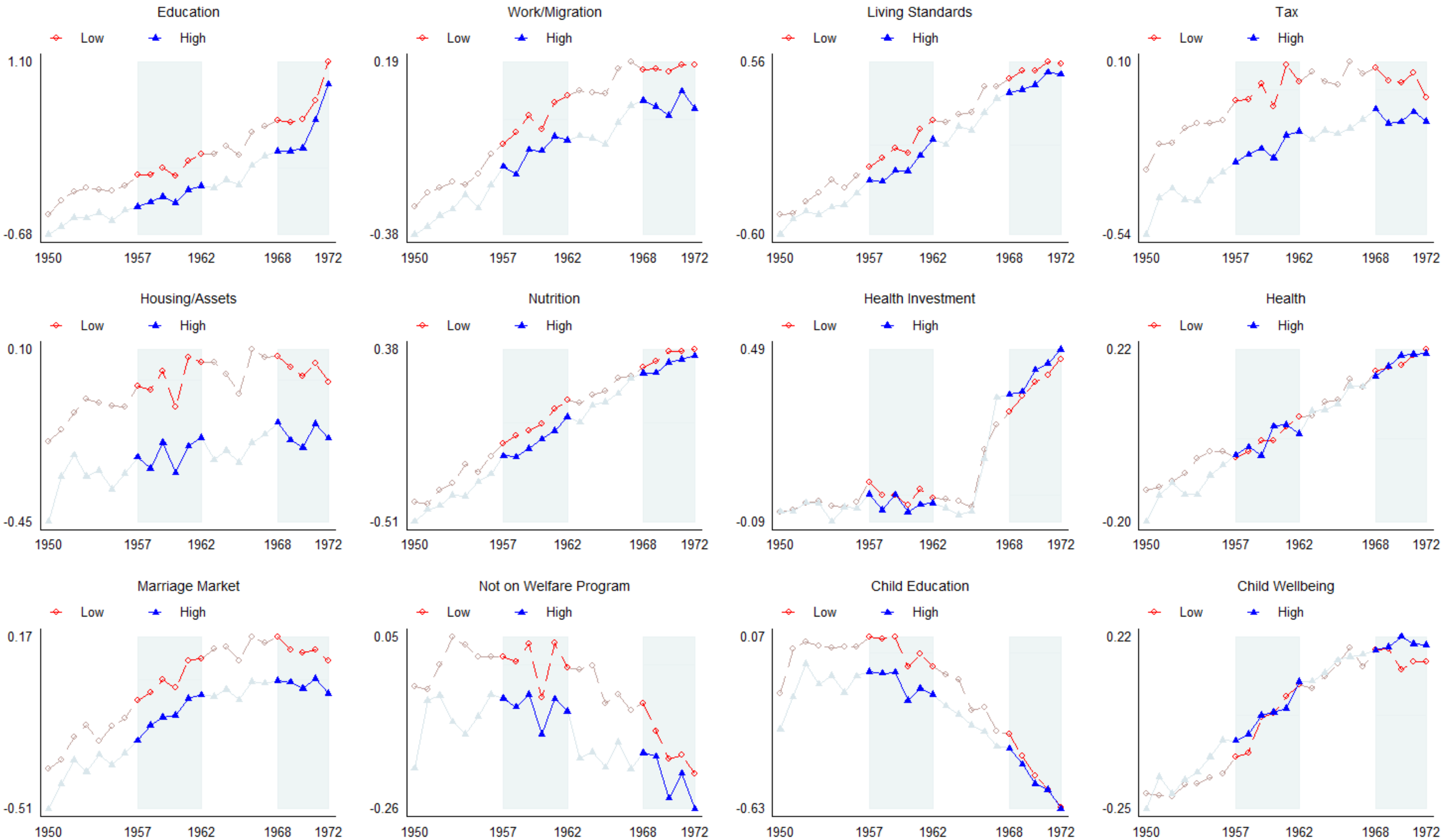
Appendix Figure A.2. Effect of school construction on first generation individual's years of schooling by age in 1974

Notes: Sample is restricted to individuals aged 2-24 in 1974 (born between 1950 and 1972). Each dot represents the interaction coefficient of the number of INPERS primary schools constructed in one’s birth district and an age in 1974 dummy. The age group 19-24 is omitted from the regression. The dashed lines represent 95% confidence bands.



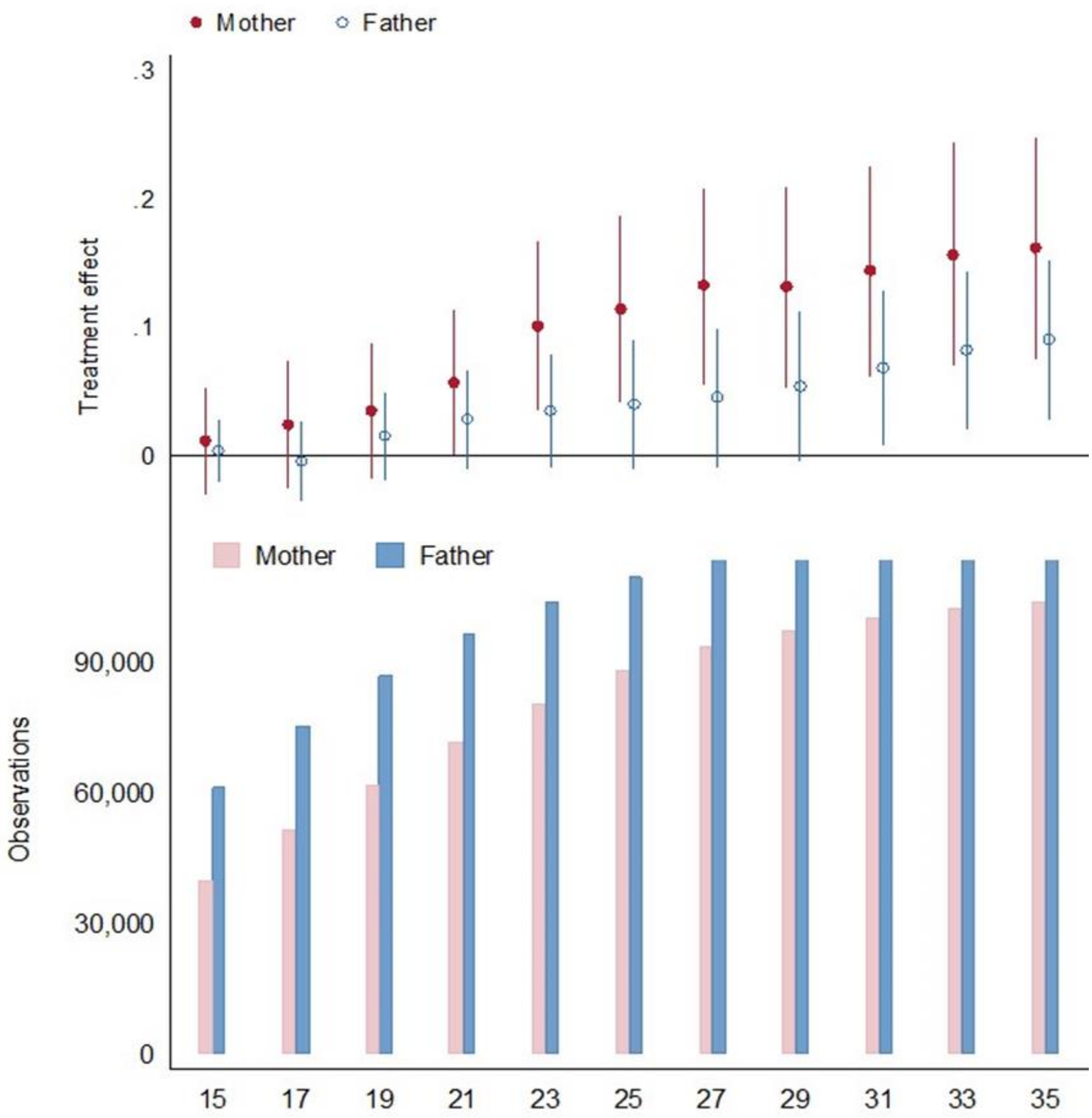
Appendix Figure A.3. Parallel trends of indexes of men born in high and low-intensity school construction districts

Note: Indexes are defined as in Figure 1. Light blue shades highlight our main experiment: contrasting old and young cohorts. “High” (“Low”) indicates high-intensity (low-intensity) school construction districts, as defined in Duflo (2001), in which the residual of a regression of the number of schools on the number of children is positive (negative).



