Improving the Early Childhood Environment: Direct and Distributional Effects on Human Capital for Multiple Generations *

Tania Barham, Brachel Champion, Gisella Kagy, and Jena Hamadani †

November 10, 2023

Abstract

This paper examines the long-term and intergenerational effects of improving the early childhood health environment on human capital in Bangladesh. In adulthood, children eligible for health promoting interventions exhibit increased height and reduced short stature, while males achieve higher levels of educational attainment. These finding are concentrated among individuals with the lowest pre-program health endowment, reducing inequality in human capital across generations, and underscoring the program's distributional implications. Intergenerational effects reveal daughters experienced increased height, reduced stunting, and improved cognitive outcomes. The findings suggest that failing to consider distributional and intergenerational effects of programs could lead to underinvestment in children.

JEL Classification: I15, J13, O12

Keywords: human capital, intergenerational, inequality, vaccination, family planning, Bangladesh

*We thank Sajeda Amin, Francisca Antman, Jason Boardman, Ryan Brown, Brian Cadena, Chloe East, Andrew Foster, Chris Jochem, Randall Kuhn, Richard Mansfield, Terra McKinnish, Jane Menken, John Maluccio, Hani Mansour, Noboko Mizoguchi, Mushfiq Mobarak, Colleen Reid, Elisabeth Root, Paul Schultz, Duncan Thomas, Patrick Turner, and Andrea Velásquez. The data collection for this project was generously funded by the National Institutes of Health, Population Research Bureau, the International Initiative for Impact Evaluation, and the CU Population Center. We are deeply grateful to our many staff partners at icddr,b for their partnership and data access. The Matlab Health and Socioeconomic Survey-wave 2 (MHSS2), will be made publicly available at ICPSR alongside the 1996 Matlab Health and Socioeconomic Survey (MHSS1). MHSS1 data are currently available at https://www.icpsr.umich.edu/web/NACDA/studies/2705. Specific icddr,b census and DSS data must be requested from the organization. More information can be found at http://www.icddrb.org/component/content/article/10003-datapolicies/1893-data-policies. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the United States Air Force Academy, the Air Force, the Department of Defense, or the U.S. Government.

†Barham: University of Colorado Boulder, Economics Building Rm 212, 256 UCB, Boulder, CO 80309 (tania.barham@colorado.edu); Champion: United States Air Force Academy, 2354 Fairchild Drive, Suite 6K-110, Colorado Springs, CO 80840 (brachel.champion@afacademy.af.edu); Kagy: University of Wisconsin - Madison, 1300 Linden Dr, Madison, WI 53706 (gkagy@wisc.edu); Hamadani: icddr,b (jena@icddrb.org).

1 Introduction

A substantial body of evidence shows clear connections between the human capital and economic opportunities of parents and their children.[1] Economic models of human capital formation and transmission suggest that these correlations could arise from a complex interplay of genetic factors, environmental influences, and investments.[2] Moreover, they predict that positive investments during sensitive periods of development, such as early childhood, can have lasting effects on human capital formation, leading to better economic opportunities in adulthood and for future generations. While many existing public programs improve the early childhood environment and global investment in maternal and child health has increased substantially in recent decades, our understanding of the long-term and intergenerational impacts of these investments and their potential to reduce persistent parent-child human capital inequalities remains limited. This understanding is crucial for designing effective policies that break the cycle of disadvantage across generations.

In this paper, we bring new panel data that spans three generations to provide evidence that public health interventions have important long-term and intergenerational effects on multiple measures of human capital. Furthermore, we show that these positive effects are strongest for those who had lower measures of maternal human capital prior to the intervention, thus reducing inequality. To identify the effects, we take advantage of the quasi-randomly placed Maternal and Child Health and Family Planning (MCH-FP) Program. The MCH-FP was rolled out over time in the Matlab district of Bangladesh in the late 1970s and 80s in a treatment area, leaving a comparison area for evaluation purposes. MCH-FP revolutionized the field by using a home-based delivery model, integrating family planning with mother-and-child health services. The services provided are typical of public health interventions world-wide, including the provision of modern contraception, maternal tentanus toxoid vaccination, and vaccination against early childhood diseases such as measles (Bhatia et al., 1980). Together these interventions represent a substantive improvement to the early childhood health environment.

We investigate the impact of the MCH-FP program on the human capital of two generations. Long-term effects are evaluated through the first-generation cohort which is comprised of individuals born between 1982 and 1988, a period when both family planning and intensive early child health interventions were accessible in the treatment area. Their longer-term human capital outcomes are assessed when they are approximately 24 to 30 years old. Intergenerational effects are measured by the second-generation cohort which includes the first-born children of females in the first-generation cohort. Their human capital outcomes are measured at 0-14 years old. We further investigate the distributional effects by the pre-program health endowment using a third generation, generation zero, who are the mothers of the first generation. We present a conceptual framework drawing from the dynamic human capital production function model of Heckman (2007) and Attanasio (2020) to illustrate how these three generations may be interconnected and how the MCH-FP program might influence human capital within and between

1. For examples on income and education see Black and Devereux (2011), Hertz et al. (2008), Richter and Robling (2013), for health Bhalotra and Rawlings (2013) and Venkataramani (2011), and for cognition Grönqvist, Öckert and Vlachos (2017).

2. For theoretical models of early-life skill formation in economics see Cunha and Heckman (2007); Heckman (2007*a*); Attanasio (2015), and for biological models of developmental origins of health and disease see (Gluckman, Buklijas and Hanson, 2016).

generations.

The scope and rigor of the study is enhanced by the rich data availability that links individuals over time across data sources connecting multiple generations. The outcomes are drawn from the 2012 Matlab Health and Socioeconomic Survey 2 (MHSS2). This survey has low levels of attrition—less than 10 percent in each generation—despite data being collected 35 years after program start in a setting with high levels of migration. Notably, more than 60 percent of first generation men had migrated out of the study area, 25 percent to international destinations.[3] To examine the robustness of the research design, these data are linked to several data sets on the study area including MHSS1 from 1996, demographic surveillance (including births, deaths, and migration) that span the pre-program period to MHSS2, and pre-program census data.