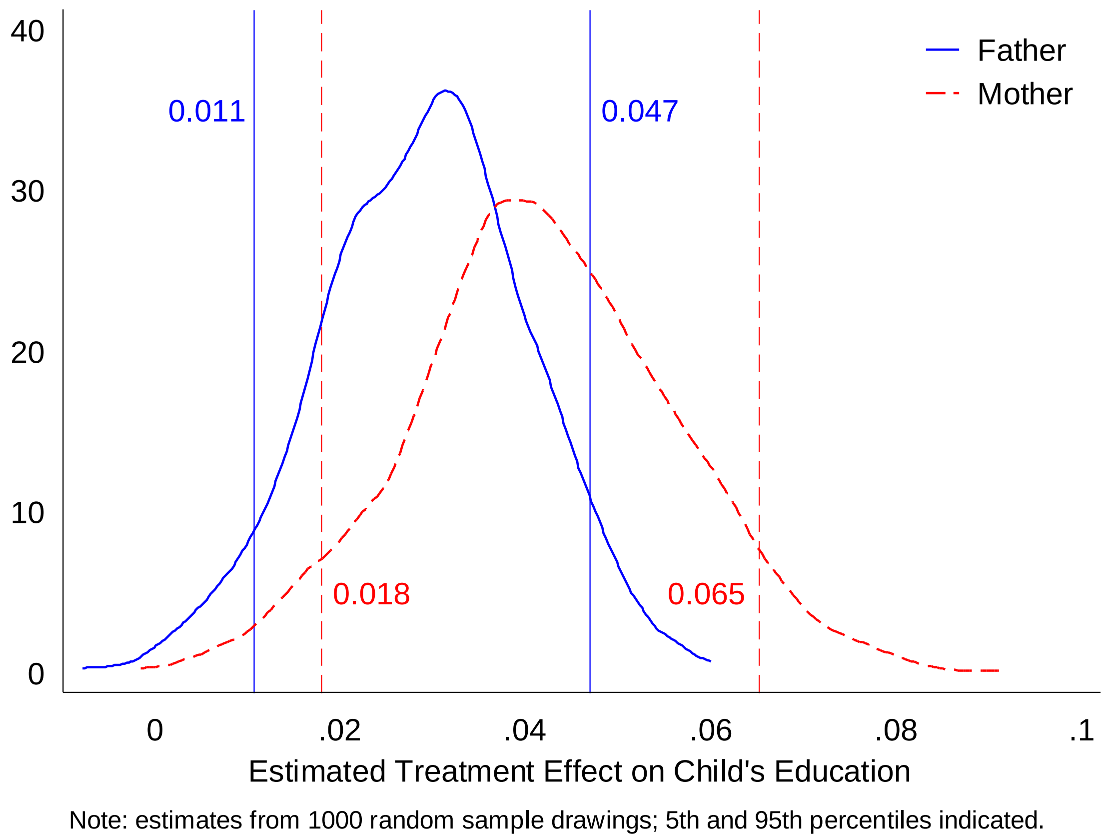
Appendix Figure A.4. Parallel trends of indexes of women born in high and low-intensity school construction districts

Note: Indexes are defined as in Figure 1. Light blue shades highlight our main experiment: contrasting old and young cohorts. "High" ("Low") indicates high-intensity (low-intensity) school construction districts, as defined in Duflo (2001), in which the residual of a regression of the number of schools on the number of children is positive (negative).



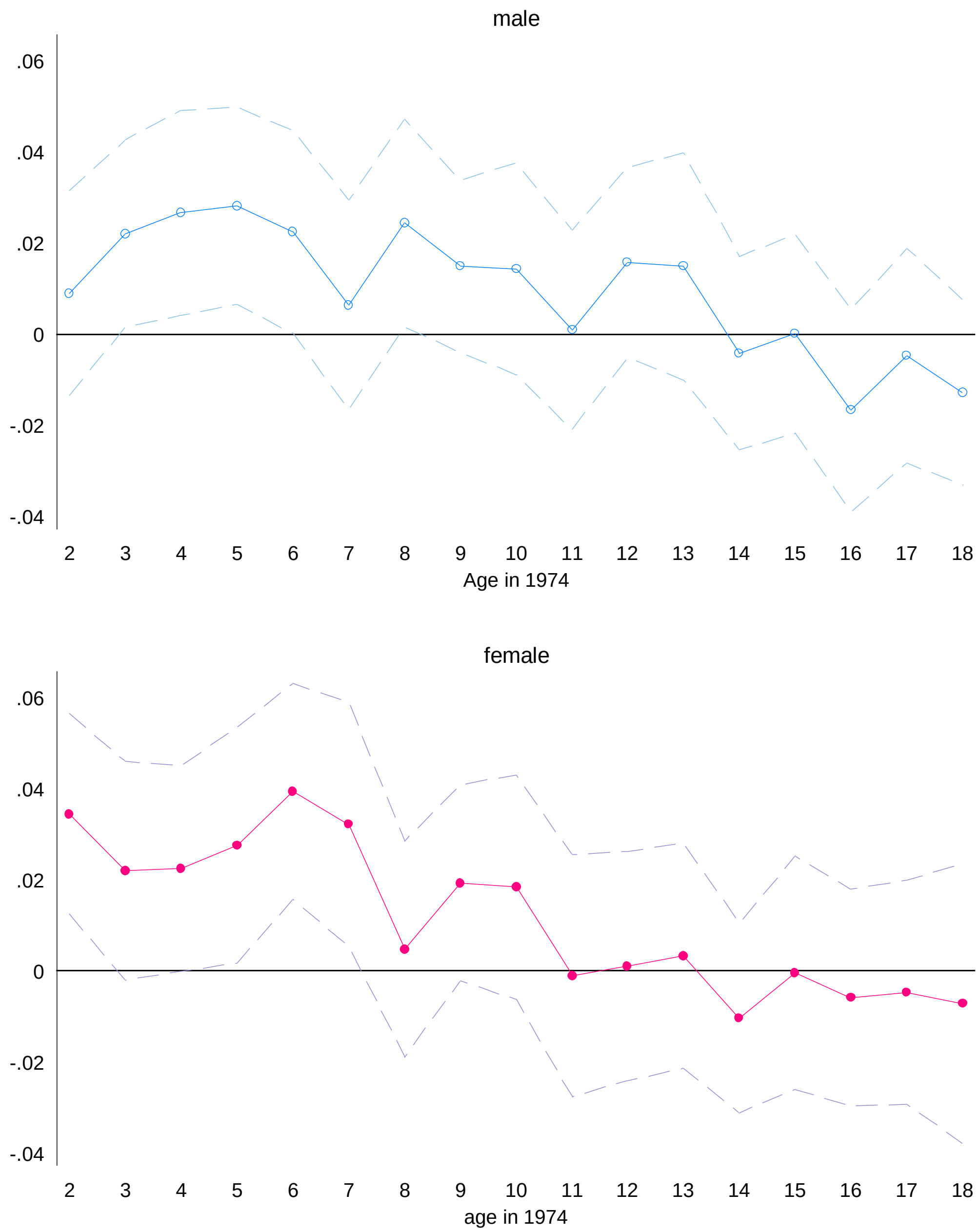
Appendix Figure A.5. Effect of school construction on second-generation's years of schooling, using alternative upper-bound age restrictions

Notes: In the top panel, treatment effects indicate the effect of one additional school constructed per 1,000 children in the mother's or father's birth district on the years of schooling for second-generation individuals. Each dot represents a coefficient in a separate regression. We show estimated regression coefficients and their respective 95% confidence intervals. Sample is restricted to children from age 5 up to the value on the x-axis. Bottom panel shows the number of additional observations added to each regression when the upper-age limit is increased.