A key feature of the MCH-FP program is that treatment and comparison areas were built into the program and designed to be economically and socially similar. Villages were not randomized into treatment, rather villages were placed into six treatment and comparison blocks to facility program delivery and limit inherent spillovers. We take advantage of the well-designed treatment and comparison areas to estimate intent-to-treat (ITT) effects on multiple human capital measures including height, short stature/stunting, cognition, and grades attained. For the first generation, we estimate double-difference models with birth-year fixed effects and pre-program controls and weight for attrition from birth to MHSS2.[4] We use individuals born between 1947-1969 as a baseline cohort because they were adults when the MCH-FP program began so the human capital measures examined in this paper were unlikely to be affected by the interventions. For the second generation, we estimate single-difference models using mother's treatment status as our primary specification because it is not possible to identify a baseline cohort that is similarly-aged and unaffected by the program. We present double-difference models as a robustness check using a less affected comparison group and results are similar. To support this empirical strategy, we demonstrate that the comparison area provides a good counterfactual: there is pre-program balance in human capital, fertility, mortality, employment, migration trends, and most individual and household characteristics. We also find no evidence of spatially correlated errors between treatment and comparison areas, or have knowledge of a disease outbreak that affected only the treatment or comparison area from the decades of demographic surveillance data on mortality and disease (Fauveau, 1994). Results are also robust to multiple hypothesis testing and randomization inference that takes into account the placement of villages into a contiguous treatment area.

Findings on the first generation reveal human capital gains documented in adolescence (Barham, 2012; Joshi and Schultz, 2013) persist into adulthood for height and grades attained but not for cognition. Both males and females are about one centimeter taller and short stature is reduced by almost 50 percent. In addition, ITT effects for grades attained are almost one year for men. There are no program effects on education for first generation females which is consistent with the availability of a successful nationwide

3. To reduce attrition migrants were tracked throughout Bangladesh, international migrants interviewed when they returned to Bangladesh for vacation, and a phone survey was implemented to contact the remaining international migrants.

4. There is no pre-analysis plan for this research as the data collection for this paper was conceived prior to their common use in economics. Matlab Health and Socio-Economic Survey 2 (MHSS2), was designed and collected by some of the authors in collaboration with a multi-disciplinary group, to study the long-term effects of the MCH-FP program. Descriptions of the planned research design and analysis are in the grant applications for the data collection.

girl's secondary school scholarship program operating in both the treatment and comparison areas at the time (Shamsuddin, 2015).

To understand the distributional effects of the program on the first generation, we decompose the treatment effect by terciles of an individual's pre-program health endowment. We use mother's height as a proxy for the first generations' health endowment since it is not affected by the program and height reflects the early childhood health environment. Results show program effects for the first generation are largest for those born to the shortest mothers, where potential gains from health improvements were the largest. This indicates that the MCH-FP contributed to breaking the parent-child human capital correlation and reduced human capital inequality for the first generation.

The second-generation findings are concentrated solely among females. ITT effects indicate that, similar to their mothers, female offspring are on average taller (1.6 cm or 0.33 SD). These height gains are observed at the lower end of the distribution, resulting in a sizable 50 percent reduction in stunting. These improvements essentially eliminate stunting within this cohort, underscoring that those who were potentially the most disadvantaged benefited from the program. Furthermore, cognition effects, which faded for the first generation, resurfaced among the 7-14 year olds in the second generation, showing a substantial effect of 0.26 SD.

Finally, we explore if second generation effects could be a result of differential selection into or investment in the cohort. We rule out that second-generation findings are a result of fertility or mortality selection, or differences in mother's human capital characteristics or empowerment. We further show that the second generation effects are not driven by differential measures of typical pre- and post-natal health investments such as number of prenatal care visits, skill of birth attendant and vaccination.

This paper makes several important advances to the literature on the long-term and intergenerational effects of early childhood health intervention. In general, this paper is unique in that it offers causal identification based on purposeful quasi-random program placement along with extensive data that connects three generations and provides baseline information. The low attrition in the outcome data mitigates concerns related to selection that are pervasive in the long-term and intergenerational literature. Moreover, we investigate the effects of widely-implemented public health interventions on multiple measures of human capital, which are typical inputs into the labor production function. This approach provides robust evidence concerning the enduring effects of a positive early childhood intervention on two generations and allows us to make a number of contributions which are important for both policy and research.

In addition to our general contribution, we contribute specifically to the limited but growing literature on long-term effects of positive early child health interventions by shedding light on how these programs disrupt the transfer of human capital from one generation to the next by identifying who benefits from the program based on their initial health endowment.[5] These findings reveal that individuals in the first

5. Evidence on the long-term effects of early child health improvements on adult human capital outcomes in the U.S. focuses on the expansions of large federally funded programs, such as Medicaid (Miller and Wherry, 2019) and food stamps (Hoynes, Schanzenbach and Almond, 2016), and show improved educational attainment and adult health including reductions in chronic disease and the incidence of metabolic syndrome. However, effects on cognition are sparse and evidence from Head Start showed early program impacts on cognition fade by adolescence (Deming, 2009). Bailey (2013) provides suggestive evidence the family planning policies can increase college completion, labor force participation and family incomes decades later. In

generation with the lowest levels of maternal health endowment benefit the most from an improved early childhood environment, thereby reducing human capital inequality in adulthood. This research offers valuable insights for policymakers aiming to break the cycle of disadvantage in human capital across generations.

We further contribute to to the emerging literature on intergenerational effects of positive early childhood investments by providing novel evidence on multiple measures of human capital and by deepening our understanding of the selection process. Notably, there are few prior studies due to a lack of well-designed programs introduced decades earlier with data that includes treatment status and links multiple generations. Previous work finds reductions in low birth weight from the roll out of Medicaid in the US (East et al., 2023), and improvements in height and weight-for-age for males from a nutrition intervention in Guatemala (Behrman et al., 2009), though results are hard to interpret due to attrition.[6] Related research on the intergenerational effects of early childhood education shows improvements in educational attainment from Head Start (Barr and Gibbs, 2017) and the Perry Preschool Project (Garcia, Heckman and Ronda, 2021).