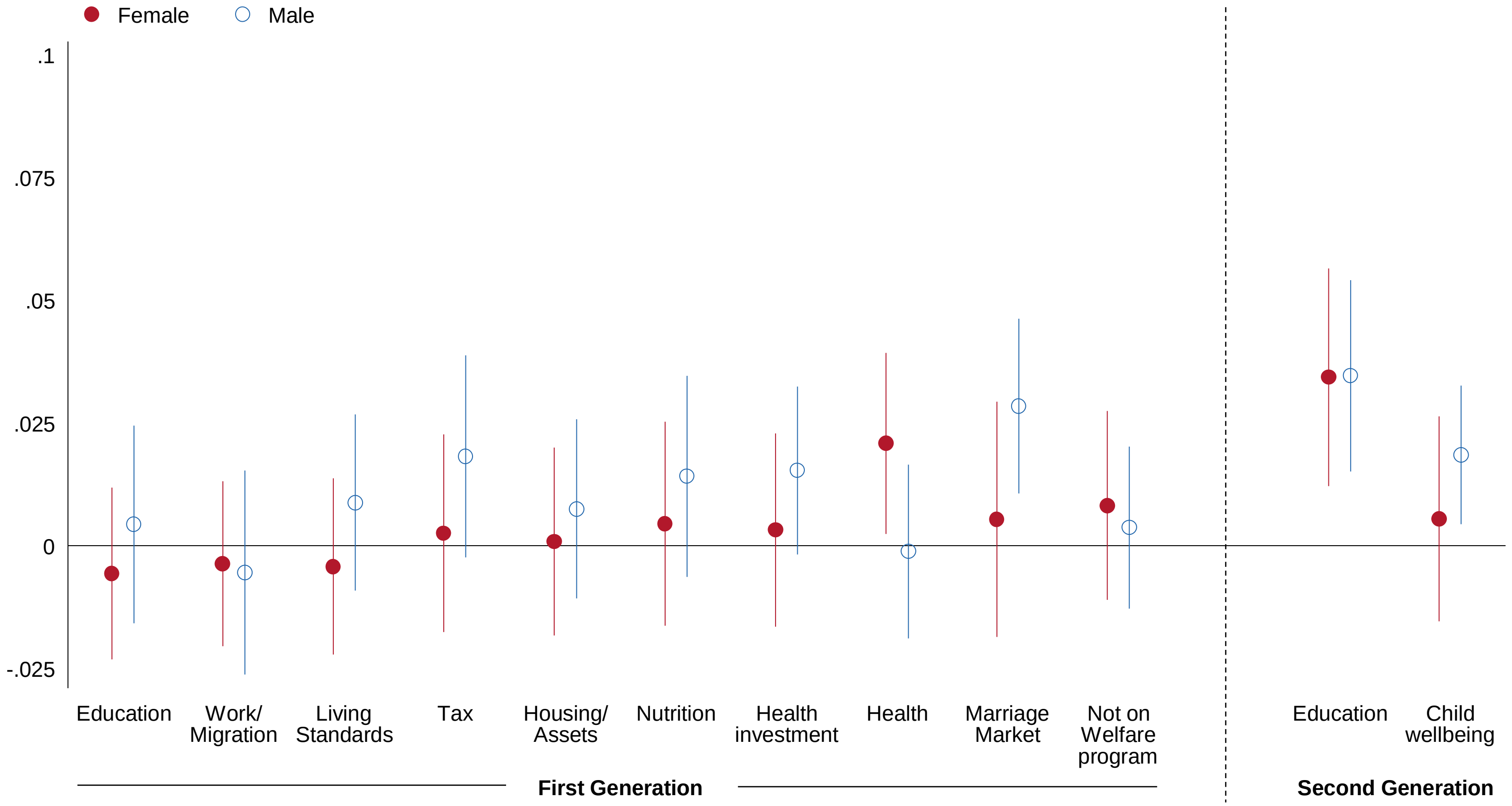
Appendix Figure A.6. Distribution of estimated treatment effects on second-generation's years of schooling from simulated exposure assignment

Note: To address the selection issue about co-resident second-generation children observed in the Susenas data, we use the IFLS to obtain the fraction of children at each age who are born to old and young cohort parents among all children no longer living with their parents. We then use these IFLS-based fractions to randomly assign non-co-resident children at each age in the Susenas data to either old or young cohort parents and to exclude the others from the regression. We then simulate this randomization assignment procedure 1,000 times and estimate the second-generation years of schooling regression. This figure plots the density distribution of estimated coefficients from these 1,000 repetitions for father and mother's exposure to the school construction. Solid lines indicate the distribution of father's effects and dashed lines indicate the distribution of mother's effects. Vertical lines indicate the 5th and 95th percentiles.



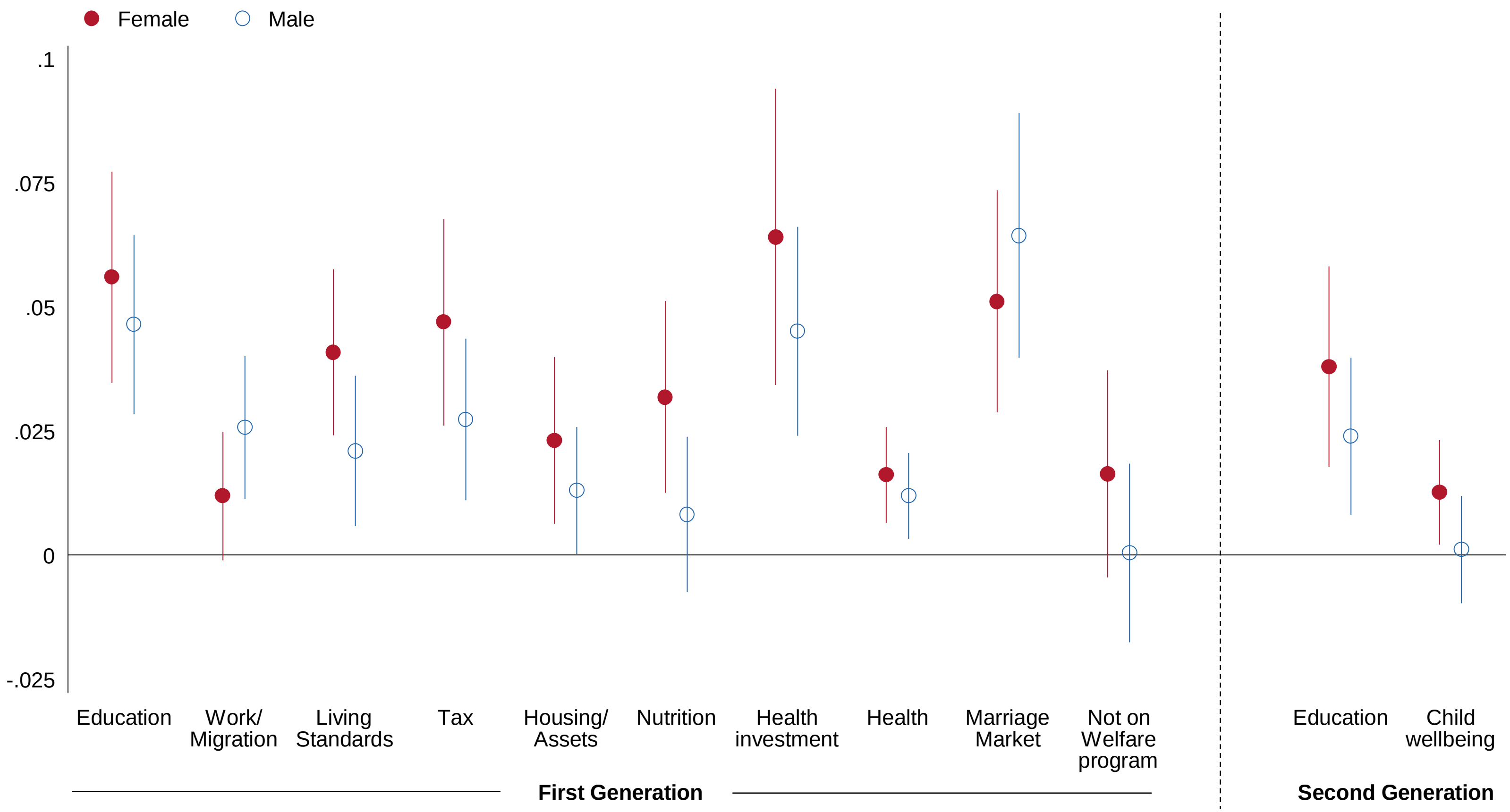
Appendix Figure A.7. Effect of school construction on household expenditures by age in 1974

Notes: Sample is restricted to individuals aged 2-24 in 1974 (born between 1950 and 1972). Each dot represents the interaction coefficient of the number of INPERS primary schools constructed in one's birth district and an age in 1974 dummy. The age group 19-24 is omitted from the regression. The dashed lines represent 95% confidence bands.



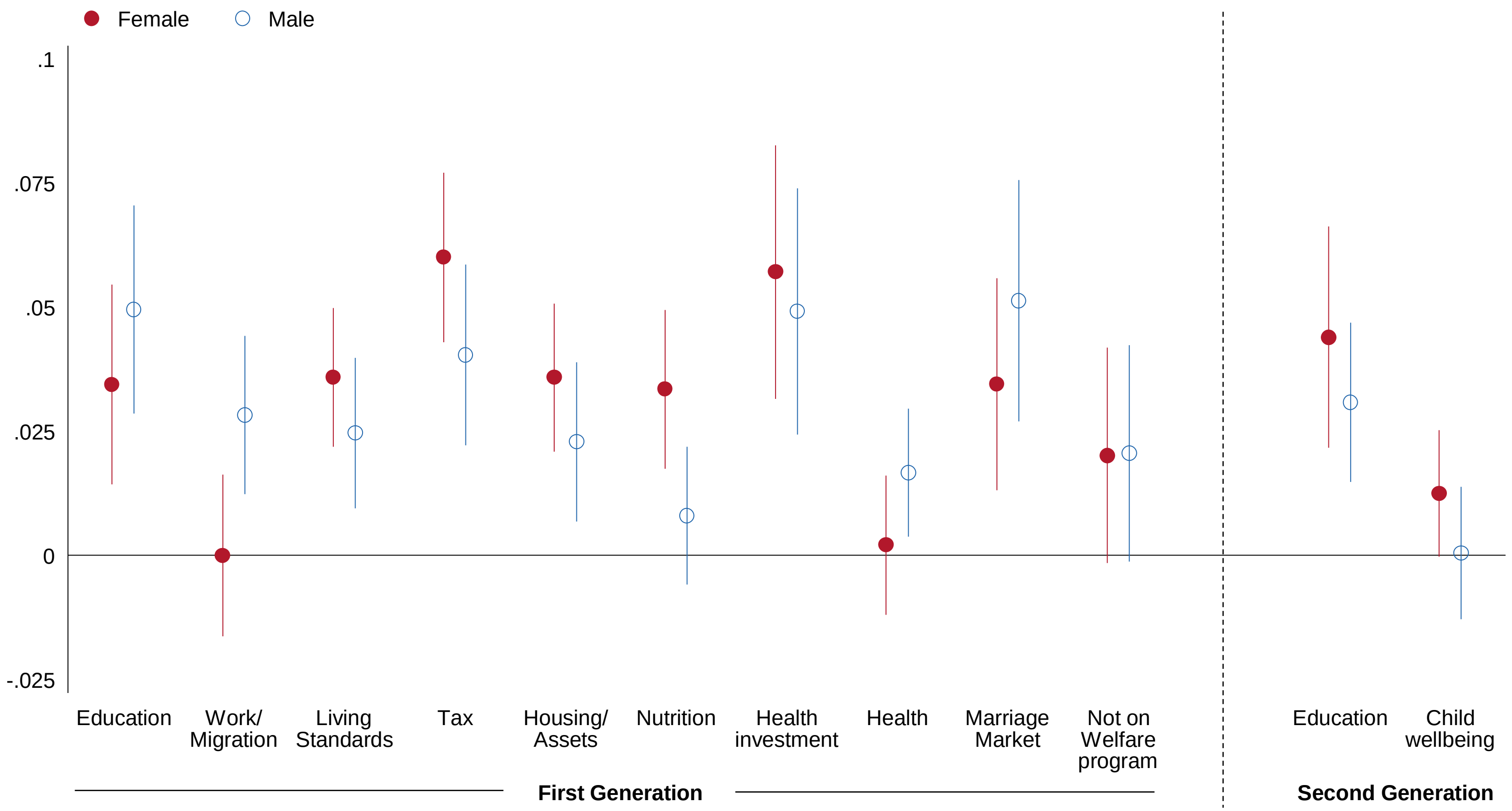
Appendix Figure A.8. Placebo effect of school construction on indexes of long-run outcomes for individuals too old to benefit from primary school construction

Notes: Similar to Figure 1, where we compare individuals born between 1957-1962 (old cohort) and 1968-1972 (young cohort), we now estimate a placebo regression by restricting the sample to individuals born between 1950-1956 (an older cohort) and 1957-1962 (old cohort). Each dot represents the interaction coefficient of the number of INPRES schools built between 1973 and 1979 in one's birth district and a dummy for being born between 1957 and 1962. The solid lines represent 95% confidence bands. This figure serves as a placebo test since the old cohort was too old to be enrolled in primary school when the schools were constructed, and thus could not benefit from the school construction. The individual outcomes making up the index for each family are listed in Tables 1-5 and Appendix Table A.1.



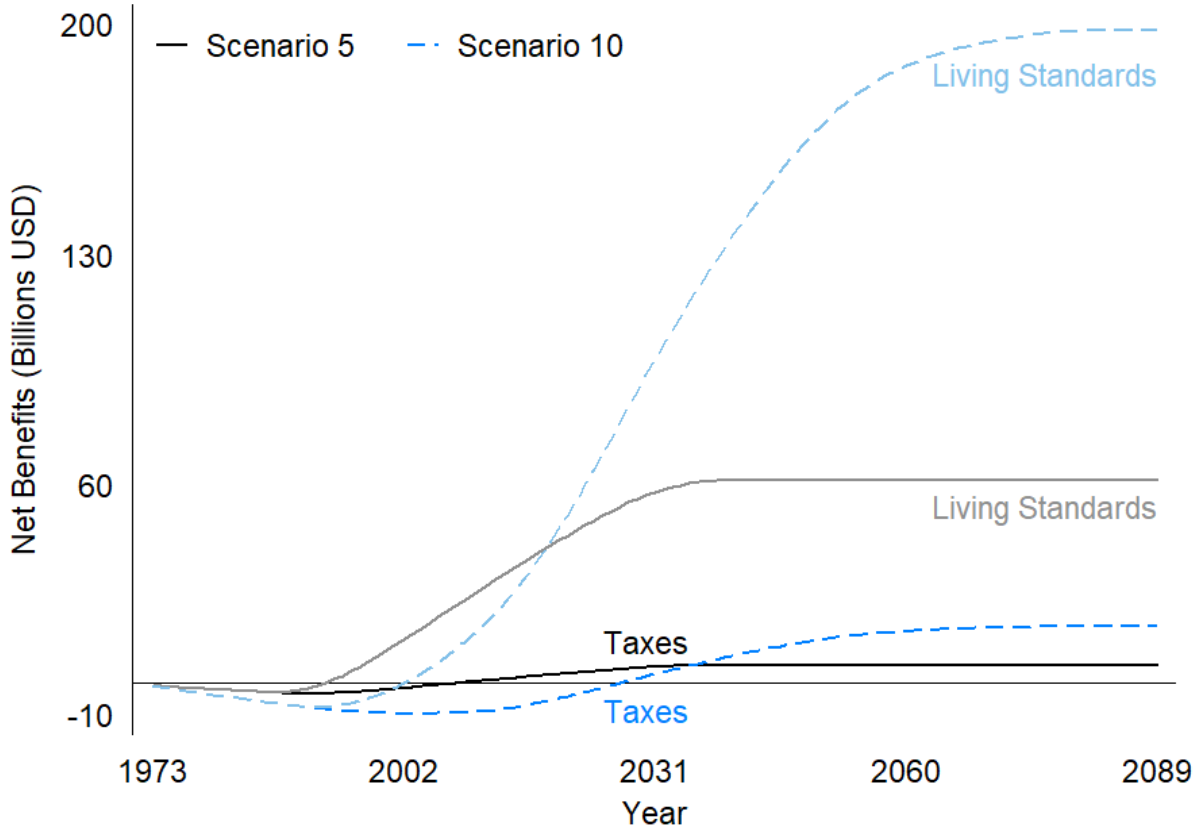
Appendix Figure A.9. Effect of school construction on indexes of long-run outcomes extending the sample to all individuals born between 1950 and 1980

Notes: Similar to Figure 1, but regressions now include all individuals born between 1950 and 1980. Each dot represents the interaction coefficient of the number of INPRES schools built between 1973 and 1979 in one's birth district and a dummy for being born between 1968 and 1980. The solid lines represent 95% confidence intervals. The individual outcomes making up the index for each family are listed in Tables 1-5 and Appendix Table A.1.