In contrast to these studies, this study offers several advantages. First, we examine multiple dimensions of human capital beyond birth, disaggregated by gender. Understanding effects beyond birth and across multiple dimensions is critical because some behavioral mechanisms linking generations manifest after birth, such as post-natal parental investments, and may vary depending on the specific measure and gender of the child. Moreover, the medical literature highlights the intricacy and gender-specific nature of some biological mechanisms influencing the subsequent generation (Drake and Liu, 2010; Sandovici et al., 2022). Second, our approach allows us to rule out common sources of bias or selection into the second generation including fertility and mortality selection, and attrition.

Finally, by examining the program's effects over a longer time-frame than many previous studies—specifically 35 years after the program's inception—we can illustrate how program effects on human capital outcomes may differ throughout the lifespan and across generations. Indeed, our study indicates that for females the effects on cognition identified in adolescence waned among adults in the first generation but resurfaced in the second generation, whereas the effects on height remained consistent in both generations.

Overall, the findings demonstrate that investments in the early health environment through common public health interventions can generate substantial and enduring effects on human capital. These benefits are particularly pronounced for individuals with shorter mothers, as they disrupt the persistent pattern of height correlation between parents and children in the first generation and contribute to a reduction in short stature across both generations, mitigating human capital disparities across generations.

developing countries, Maluccio et al. (2009) and Walker et al. (2021) demonstrate an early childhood nutrition intervention in rural Guatemala lead to higher educational attainment, reading comprehension and cognition in adulthood, though attrition was high (40 percent).

6. The program provided a protein enriched beverage in two randomly chosen villages out of four with the other villages receiving a sugary beverage. Results are based on interviews with approximately 60 percent of the original sample.

2 MCH-FP Program

2.1 MCH-FP Intervention

The MCH-FP program was initiated in October 1977 in the rural Matab district of Bangladesh and administered by icddr,b, (formally known as International Center for Diarrhoeal Disease Research, Bangladesh). The program included integrated family planning and maternal and child health services. A key feature of the program was that interventions were free and administered in the beneficiary's home during monthly visits made by local female health workers hired and trained by the program (Bhatia et al., 1980). The study area covered about 200,000 people in 149 villages. To facilitate evaluation these villages were split into a treatment and comparison area. Individuals living in the treatment villages were eligible for health and family planning interventions provided by the program while those living in the comparison villages were not, but they had access to regular government services (Fauveau, 1994).

Services rolled out over two main phases: October 1977 to February 1982 and March 1982 to December 1988. During the first phase, program interventions focused on family planning and maternal health through the provision of modern contraception, tetanus toxoid vaccinations for pregnant women (starting June 1978), and iron and folic acid tablets for women in the last trimester of pregnancy (Bhatia et al., 1980). Tetanus toxoid was expanded to all women of reproductive age in 1982. During these visits, the female health workers also provided counseling on contraceptives, nutrition, hygiene, and breastfeeding, and provided instructions on oral rehydration solution preparation. Follow-up and referral systems ensured management of side effects and aided in the continued use of contraceptives (Phillips et al., 1984).

During the second phase starting in March 1982, child health interventions were intensified for children under the age of five. The measles vaccine was introduced in half the treatment area and expanded to the other half in November 1985 (Koenig et al., 1990). Additional child health interventions were phased in between 1986 and 1988. In January 1986, DPT (diphtheria, pertussis, and tetanus), polio, and tuberculosis immunizations were added, and, later in that year, vitamin A supplementation. Curative care, such as nutrition rehabilitation for the nutritionally at risk, was introduced in the late 1980s.

During the program roll-out, the comparison area had access to then-standard government health and family planning services. These services included access to modern family planning in the clinics (rather than in the home) but childhood vaccines were not yet available in the comparison area, or most of Bangladesh. Differences between the treatment and comparison areas narrowed after 1988 as lessons of the Matlab success were incorporated into the national plan (Cleland et al., 1994) providing an approximately 10-year experimental period. Starting in 1989, similar vaccines became more readily available nationwide, and the number of government community health workers delivering in-home services increased throughout the country over time. For example, in the comparison area, the worker-client ratio increased from 1 per 8000 in 1987-1988 to 1 per 5,000 in 1989-1990 in (Cleland et al., 1994), though this ratio was still lower in the treatment area at 1 per 1,300 in 1990.

2.2 MCH-FP Rollout

Program implementation followed the planned timeline and uptake was rapid. For example, Figure 2 indicates that the contraceptive prevalence rate (CPR) for married females 15-49 was similarly low prior to the program (<6%) in both areas. There was a large increase in the CPR to 30% in the treatment area during the first year of the project and it increased steadily, reaching almost 50% by 1988. Due to availability of contraceptives from government services, the CPR did increase in the comparison area over time, though not as quickly, and rates still remained below 20% in 1988.

The measles vaccination rate also rose rapidly to 60% in 1982 after it was introduced in half of the treatment area and in 1985 when it was introduced in the other half (Figure 2). By 1988, coverage rates for children aged 12–23 months living in the treatment area were 93 percent for the BCG vaccine against tuberculosis, 83 percent for all three doses of DPT and polio, 88 percent for measles, and 77 percent for all three major immunizations ((Icddr, 2007). Data on vaccination rates in the comparison area do not exist, but are believed to be near zero. Government services did not regularly provide measles vaccination for children until around 1989, so the comparison area was viewed as a largely unvaccinated population (Koenig, Fauveau and Wojtyniak, 1991). This is in line with national data that indicate the measles vaccination rate for children under the age of five in Bangladesh was less than 2% in 1986 (Khan and Yoder, 1998). While vaccinations became available in the comparison area after 1988, vaccination rates were still relatively low (below 40%) in the comparison area in 1990 (Fauveau, 1994).

2.3 MCH-FP Placement

To ensure a rigorous evaluation of the program, treatment and comparison areas were built into the design of the MCH-FP program. Villages were not randomly assigned a treatment status but instead were placed into blocks to facilitate access and delivery of program health services and to limit spillovers from vaccination and family planning. Villages were assigned to one of either four contiguous treatment blocks or two comparison blocks that flanked the treatment area (Figure 1) with the population being split fairly evenly into 70 treatment and 79 comparison villages. The treatment blocks were each organized around a MCH-FP clinic created by the program and staffed by a paramedical personal and a male health assistant (Fauveau, 1994). Each comparison area block also has a government run health clinic. Local female health workers who were responsible for in-home visits were linked to an MCH-FP clinic. These six blocks were created specifically for the program and did not represent any prior level of government or health service delivery area. The block design was important to mitigate potential spillovers into the comparison area of information about the family planning interventions (Huber and Khan, 1979) or from the positive externalities generated by vaccination. The comparison and treatment areas were also viewed to be socially and economically similar and geographically insulated from outside influences at the time (Phillips et al., 1982). Section 5.2 confirms that the treatment and comparison areas were indeed similar.