Appendix Figure A.10. Effect of school construction on indexes of long-run outcomes using alternative control variables

Notes: Similar to Figure 1, but regressions now exclude the interaction of birth year dummies and water and sanitation programs from the control variables. Each dot represents the interaction coefficient of the number of INPRES schools built between 1973 and 1979 in one's birth district and a dummy for being born between 1968 and 1972. The solid lines represent 95% confidence intervals. The individual outcomes making up the index for each family are listed in Tables 1-5 and Appendix Table A.1.



Appendix Figure A.11. Discounted net benefits of school construction in Indonesia

Note: We plot net benefits (the difference in discounted total benefits and total costs) over time. Benefits are either tax receipts collected by the government or improved living standards of the citizens. Net benefits are reported in billions of 2016 USD. We present two scenarios using the parameters from the cost-benefit model in column (5) and column (10) of Table 9. Solid lines indicate net benefits—in taxes and living standards—under Scenario 5. Dashed lines indicate net benefits under Scenario 10.

Appendix Table A.1. Description and treatment effects of variables used to construct summary indexes

		Mean / SD		Treatment effect on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Panel A: Tax					
Total (Rp10k)	Self-reported tax expenditures include the following components and "other"	4.749	4.552	0.078***	0.123***
		(11.43)	(10.74)	(0.017)	(0.019)
				[0.000]	[0.000]
Land & building (Rp10k)	Taxes on land and/or building ownerships.	0.465	0.506	0.041*	0.075***
		(2.742)	(2.446)	(0.022)	(0.021)
				[0.120]	[0.000]
Vehicle (Rp10k)	Motorized and non-motorized vehicle license fees.	3.610	3.398	0.154***	0.267***
		(8.076)	(7.821)	(0.047)	(0.052)
				[0.003]	[0.000]
Local (Rp10k)	Levies/retributions; examples include neighborhood/citizen associations, garbage, security, cemetery, parking fees.	0.469	0.468	0.048	0.082**
		(2.259)	(2.074)	(0.033)	(0.039)
				[0.148]	[0.036]
Tax index				**0.036*** **	**0.059*** **
				(0.009)	**(0.010)**
Panel B: Housing/Assets					
Urban	Indicator for residing in an urban area.	0.425	0.438	-0.001	0.002
		(0.494)	(0.496)	(0.004)	(0.004)
				[0.822]	[0.576]
Rent equivalent (Rp10k)	Actual monthly rent, if house is rented, or estimated rent value if house is owned or leased by the employer.	42.99	43.08	0.012	0.028***
		(56.34)	(56.57)	(0.008)	(0.008)
				[0.293]	[0.001]
Floor area (sq.-m)	House's floor area in squared meters.	79.894	81.355	1.229**	1.480***
		(58.651)	(59.726)	(0.566)	(0.510)
				[0.119]	[0.011]
Utilities (Rp10k)	Expenditure on electricity, water, gas, and kerosene.	15.71	15.73	0.051**	0.085***
		(20.98)	(21.80)	(0.022)	(0.024)
				[0.102]	[0.002]
Asset index	PCA index on binary ownerships of motorcycle, car, home phone, computer, TV, jewelry, refrigerator, water heater, LPG gas tube, boat, and air conditioner	-0.035	-0.069	0.030*	0.040**
		(1.868)	(1.882)	(0.017)	(0.015)
				[0.223]	[0.020]
Housing/Assets Index				**0.021** **	**0.035*** **
				(0.009)	**(0.009)**
Panel C: Nutrition					
Calories	Household's accounts of units of food consumed in the past week (e.g. 5 kg of rice) are converted into nutritional intake by the Central Statistics Agency. Following their procedure, we convert the weekly intake to monthly intake. In the regressions, we apply an inverse hyperbolic transformation for reasons discussed above. The mean of calories intake is reported in 1 kcal increments. The means of protein, fat, and carbohydrate intakes are reported in 1 kg increments.	260.92	249.70	0.005	0.018***
		(106.00)	(109.83)	(0.004)	(0.005)
				[0.301]	[0.001]
Protein		7.116	6.831	0.006	0.018***
		(3.254)	(3.330)	(0.005)	(0.005)
				[0.301]	[0.001]
Fat		6.074	5.810	0.011**	0.023***
		(3.110)	(3.150)	(0.004)	(0.006)
				[0.061]	[0.000]
Carbohydrates		40.87	39.04	0.005	0.017***
		(17.73)	(18.24)	(0.004)	(0.005)
				[0.301]	[0.001]
Nutrition index				**0.014**	**0.039*** **
				(0.009)	**(0.010)**

Outcome	Description	Mean / SD: Men	Mean / SD: Women	Treatment effect on: Men	Treatment effect on: Women
Panel D: Health Investments					
Total health expenditure (Rp10k)	Total monthly household health expenditures, which aggregates curative, medicine, and preventive health expenditures	7.517 (34.13)	7.961 (35.24)	0.071* (0.038) [0.114]	0.055 (0.041) [0.185]
Preventive measures (Rp10k)	Consist of pregnancy checks, immunizations, medical check-ups, family planning, and other expenditures, e.g., vitamins, massage, gym memberships	0.744 (3.225)	0.671 (3.135)	0.242*** (0.068) [0.002]	0.193*** (0.071) [0.013]
Family planning (Rp10k)	A sub-category under preventive health expenditures, which includes costs of contraceptives and consultations	0.286 (0.872)	0.219 (0.856)	0.321*** (0.061) [0.000]	0.226*** (0.071) [0.008]
Private hospital (Rp10k)	A sub-category under curative health expenditures and is distinct from expenditures on public hospitals, clinics, and traditional healers	2.101 (20.72)	2.200 (22.27)	0.048** (0.023) [0.114]	0.075*** (0.024) [0.008]
Health insurance (Rp10k)	Health insurance is distinct from life, accidental, vehicle, and house insurances	3.821 (16.43)	3.635 (14.05)	0.083 (0.055) [0.134]	0.142*** (0.048) [0.009]
Health investment index				**0.065*** (0.015)**	**0.063*** (0.016)**
Panel E: Health					
No health complaint	Self-reported indicator taking the value of 1 if did not experience a health complaint in the past month.	0.690 (0.463)	0.646 (0.478)	0.004 (0.004) [0.352]	0.003 (0.004) [0.771]
Non-disrupted days	Self-reported number of days in the past month (maximum of 30 days) that a health complaint did <u>not</u> disrupt daily activities.	28.85 (4.012)	28.80 (4.064)	0.042 (0.028) [0.266]	0.027 (0.033) [0.771]
No severe health complaint	Self-reported indicator taking the value of 1 if did not experience a severe health complaint in the past month.	0.951 (0.216)	0.949 (0.221)	0.005*** (0.002) [0.025]	-0.001 (0.002) [0.771]
Health outcomes index				**0.015** (0.007)**	**0.004 (0.007)**
Panel F: Welfare Program					
Cash Transfer	Unconditional cash transfer to compensate for the removal of gas price subsidy for poor households	0.041 (0.197)	0.039 (0.194)	-0.002 (0.002) [0.742]	-0.001 (0.002) [0.914]
Rice for Poor	Monthly rice allowance for poor households	0.392 (0.488)	0.406 (0.491)	0.002 (0.004) [0.850]	-0.009* (0.005) [0.200]
Poor Student's Assistance	Cash transfer conditional on school enrollment	0.056 (0.363)	0.127 (0.333)	-0.001 (0.004) [0.850]	0.000 (0.004) [0.914]
Social Protection Card	Card provided to poor households, which entitles them to social welfare programs mentioned above	0.186 (0.389)	0.180 (0.384)	-0.001 (0.004) [0.850]	-0.000 (0.004) [0.914]
Welfare program non-participation index	For the index, we reverse the sign for the 4 welfare programs to indicate a positive outcome.			**0.006 (0.011)**	**0.010 (0.012)**

Outcome	Description	Mean / SD		Treatment effect on:	
		Men	Women	Men	Women
Panel G: Child Wellbeing					
Non-work days	Number of days not worked in the past week by the child unconditional on work, i.e. 7 for non-working individuals	5.317 (2.670)	4.820 (2.865)	0.044** (0.021) [0.136]	0.031 (0.019) [0.463]
Non-work hours	Number of hours not worked in the past week by the child unconditional on work, i.e. 168 for non-working individuals	156.68 (19.70)	153.05 (21.60)	0.299* (0.157) [0.173]	0.215 (0.151) [0.463]
No health complaint	Self-reported indicator defined as 1 if child did not experience a health complaint in the past month	0.797 (0.402)	0.823 (0.382)	-0.008*** (0.003) [0.042]	0.004 (0.003) [0.463]
Non-disrupted days	Self-reported number of days in the past month (maximum of 30 days) that a health complaint did not disrupt child's daily activities	29.49 (2.086)	29.55 (2.067)	-0.026* (0.016) [0.198]	0.007 (0.015) [0.893]
No severe health complaint	Self-reported indicator defined as 1 if child did not experience a severe health complaint in the past month	0.978 (0.147)	0.980 (0.140)	-0.000 (0.001) [0.751]	-0.000 (0.001) [0.893]
Second-generation wellbeing index				**-0.005** **(0.007)**	**0.014**** **(0.007)**

Notes: Panel headings correspond to the summary indexes in Figure 1 not shown earlier, and the listed outcomes refer to the outcomes used to construct the summary index. Treatment effects report regression coefficients of young cohort dummy interacted with the number of schools constructed in region of birth. Treatment effects for the summary indexes are shown in bold. Standard errors clustered at region of birth are shown in parentheses. FDR q-values are computed over variables under the same heading and are shown in square brackets. Stars denote statistical significance at 1, 5, and 10% levels based on regular p-values. All expenditure values refer to the household's average monthly expenditure; means are reported in 10,000 Indonesian Rupiah (IDR) increments. In 2016, the average daily exchange rate was 1 USD=13,308 IDR. In the regressions, we apply an inverse hyperbolic sine transformation to the nominal values since expenditure data tends to be skewed and a log transformation would not be defined for zero expenditures. For small values, inverse hyperbolic sine function approximates logarithmic function. Similarly, our expenditure estimates can be interpreted as percentage changes.

Appendix Table A.2. Spouse's characteristics as mediators of the effect of school construction on first generation's living standards

	Dependent Variable: Living standards									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mediator:	None	Years of Schooling	Completed Primary	Literate	Work	Formal worker	Non-agriculture sector	Migrant	No health complaint	All
Panel A: Men										
Schools constructed * Young cohort	0.021***	0.009	0.007	0.010	0.020***	0.019**	0.017**	0.018**	0.020***	0.008
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Mediator		0.062***	0.337***	0.374***	0.002	0.371***	0.480***	0.273***	-0.009	
		(0.001)	(0.009)	(0.014)	(0.008)	(0.014)	(0.010)	(0.011)	(0.006)	
Observations	68,687	64,416	64,416	64,416	64,416	39,545	39,545	64,416	64,416	39,545
Mean	8.011	8.068	8.068	8.068	8.068	8.007	8.007	8.068	8.068	8.007
Panel B: Women										
Schools constructed * Young cohort	0.032***	0.016**	0.015**	0.017**	0.023***	0.018**	0.018**	0.021***	0.023***	0.012
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Mediator		0.063***	0.351***	0.380***	-0.057***	0.319***	0.412***	0.293***	0.030***	
		(0.001)	(0.009)	(0.015)	(0.013)	(0.012)	(0.009)	(0.012)	(0.007)	
Observations	66,249	55,449	55,449	55,449	55,449	50,884	50,884	55,449	55,449	50,884
Mean	7.152	7.313	7.313	7.313	7.313	7.323	7.323	7.313	7.313	7.323

Notes: Each column shows a regression of the first generation's living standards on exposure to the school construction program and includes a potential mediator variable. These mediator variables (as indicated by the column heading) are spouse's characteristics in Table 4. Regressions are as in row 1 of Table 3. Effect of program exposure are the regression coefficients of father or mother's young cohort dummy interacted with the number of schools constructed in father or mother's district of birth. All regressions control for parent's district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at parent's district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values.

Appendix Table A.3. Effect of school construction on spouse's education

		Mean / SD		Effect of Program Exposure on:	
Outcome	**Description**	**Men**	**Women**	**Men**	**Women**
Spouse's years of schooling	Based on highest education level and grade attended. Standard durations of study are assumed; grade retentions are not counted	7.635 (4.081)	7.426 (4.192)	0.180*** (0.046) [0.000]	0.116*** (0.043) [0.028]
Spouse's years of schooling percentile	Spouse's years of schooling percentile (relative to individuals born between 1950-1980)	40.62 (28.49)	32.44 (29.29)	0.625** (0.303) [0.115]	0.448 (0.298) [0.133]
Education gap	Husband's years of schooling minus wife's years of schooling	0.432 (3.707)	0.112 (3.600)	0.093* (0.049) [0.115]	-0.078** (0.038) [0.086]
Education percentile gap	Husband's education percentile minus wife's education percentile	36.96 (31.18)	34.39 (30.08)	0.461 (0.402) [0.252]	-0.738** (0.322) [0.065]

Notes: Effect of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. FDR q-values are computed over all 4 outcomes and are shown in square brackets. FDR q-values indicate the probability of false positives among *significant* tests. Spouse's characteristics and education gaps are defined for household heads and spouses in the sample. There are 64,422 and 55,468 observations for men and women, respectively, because it is set to missing if the spouse does not currently live in the household (divorced, widowed).

Appendix Table A.4. Effect of school construction on second-generation's years of schooling on various samples

	Susenas	Susenas with Extreme Assumptions		IFLS		
	(1)	(2) Assume Not Exposed	(3) Assume Exposed	(4) All	(5) Stayers	(6) Movers
Panel A: Father						
Schools constructed *	0.097***	0.021	0.000	0.103	0.030	-0.020
Young cohort	(0.032)	(0.016)	(0.014)	(0.104)	(0.109)	(0.251)
Observations	120,838	644,675	644,675	6,186	4,048	2,138
Mean	7.967	7.731	7.731	7.807	6.434	10.396
Panel B: Mother						
Schools constructed *	0.169***	0.052***	0.030*	0.300**	0.539***	0.126
Young cohort	(0.045)	(0.017)	(0.017)	(0.147)	(0.128)	(0.239)
Observations	105,523	644,675	644,675	7,227	3,756	3,471
Mean	8.854	7.731	7.731	9.038	8.097	10.034

Note: Column (1) is from Table 5. Column (2) and (3) estimate extreme bounds in which all non-co-resident children aged 0-40 are assumed to have parents who are either exposed or not exposed (Manski, 1990). Columns (4)-(6) use the IFLS 2014 Round 5 data. We match parents to their co-resident children ("Stayers") found in the household roster and to their non-co-resident children ("Movers") in the respective module.

Appendix Table A.5. Effect of school construction on first generation's years of schooling (extended cohort definitions)

	(1)	(2)	(3)	(4)	(5)
Cohorts Included:	1957-1962 and 1968-1972	… + 1950-1956	… + 1963-1967	… + 1973-1980	1950-1980
Panel A: Male					
Schools constructed * Young cohort	0.268***	0.267***	0.221***	0.211***	0.172***
	(0.047)	(0.039)	(0.037)	(0.044)	(0.032)
Observations	72,367	98,895	98,781	138,617	197,951
Mean	8.022	7.500	7.938	8.478	8.047
Panel B: Female					
Schools constructed * Young cohort	0.234***	0.219***	0.209***	0.245***	0.210***
	(0.042)	(0.044)	(0.039)	(0.044)	(0.045)
Observations	71,423	97,268	99,843	140,142	200,644
Mean	7.105	6.496	6.901	7.790	7.194

Notes: Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. Column (1) sample is restricted to individuals born in the sample period 1957-1962 (old cohort) and 1968-1972 (younger cohort) and is the sample used in the analysis in the rest of the paper. Columns (2) to (5) extend the sample as indicated in the column headings. Panel A looks only at males and Panel B only at females. School constructed denotes the number of INPRES schools constructed per 1,000 children in one's birth district. Young cohort is an indicator defined as 1 for being born after 1967.