2.4 Short-term Mechanisms of the MCH-FP Program on Human Capital

Children from low-income families often face a myriad of interrelated obstacles such as infectious diseases, malnutrition, and high levels of fertility in reaching their full human capital potential. At the time of the MCH-FP rollout, Matlab was a poor rural area characterized by high rates of disease and fertility as reflected by an under-five mortality of 22 per thousand and a total fertility rate of 6 (Bhuiya et al., 1987). These issues were compounded by poor nutritional status and low levels of education for both males and females. In 1980, 56 percent of under-five children in the study area were less than 70 percent of weight-for-age and 45 percent of males and 74 percent of females over the age of 15 had no formal schooling (Bhuiya et al., 1987). By 2000 when second generation children were beginning to be born, the environment in Matlab had improved substantially and rates of under-five mortality, underweight and stunting had declined to 15.7, 52, and 35 percent respectively, and total fertility halved to 3.2 (Alam, Zahirul Haq and Kim Streatfield, 2010; Das et al., 2015). These rates of malnutrition are comparable to lower-income countries today, but Matlab's mortality and fertility rates are low and closer rates in middle-income countries.

The MCH-FP program provided a package of interventions—vaccinations, family planning, and maternal health—that are known to improve the early childhood health environment and lead to better nutrition, health, and cognitive development, all of which can later enhance educational outcomes. In fact the dissemination of early childhood vaccines was in general viewed as a turning point in public health in the twentieth century because their widespread accessibility lead to rapid declines in rates of morbidity and mortality from contagious diseases.[7] Some of these now vaccine-preventable diseases are particularly damaging because of the longevity of the impact of the virus. For example, the measles virus can cause immunosuppression for up to five years, and effectively resets previously acquired immunity (Gadroen et al., 2018). This "immune amnesia," creates a highly-compromised immune system during childhood and is associated with increased non-measles related mortality for a period of two years after a measles infection, and increased non-measles related morbidity for up to 5 years afterwards (Gadroen et al., 2018; Mina et al., 2015). Not only are there the direct health effects from the vaccine-preventable diseases, but morbidity from the dieseases also causes general malaise, and apathetic children typically receive less stimulation from adults and less learning opportunities, further hindering human capital development (Walker et al., 2007).

While these diseases directly effect the health and cognitive development of a young child (e.g. pertussis can lead to encephalitis), they also affect children's nutritional status from the disease itself and from secondary complications such as pneumonia and diarrhea (Reddy, 1987). Both non-randomized and randomized studies document that undernutrition—especially before the age of three—affects the growth and cognition of young children (Grantham-McGregor, Fernald and Sethuraman, 1999*a*,*b*; Walker et al., 2007). While children's physical and developmental growth can catch up once the illness has passed, children may experience a number of episodes of illness in combination or close succession, reducing the time for catch-up growth. For girls, catch-up growth that takes the form of accelerated growth may

7. For example, prior to the measles vaccine it is estimated that by age sixteen 95% of a population would have contracted measles at some point in their lives (Miller, 1964; Perry and Halsey, 2004)

also trigger early puberty which limits final height (Proos, 1993). The negative impacts of disease on human capital development may be stronger in areas with higher levels of undernutrition, such as lower income countries, because malnutrition weakens the immune system, leaving children more susceptible to disease and lengthening recovery times.

Finally, the non-child health components of the MCH-FP program could also affect human capital. Longer birth spacing resulting from family planning interventions and maternal health inputs (e.g., iron and folate supplementation, nutritional counseling, referral for pregnancy complications) may affect the human capital of a child through the improved nutrition and health of the mother while the child is in utero (Walker et al., 2007; Almond and Mazumder, 2011). Furthermore, reduction in family size from access to modern family planning can lead to a quantity-quality trade-off, with lower-fertility parents investing more in their children both in terms of time spent together and access to resources, such as better nutrition or more schooling.

2.5 Effect of the MCH-FP Program on Human Capital in Childhood

Previous literature documents that the MCH-FP program had important effects on the human capital attainment of the first generation in late childhood using survey data from 1996. Children born during the roll-out of the intensive child health intervention (1982-1988) had improved human capital as measured by height (0.22 SD), cognitive functioning (0.39 SD), and years of education (0.17 SD) at ages 8-14 (Barham, 2012). The increase in schooling for these children was concentrated in males (Joshi and Schultz, 2013). There was no improvement for children born prior to the availability of the intensive child health interventions (1977-1981) at ages 15-19 (Barham, 2012). Driessen et al. (2015) demonstrates that improvements in schooling are associated with measles vaccination take-up, consistent with Barham (2012).[8]

3 Conceptual Framework

We specify a conceptual model to illustrate the channels through which the MCH-FP program could have affected the process of human development in adulthood and across generations. We follow Heckman (2007*b*) and Attanasio, Meghir and Nix (2020) and model the dynamic evolution of human capital, but restrict the model to a three-stage framework, encompassing early childhood ($t = 1$), late childhood ($t = 2$), and adulthood ($t = A$). We include j types of human capital in the model to capture the multiple dimensions of human capital our analysis considers. The production function for adult human capital of type j, θ_{jA}, can be written as:

$$\theta_{jA} = f_j(\theta_{j1}, \theta_{-j1}, I_1, I_2, Z_1, \theta_{jp}, \theta_{-jp}) \tag{1}$$

where θ_{j1} are initial conditions which we simplify to be a child's endowment of type j of human capital, θ_{-j1} a child's endowment of all other measures of human capital except j, I_t represents either public

8. The program also reduced family size by almost one child (Phillips et al., 1984; Joshi and Schultz, 2013) though the reduction in family size depends on length of exposure of the mother to the program (Barham et al., 2021).

or private investments made during early childhood ($t = 1$) or late childhood ($t = 2$), Z_1 is a vector of background characteristics (such as gender and religion), θ_{jp} is parental human capital of type j and θ_{-jp} is parental endowment of all other measures of human capital except j. Given conditions outlined in Heckman (2007*b*), this framework allows for a dynamic human capital production function that embodies qualities of self-reinforcement, dynamic complementarity, and sensitive periods of development.