Appendix Table A.6. Effect of school construction on first generation's household expenditures (various transformations)

	Total expenditure				Education expenditure			
	IHS Total	Log	Nominal	IHS Per-capita	IHS Total	Log	Nominal	IHS Per-capita
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Father								
Schools constructed *	0.021***	0.021***	9.882***	0.016**	0.160**	0.013	0.309	0.140**
Young cohort	(0.007)	(0.007)	(3.628)	(0.007)	(0.064)	(0.010)	(0.309)	(0.056)
Observations	68,687	68,687	68,687	68,687	68,687	48,123	68,687	68,687
Mean	391.649	391.649	391.649	391.649	13.971	13.971	13.971	13.971
Panel B: Mother								
Schools constructed *	0.032***	0.032***	11.022***	0.018***	0.193**	-0.010	-0.191	0.167**
Young cohort	(0.007)	(0.007)	(2.583)	(0.007)	(0.076)	(0.014)	(0.383)	(0.067)
Observations	66,249	66,249	66,249	66,249	66,249	39,492	66,249	66,249
Mean	375.616	375.616	375.616	375.616	12.202	12.202	12.202	12.202

Notes: Effect of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. All regressions control for district of birth and cohort of birth fixed effects, children's population and enrollment in 1971, and water and sanitation program intensities that vary by district of birth interacted with birth year dummies. Robust standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on unadjusted p-values. All expenditure values are defined at the household level and refer to the household's average monthly expenditure. Nominal values are reported in 10,000 Indonesian Rupiah (IDR) increments. In 2016, the average daily exchange rates was 1 USD=13,308 IDR. Columns (1)-(4) examine total household expenditure; columns (5)-(8) examine education expenditure. Inverse hyperbolic sine (IHS) transformations are applied to total and per capita household expenditures (columns 1 and 4) and to total and per-capita education expenditures (columns 5 and 8). Log transformations are applied in columns (2) and (6). Column (1) and (5) are the preferred specification and are the same as Table 3, rows 1 and 5.

B. Data Appendix

Two critical data issues about the Susenas 2016 survey are relevant for our analysis. First, to estimate the difference-in-differences specification described in Section 3.1, it is necessary to have information about an individual's residence at birth. Current residence could be endogenous to the school construction program as households might move to provide access to schools to their children. Location of birth and location where the individual obtains their education are highly correlated.[1] However, birth location is not endogenous with respect to the school construction since all of the individuals in the analysis were born before the program started. Given the importance of knowing where the individual was born, it is unfortunate that most household surveys in Indonesia only provide information about the individual's current location of residence. This lack of information about an individual's birth location is the case for the Indonesian Labor Force Survey (Sakernas), the Indonesian Demographic and Health Surveys (DHS), and many other rounds of the Susenas data, making them unavailable to use to analyze the impacts of the school construction program. However, the Susenas 2016 is one exception to this, as there is information on every individual's district of birth.

Second, it is important that the data include a sufficiently large sample of individuals from these specific birth cohorts (1957-1962 and 1968-1972). The Indonesia Family Life Survey (IFLS) does contain information on each individual's district of residence at birth, thus satisfying the first criteria we outline above. We use the IFLS data to estimate our main difference-in-differences specification exploiting variation across birth cohorts and districts in the number of schools built. The IFLS is a longitudinal survey, and the first round was collected in 1993/1994. Subsequent rounds were collected in 1997, 2000, 2007/2008, and most recently in 2014/2015. Tracking across rounds has been extremely successful, with rates between 92 to 95 percent for each IFLS round (Thomas et. al., 2012). Almost 88 percent of households in survey round one were subsequently interviewed in all of the five survey rounds. In columns 1-3 of Appendix Table B.1 and B.2, we use the most recent survey round collected in 2014/2015 (IFLS 5) and include all individuals interviewed in that round in the regressions. In columns 4-6 of these tables, we begin with the IFLS 5 and then add in any other individuals from the other four rounds who might no longer be present in the final round of the panel survey. We estimate regressions with different control variables to see if that has any influence on the results. Appendix Table B.1 examines years of schooling as the dependent variable, while Appendix Table B.2 examines completed primary. Column 3 (IFLS 5 only) and column 6 (IFLS 5 plus last observed round) in each table correspond with our main results for men and women in Table 1 row 1 (years of schooling) and row 2 (completed primary). We do not observe any statistically significant relationship for men or women between exposure to school construction and increased years of schooling. Similarly, when examining whether women exposed to school construction completed primary school we do not observe any statistically significant relationship. However, for men, that relationship is negative indicating school construction exposure lowers the likelihood of completing primary school, and depending on the control variables included in some cases it is statistically significant.

[1] Based on the IFLS data, almost 92 percent of children at age 12 still live in the same district where they were born (Duflo, 2001). Likewise, in the Susenas 2016, 93.2 percent of children at age 12 live in the same district where they were born.

The IFLS and Susenas data have two key differences that might be relevant to explain this situation. First, the Susenas data is nationally representative covering all 34 provinces and all 511 districts in the country. IFLS is representative of only 83% of the Indonesian population and covers individuals living in 13 out of 27 provinces in the country. Appendix Figure B.1 shows a map of Indonesian districts with the districts shaded in gray indicating which ones the IFLS survey covers. Comparing Appendix Figure A.1 (map of Indonesia indicating the spatial distribution of school constructed per 1,000 children) and Appendix Figure B.1 highlights that many of the districts that had large numbers of schools constructed are not included in the IFLS survey. In column 7 of Appendix Tables B.1 and B.2, we present results using the Susenas 2016 data but restricting the analysis to only those districts covered in the IFLS survey. The coefficients from the regression with this restricted sample with years of schooling as the dependent variable are somewhat smaller (0.194 for men and 0.159 for women) compared to the full sample from Table 1, row 1 (0.268 for men and 0.234 for women), but the results are still statistically significant and economically meaningful. Likewise, the regressions in column 7 of Appendix Table B.1 using completed primary as the dependent variable also show smaller effect sizes compared to the full sample. This is evidence that the different geographic coverage of the IFLS and the Susenas is unlikely to explain the lack of relationship between school construction and years of schooling in the IFLS data (columns 1-6). Second, note that the number of observations in the IFLS regressions for women is only 2,473 if using only IFLS 5 or 2,659 if using IFLS 5 plus the last observed round for any individual.[2] This compares with 71,423 observations for women in the regression using the Susenas data. While the point estimates for women are similar across the two datasets, this difference in sample size could explain the much larger standard errors in the regressions using IFLS data.

[2] Using the extended cohort of individuals born between 1950 and 1980 roughly triples the sample size (for men to 6,586 and 7,050 and for women to 7,138 and 7,640 in the IFLS 5 and IFLS 5 plus last observed round, respectively), but the results are still not statistically significant.