Self-reinforcement means that, for a given level of investment, higher levels of human capital in one period create higher levels of human capital in the next period, within and across human capital types, i.e., $\partial f_j(\cdot)/\partial \theta_{j1} > 0$. Dynamic complementarity arises when investments are more productive because the previous period's stock of human capital is higher i.e., $\partial^2 f_j(\cdot)/\partial \theta_{j1} \partial I_2 > 0$. Finally, a period is defined as sensitive if a given level of investment has a higher return in this period than any other period. The relevance of these three concepts in any period could vary by background characteristics such as gender and human capital type.

For the first generation, the MCH-FP program represents an increase in public investment, I_1, in early childhood, the first period of the model. Early childhood is viewed as a sensitive period for human capital measures such as height and cognition. For the same sample considered in this paper, Barham (2012) shows that the first period investment was successful at increasing human capital in the second period, late childhood, for three measures of human capital: height, cognition, and grades attained (see Section 2.5). However, whether these improvements in human capital caused by first period investments persist into adulthood may depend on dynamic complementarity and may require further investment in the second period. The persistence of impacts in adulthood may vary by the type of human capital, and could depend on whether period 1 is the only sensitive period for that type of human capital. Investments made in period 1 may also affect adult human capital through epigenetics as the biomedical literature shows one's first period environment can alter gene expression in adulthood.[9]

For the second generation, the framework provides several pathways for intergenerational transmission of MCH-FP benefits. First, parental human capital (θ_{jp} and θ_{-jp}) is directly altered by the MCH-FP program, and this may lead to biological pathways through which higher human capital of the first generation may impact the second generation. The biological pathways are complex and understanding these mechanisms is an active area of research in both human and animal studies. These pathways include anatomical, physiological, or epigentic pathways(Drake and Liu (2010), Fitzgerald, Hor and Drake (2020), Gluckman, Buklijas and Hanson (2016)). For instance, physical growth of first generation mothers in period 1 could allow them to physically bare larger children, or both parents from the first generation may pass on height and health-determining genes to their children (the second generation) including epigenetic marks resulting from the intervention.

Second, behavioral mechanisms can arise when parental human capital is correlated with the level and productivity of investments made in early and late childhood (I_1 and I_2). Higher human capital parents may be healthier and more able to spend time actively engaging with their children, thus providing a more stimulating environment. Tastes and preferences for health and health-promoting behaviors may also be altered by higher parental human capital. In addition, higher parental human capital could lead to

9. See Gluckman, Hanson and Beedle (2007) and Gluckman, Buklijas and Hanson (2016) for a survey of this literature.

increased income or financial support for investments in children.

It is important to note that both biological and behavioral mechanisms can vary by the sex of the child, making it important to dissagregate effects by sex. Again, while biological mechanisms are not well understood, there are documented difference by sex in biological mechanisms including the susceptibility to disease and intergenerational inheritance of traits, making it important to examine effects by sex (Sandovici et al. (2022) Drake and Liu (2010)). On the behavioral side, parental investments may vary by the sex of the child due to persistent environmental factors such as cultural practices, or as a result of program impacts on the first generation outcomes such as their human capital or agency.

This framework also conveys a conceptualization of the interrelation of human capital across three generations. For the first generation, their parents' human capital (θ_{jp} and θ_{-jp}), represents their pre-program health endowment, and is a direct input into the their human capital production function. Effects may be heterogeneous depending on the interaction between the parental health endowment and increased public investment (the MCH-FP program) or private investment (I_1). For the second generation, equation 1 predicts that individuals are influenced by their pre-program health endowment, which is the human capital of generation zero (their grandparents). The influence of the pre-program health endowment, however, is mediated through its effect on the human capital of the first generation (their parents), affecting both the health endowment and parental investments of the second generation.

4 Data and Attrition

4.1 Data Sources

We draw on rich panel data from the Matlab district. Unique identification numbers allow for linking individuals throughout time from four main data sources: the 2012–2015 Matlab Health and Socioeconomic Survey wave 2 (MHSS2), the 1996 Matlab Health and Socioeconomic Survey wave 1 (MHSS1) (Rahman et al., 1999), periodic censuses of the study area conducted by icddr,b in 1974 and 1982 (icddrb, 1974, 1982), and 1974–2014 Matlab demographic surveillance site (DSS) data on vital events (e.g., births, marriages, deaths, in- and out-migrations) collected by icddr,b.[10] MHSS1 and MHSS2 are random samples of the study area, while the periodic censuses and DSS data cover the entire study area. The census and DSS data are known for their high quality (they were collected in the home bi-weekly or monthly) and allow for determination of exact birth dates, treatment status, migration status, and testing of pre-program balance. These data sets are further linked (usually at the village level) to data on potentially confounding programs such as access to microcredit, primary schools, health facilities, flood mitigation, and arsenic exposure. More details on data construction are in Appendix B and potential confounders in Appendix C.

Outcomes are drawn from MHSS2 which is a large socioeconomic survey designed as a panel follow-up of all individuals in the MHSS1 primary sample and their descendants. The MHSS1 primary sample is representative of the study area's 1996 population, but does not include individuals who migrated between

10. MHSS2 data linked with the baseline data will be publicly available in the future. It was collected by the authors together with a team of researchers from the University of Colorado Boulder, Brown University, and icddr,b. Specific icddr,b census and DSS data must be requested from the organization. More information can be found at http://www.icddrb.org/component/content/article/10003-datapolicies/1893-data-policies

the start of MCH-FP and MHSS1. To address this bias, MHSS2 also includes individuals born to an MHSS1 household member between 1972 and 1989 who had migrated out of Matlab between 1977 and 1996, which we refer to as pre-1996 migrants. Finally, MHSS2 sample does not include anyone from households where the whole household and all the descendants migrated between 1977 and 1995, because there was no one left to sample in MHSS1. The annual whole household migration rate from the entire study area over this time period was low, 0.66 percent, and most exiting households were Hindu.

MHSS2 was collected between 2012 and 2015 throughout Bangladesh and included interviews of international and difficult-to-track migrants when they returned to the study area to visit family, particularly during Eid celebrations. A phone survey was also administered to international migrants who did not return to Bangladesh during data collection. Phone survey respondents are all men and represent fifteen percent of all men in the sample. In-person survey respondents were interviewed and tested in their homes so the survey does not suffer from attrition that occurs when test measures are collected in institutional settings such as schools. Individual test and health measurements were collected by enumerators who were rigorously trained by psychologists and public health professionals working for icddr,b. While the phone survey is shorter than the main survey, it does include self-reported measures of height and education. However, only one cognitive test was included in the phone survey (digit span forward), so a cognitive index is not available for the phone survey respondents (all men), resulting in smaller sample sizes for cognition for men.

Attrition rates for MHSS2 are less than 10 percent of the target sample which is extremely low for a 35-year follow-up survey. Attrition, including death, for the first generation sample is 10 percent for men and 7 percent for women (Table B1). Second generation sample attrition is less than 7 percent, with attrition from mortality accounting for about 4 percent (Table B2). Attrition is balanced between the treatment and comparison groups for all age and sex cohorts (Table B3). See Appendix B.2 for details.

4.2 Outcomes

This paper focuses on four main measures of human capital: height, short stature[11] or stunting, cognition, and grades attained. These outcomes are referred to as stock measures of human capital because they tend to be stable by adulthood. Height and cognition are also used as indicators of early childhood health and nutrition since early childhood is a sensitive period of physical growth and brain development. [12]

To help with interpretation and account for differences in outcomes with age and gender, outcomes are standardized into z-scores by age and sex. We use the WHO international standards for height and stunting of the second generation, otherwise, z-scores are created by internally standardizing by age and sex using the comparison group mean with six-month age bins for those age 6 and younger, and one year age bins for those age 6 and older. Cognition measures are all internally standardized into z-scores because norms for the cognitive tests do not exist for Bangladesh.

MHSS2 includes a battery of cognitive tests covering multiple domains so we create a cognition index

11. Short stature is used for the first generation and is define as height less than 155cm for males or 145cm for females.
12. We do not present program effects on metabolic syndrome, a cluster of conditions that indicate an increased risk of heart disease, stroke and Type 2 diabetes, because the prevalence of these risk factors, while growing, is still low among the less than thirty population in Bangladesh.

by averaging the standardized test scores to reduce concerns of multiple hypothesis testing. For the first generation, the cognition index includes the Mini Mental State Exam (MMSE), Digit Span Forwards and Backwards, and Raven's Colored Progressive Matrices (Ravens). Digit Span Forwards and Backwards are tests of working memory which is thought to increase with intelligence; Ravens is a non-verbal and simple measure of general intelligence and perhaps the most common and popular test for people above the age of five (Kaplan and Saccuzzo, 2009); the MMSE was developed as a global assessment of cognitive status and is also a frequently-used, brief cognitive screening test (Ismail, Rajji and Shulman, 2010). Digit Span Forwards is the only cognitive test available in the phone survey so we do not create a cognition index for phone survey respondents, resulting in a lower sample size for men for this outcome. More details on how cognition outcomes are defined are in Appendix B.3. For the second generation, cognitive tests for children under age seven are drawn from the Denver Developmental Screening Test (Frankenburg et al., 1975). We use the subcomponents on language, fine motor, and gross motor skills to create a Child Development Index. For respondents ages seven and older, the cognition index is comprised of the MMSE, Digit Span Forwards and Backwards, Ravens, a test of memory drawn from a subtest from the Woodcock Johnson IV Test of Cognitive Abilities (Memory for Names), and a three minute timed visual matching test to measure perceptual speed.

Finally, we examine effects on grades attained for the first generation, but not the second generation. About half the second generation sample are too young to attend school, and most school age respondents had not yet reached the age of school drop out. Grades attained is taken from the household survey rather than the testing module where the other outcomes were collected, and the sample size is slightly larger as a result.[13]

4.3 Intent-to-Treat Indicator and Linking to Baseline Variables

Access to the MCH-FP program was based on the village of residence of the individual during the program period. Because a person's residence after the program start is potentially endogenous, we use DSS and census data to create an intent-to-treat indicator based on the village of residence for an individual's first household head prior to 1977.[14] Individuals in MHSS1/2 are linked with the DSS and census data by a unique ID available in each of the datasets using the following sequence of linkages. First, we link our respondents to the 1974 census through the household head of their first residence in the DSS area. If their household head was not present in the 1974 census, we identify that person's first household head in the DSS area and link that new person to the 1974 census. Finally, remaining unlinked individuals are assigned a treatment status using the location of their household head in the DSS area after the 1974 census, but before the inception of MCH-FP in 1977.[15] The ITT variable, Treat, takes a value of one if the

13. Grades attained are taken from individual reports in the household survey. Results are similar when missing values are filled in with reports from the household roster.

14. The treatment indicator is nearly identical if individuals were linked to 1974 through their fathers and grandfathers. Less than 0.5% of the sample would have been assigned a different treatment status. We use household head because this sequence of linkages results in more direct links to the 1974 census, and therefore fewer missing treatment status and baseline characteristics.

15. We link over 96% of individuals in our sample to the 1974 census through their first household head. An additional 3 percent link to the 1974 census through that person's first household head. The remaining less than 1 percent link through their household head's location in the DSS after the 1974 census, but before program inception in October 1977.

1974 census-linked household head was living in a village in the treatment area in 1974 or migrated into a village in the treatment area between 1974 and 1977 and zero otherwise.

Baseline characteristics from the 1974 census are linked to individuals in the same manner used to construct treatment status. For the few individuals that could not be linked to the 1974 census, missing baseline characteristics are assigned means based on treatment status, sex, and cohort.[16] Finally, the village from the 1974 census link is used to cluster standard errors in our analysis.