Appendix Table B.1. Effect of school construction on first generation's education using IFLS data

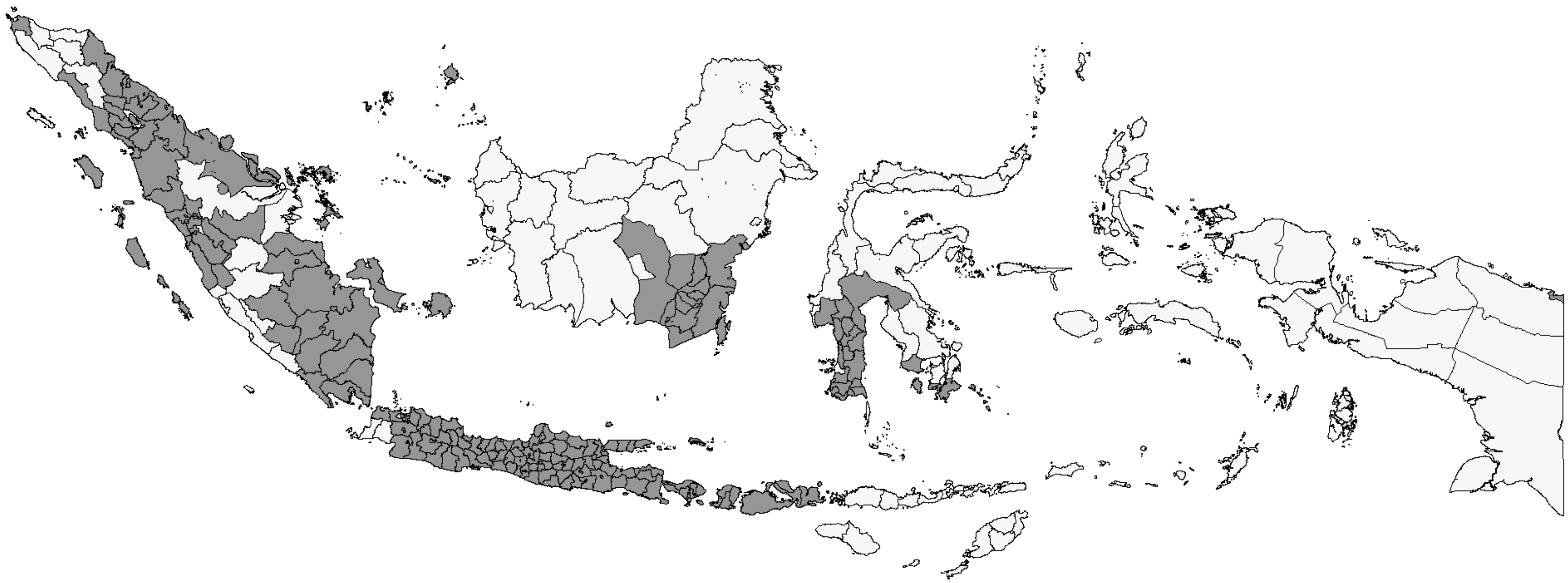
Data source:	IFLS 5 (2014/2015)			IFLS 5 + last observed round			Susenas 2016 restricted to IFLS districts
Dependent variable: Years of schooling	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Male							
Schools constructed * Young cohort	-0.205	-0.166	-0.132	-0.282	-0.231	-0.279	0.194***
	(0.159)	(0.164)	(0.204)	(0.183)	(0.189)	(0.226)	(0.065)
Observations	2,230	2,230	2,230	2,408	2,408	2,408	52,461
Children population in 1971	X	X	X	X	X	X	X
Enrollment in 1971		X	X		X	X	X
Water and sanitation program			X			X	X
Panel B: Female							
Schools constructed * Young cohort	0.052	0.091	0.205	0.211	0.238	0.298	0.159***
	(0.170)	(0.165)	(0.212)	(0.155)	(0.147)	(0.196)	(0.060)
Observations	2,473	2,473	2,473	2,659	2,659	2,659	52,208
Children population in 1971	X	X	X	X	X	X	X
Enrollment in 1971		X	X		X	X	X
Water and sanitation program			X			X	X

Notes: Effects of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. Standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on regular p-values. Columns 1-3 uses Indonesia Family Life Survey data, Round 5 (2014/2015). Columns 4-6 uses IFLS round 5 data plus the observation from the last observed round for any individual not in round 5. Column 7 uses Susenas 2016 data that is restricted to the IFLS districts, which cover 83% of the Indonesian population.

Appendix Table B.2. Effect of school construction on first generation's completed primary using IFLS data

Data source:	IFLS 5 (2014/2015)			IFLS 5 + last observed round			Susenas 2016 restricted to IFLS districts
Dependent variable: Completed primary	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Male							
Schools constructed * Young cohort	-0.013	-0.013	-0.023**	-0.016*	-0.017*	-0.029***	0.015*
	(0.008)	(0.010)	(0.011)	(0.008)	(0.010)	(0.011)	(0.008)
Observations	2,230	2,230	2,230	2,408	2,408	2,408	52,461
Children population in 1971	X	X	X	X	X	X	X
Enrollment in 1971		X	X		X	X	X
Water and sanitation program			X			X	X
Panel B: Female							
Schools constructed * Young cohort	0.013	0.009	0.006	0.010	0.007	0.002	0.029***
	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.014)	(0.009)
Observations	2,473	2,473	2,473	2,659	2,659	2,659	52,208
Children population in 1971	X	X	X	X	X	X	X
Enrollment in 1971		X	X		X	X	X
Water and sanitation program			X			X	X

Notes: Effects of program exposure are the regression coefficients of young cohort dummy interacted with the number of schools constructed in district of birth. Standard errors clustered at district of birth are shown in parentheses. Stars denote statistical significance at 1, 5, and 10% levels based on regular p-values. Columns 1-3 uses Indonesia Family Life Survey data, Round 5 (2014/2015). Columns 4-6 uses IFLS round 5 data plus the observation from the last observed round for any individual not in round 5. Column 7 uses Susenas 2016 data that is restricted to the IFLS districts, which cover 83% of the Indonesian population.



Appendix Figure B.1. Map of Indonesia with districts shaded in gray indicating coverage in Indonesia Family Life Survey (IFLS)

Notes: IFLS survey is representative of 83% of the Indonesian population and covers individuals living in 13 of the 27 provinces in the country. The districts shaded in gray are included in the IFLS household survey, while the Susenas 2016 used in the main analysis in the paper is nationally representative and includes all districts in the country.

C. Cost-benefit Calculations Appendix

Discount rate

World Development Indicators collects real interest rates in Indonesia between 1987 and 2017. It averages 5.77 percent per year. Since it does not extend as far as our sample period in 1973, we assume a constant annual discount rate of 5 percent.

Teachers' salary growth

We first assume there is no real salary growth over the years and use Duflo (2001)'s reported teacher's salary in 1973. Subsequently, we allow for linear growth using teacher's salary observations in 1970 by Daroesman (1972), 1973 by Duflo (2001), Intercensal Surveys 1976 and 1995, and Labor Force Surveys 2000, 2005 and 2010.[3] Teachers are paid for the lifetime of the schools

Lifetime curvature

Individuals' tax payments and living standards generally follow an inverted-U shape, where they peak at around age 40-50. In our Susenas data, we observe individuals at their peak. To model the lifetime curvature of tax payments and living standards, we assume the same average effect on taxes and living standards across ages but different means at different ages. A 20-year old male, for instance, only spends $2,373 annually, compared to the mean in of our observed sample, $3,531, as implied in Table 3.

GDP/capita growth

GDP per-capita growth is obtained from the World Bank's World Development Indicators. We took the average between 1961 and 2017: 3.25 percent per year.

Number of students and teachers per school and recurrent costs/salaries multiplier

We follow Duflo (2001) in assuming 120 students/school, 3 teachers/school, and 25% recurrent administrative costs in addition to teachers' salaries. These imply a class size of 20 students across six grades of primary education and one teacher per grade. The latter is reasonable given that schools often run two sessions per day: morning and afternoon classes.

Individuals start paying taxes after age 18

We first assume that individuals start paying taxes after finishing Upper Secondary education at age 18. We subsequently relax this assumption to age 22, after individuals finish Tertiary education.

School lifetime

Daroesman (1971) and Duflo (2001) report that schools were expected to last for 20 years. We first use this assumption. We subsequently relax this assumption to 40 years because many INPRES schools are still in-use as of 2016.

[3] We drop Duflo (2001)'s reported salary in 1995 because it implies a 9 percent real growth per year and it is much higher than the linear fit would have predicted. It is also higher than observations in 2000, 2005, and 2010.

Life expectancy
World Development Indicators suggest an average of 56.6 years of life expectancy at birth for individuals born between 1968 and 1980. Conditional on making it to primary school age, the life expectancy is likely higher. We assume a life expectancy of 60 years throughout and then relax this assumption in the final column.

Share of men and women in affected cohorts
We construct a weighted average of the treatment effects on men and women. The share of women in the affected cohort is 0.498. For simplicity, we assume an equal share of men and women.

Intergenerational benefits
While we do not directly observe living standard and tax effects on the second generation, we do observe the education benefits on the second generation. We scale the second generation's living standard benefits by the relative ratio of second generation education effect and first generation education effect. We further assume that the first child is born one year after the average age of first marriage (i.e. 25 for fathers and 21 for mothers). Since the average household size in the sample is 4.09, we assume two children per household and that the second child is born 3 years after the first child.