5 Empirical Strategy

5.1 Analysis Samples

This paper analyzes the impact of the MCH-FP program on the early childhood health environment on the human capital of two samples—the first and second generation—to determine the longer-term and intergenerational effects, respectively. The first generation sample includes those born during the experimental phase of MCH-FP when both family planning and child health interventions were provided in the treatment area. This cohort was born between 1982–1988 (ages 24–30 in MHSS2) and hence received early childhood investments during a sensitive period of development, namely $t = 1$ in the model.

The second generation sample is comprised of firstborn children of females from the 1982–1988 first generation cohort. The second generation sample were born between 2002–2015 and are aged 0-14 in MHSS2. We identify children based on pregnancy histories collected in MHSS2 which includes a listing of all live births, stillborn children, or lost pregnancies.[17] We follow the children of the females in the first generation cohort because they are the cohort that experienced human capital gains from the MCH-FP program as children (Barham, 2012), and analysis from this paper finds gains persisted into adulthood. These females are also typically married to older men (mean age 34) who were born prior to the roll out of the intensive child health interventions in 1982 (69% of husbands), which helps make the interpretation of the treatment effect clearer. The MCH-FP program did not affect the human capital of people born prior to 1981 on average as children (Barham, 2012) or adults (Appendix Table A2), so the age difference between spouses helps isolate from which spouse any intergenerational effects on human capital from the program are likely being passed. We focus specifically on firstborns because first generation females are early in their fertility: 83 percent of the sample have a firstborn child and 47 percent have a second-born child by the time of the MHSS2 survey. Who has a second-born child is likely to be selected so we do not include this group of children.

While we would like to examine effects on children born to fathers in the 1982–1988 cohort, only 34 percent of men became fathers, as men have children later than females in this context. Furthermore, the majority of these men married females born after 1988, a time when the interventions became increasingly available in the comparison area, making it difficult to isolate the treatment effects because the spouses of these men, both in the treatment and comparison areas, are likely treated. While it is possible that we

16. Only 12 male and 13 female respondents have missing baseline data.

17. While the pregnancy history provides the main way to identify children, we check for any children missing from the listing, perhaps because they died when young, using the DSS data and the MHSS2 siblings module.

miss important intergenerational effects by not examining effects on children of first generation fathers, previous research shows intergenerational correlations are stronger between mother and child than father and child (Thompson, 2014; Bhalotra and Rawlings, 2013).

5.2 Baseline Balance

This paper takes advantage of the treatment and comparison areas built into the design of the MCH-FP program to estimate ITT effects. Villages were not randomized into treatment and comparison areas, but rather placed into four treatment and two comparison blocks that were designed to be socially and economically similar to aid program implementation and limit spillover effects (see Section 2.3 for more details). In this section, we confirm the areas were similar at baseline for our sample and that the pre-program human capital endowment was also balanced. Past research has demonstrated that the treatment and comparison areas were balanced not only in mortality and fertility (outcomes directly related to the program interventions) but also in migration stocks and flows, employment sector of the household head, and many household and household head characteristics.[18] To test if the baseline characteristics are balanced for the sample used in this paper, we link the 1974 census data to MHSS2 as discussed in Section 4.3. Table 1 presents differences in means between the treatment and comparison areas and normalized differences in means (difference in the means divided by the standard deviation of the mean for the comparison group) that are not influenced by sample size (Imbens and Wooldridge, 2009).[19]

Ideally, we would like to show that human capital was similar between the treatment and comparison areas prior to the intervention. Since first generation children were born during the program, we first examine the human capital of the parents of the first generation sample – the pre-program endowment. Given a lack of pre-program data on human capital for this sample, we use MHSS1 to examine the balance in human capital for this group. While MHSS1 was collected in 1996 after the program was implemented, the parent's of the first generation sample had mostly completed schooling and growing prior to the start of the program in 1977, so it is unlikely that the program would affect height, education, or cognition. Table 1 Panel A indicates that differences in mother's and father's height, years of education, and cognition are balanced between the treatment and comparison area with the differences being small and not statistically different from zero. To further test the baseline balance of human capital, we also show in Table 1 Panel B that these same measures are balanced for the pre-program cohort born between 1947-1969 that we use in the double-difference estimator for the first generation. Finally, we use 1974 pre-program census data to examine education of the household head and spouse (there are no other pre-program measures of human capital). Table 1 Panel C shows that education was low, but balanced, in 1974 with 60 percent of heads and 87 percent of spouses having less than 2 years of education.

Lastly, we find individual and pre-program household characteristics between the treatment and comparison areas are balanced for the 1982–1988 first generation cohort pooled by sex in Table 1 in panel C. All normalized differences are small (less than 0.12) suggesting the differences are not substantial and only 5 out of the 23 baseline characteristics are significantly different below the 5 percent level for the

18. Koenig et al. (1990); Barham (2012); Menken and Phillips (1990); Barham and Kuhn (2014); Barham, Kuhn and Turner (2019)
19. Normalized differences bigger than 0.25 standard deviations are generally thought to be substantial.

pooled sample.[20] Characteristics that are significantly different include religion, drinking from a tubewell, latrine ownership, and age of the household head and spouse.

The two largest normalized differences are religion and access to tubewell water. The treatment area had fewer Muslims than the comparison area (84 versus 95 percent of household heads), the remainder being Hindu. All specifications control for if an individual is Muslim and robustness analysis shows that the results remain similar if only Muslims are included in the analysis. In the treatment area, 30 percent of households use tubewells compared to only 16 percent of households in the comparison area. The difference in tubewell water access in 1974 is the result of a government program rather than differences in household income, propensity to drill a tubewell, a household's concern about child health, or potentially other unobservables that could be correlated with a person's anthropometrics or cognition. Regardless, the difference in access to tubewell water for drinking is concerning since tubewell water is often thought to be cleaner than other sources of water. However, there is widespread groundwater arsenic contamination in the tubewells in Bangladesh (Chowdhury et al., 2000) which is a serious health concern and has been shown to reduce IQ among school-aged Bangladeshi children (Wasserman et al., 2004). As a result, the direction of the bias is unclear. In robustness analyses, we control for the level of arsenic in the tubewells to account for this potential bias.

These findings together with previous results strongly suggest that the two areas had very similar observable characteristics. To account for any differences in baseline characteristics we include observables listed in Table 1 Panels C as controls in the regressions, as well as provide robustness analysis to further explore the imbalance between religion and tubewell water as explained above.

5.3 Empirical Specification

We exploit the quasi-random nature of the treatment and comparison areas to estimate ITT effects of the MCH-FP program on first and second generation human capital. For the first generation, we estimate double-difference models that use a cohort born between 1947-1969 to account for any pre-existing differences as their human capital was unlikely to be impacted by the program. The 1947-1969 cohort was too old when the MCH-FP interventions were implemented for the program to affect their height, education or cognition directly and they are not likely to be a sibling of those impacted directly, thus minimizing sibling competition effects.

For the first generation, the double-difference model for individual i, from household h located in village v in 1974, is specified as follows:

$$Y_{iv} = \beta_0 + \beta_1 \text{Treat}_v \times \mathbb{1}(\text{Born 1982–88})_i + \beta_2 \text{Treat}_v + \alpha_t + X_i \Gamma + \varepsilon_{iv} \quad (2)$$

where Y_{iv} is the outcome of interest measured in MHSS2, Treat_v is the binary treatment status variable, and $\mathbb{1}(\text{Born 1982–88})_i$ is a cohort indicator that is one if person i was born between 1982–1988 and zero if born between 1947-1969. Birthyear fixed effects, α_t, are included to control for differences in the outcome due to year of birth as well as other events that may be correlated with birth year. The vector X_i controls

20. Results are similar if examined by sex, though there are only three significant differences.

for individual (sex and religion) and baseline household and household head characteristics in Table 1 panels C and their interaction with the cohort dummy. We cluster standard errors at the pre-program village level and weight all models to account for attrition from birth to MHSS2 (see Section Appendix D for details on the weights). In addition, we present two forms of inference adjustments. First, we display the p-value from a randomization-based inference test to account for the placement of the treatment area across a contiguous group of villages.[21] Second, we report Anderson's sharpened q-values that control for the false discovery rate when testing multiple hypotheses (Anderson, 2008).

In this model, β_1 is the ITT double-difference estimate of the long-term effect of early childhood investments from MCH-FP for those born between 1982–1988 when they are adults aged 24-30 in MHSS2. It represents the combined effect of all program interventions. Identification of causal effects for the double-difference model assumes that the treatment and comparison areas would have had the same trends in outcomes in the absence of the program. This is not a testable assumption but seems reasonable given the balance between treatment and comparison areas prior to the intervention discussed in Section 5.2 and the similarity in outcomes for the pre-program cohort between the treatment and comparison group, demonstrated by the coefficient on Treat_v. Furthermore, we discuss potential threats to identification in Section 8 and perform robustness checks to show results are robust to these threats. For example, we demonstrate there are no pre-existing differences in human capital outcomes of a similarly-aged cohort using data from an earlier period, there is no evidence of spatially correlated errors, other programs and changes in the area are not driving results, and that results are robust randomization-based inference and the false discovery rate. Finally, we present double-difference models with village fixed-effects to control for potential non-time varying village unobservables and results are similar.

For the second generation, we use a single-difference model as the main specification, which assumes that the means between the treatment and comparison areas would have been the same in the absence of the program. Again, this seems like a reasonable assumption given the similarity of the treatment and comparison areas. We do not use a double-difference model because there is not a suitable group that is at a similar point of development and unaffected by the program. However, as robustness we present a double-difference model that uses a less affected cohort and results are similar (see Section 8 for details).

The single-difference model for individual i, born to mother m from village v is specified as follows:

$$Y_{iv} = \beta_0 + \beta_1 \text{Mother's Treat}_v + \alpha_{\text{a}} + X_i \Gamma + \varepsilon_{iv} \tag{3}$$

where Y_{iv} is the outcome of interest, Mother's Treat_v is the binary ITT treatment status variable of the mother (who is in the first generation). Six-month age fixed effects, α_{a}, are included to control for differences in the outcome due to age as well as other events that may be correlated with age. X_i is a vector of individual (sex and religion), and baseline household and household head characteristics of the mother used in Equation 2. The coefficient β_1 is the single-difference ITT estimator. Similarly to the first generation, we cluster standard errors at the mother's pre-program village level, weight all models

21. With any assignment of village-level treatment status, significant treatment effects could occur simply by chance. Following Athey and Imbens (2017), we simulate the distribution of treatment effects that would occur from randomly assigning a fixed number of villages to treatment while maintaining the contiguity of treatment and comparison areas.

to account for attrition from birth to MHSS2 using the mother's weight, and present p-values for tests of random inference and sharpened q-values.

6 First Generation Results

6.1 ITT Effects on Human Capital

We present ITT effects of the MCH-FP program on the human capital measures(height, the cognition index, and grades attained) in Table 2. Pooled results are in Panel A, and results disaggregated by sex in Panels B and C. Comparison group means of the 1982-88 cohort are presented at the bottom of each panel for variables that are not z-scores along with the randomization inference test p-value and sharpened q-value. Based on the comparison group means, short-stature is still an issue for approximately 8 percent of men and 11 percent of women and average grades attained, while growing, is still low compared to developed countries at an average of 7 years for both sexes.

In the double-difference model, the point estimate on the treatment variable, *Treat*, is the difference in means between the treatment and comparison areas for the pre-program 1947-1969 cohort, given the control variables. This estimate is close to zero and not significantly different from zero for any of the outcomes in Table 2, providing further evidence of similarity between treatment and comparison areas prior to MCH-FP.

Results pooled by sex demonstrate important ITT effects among the 1982-88 cohort for height and grades attained, but not for cognition.[22] Specifically, the program led to a 0.98 cm (1 percent) increase in height (significant at the 1 percent level), a 4 percentage point (44 percent) decrease in short stature (significant at the 10 percent level), and a 0.42 (6 percent) increase in grades attained (significant at the 5 percent level). Effects for cognition are close to zero (-0.01 SD) for the overall index and for each component of the index (Table A1). While growth in height is hard to reverse and educational attainment is necessarily non-decreasing, the positive effects of early childhood investments from MCH-FP on medium-term cognition in 1996 have faded out by 2012 (Barham, 2012). The lack of effects in early adulthood cognition could indicate the need for dynamic complementarity through investment in later childhood ($t = 2$), or the difficulty in measuring differences between treatment and comparison areas at an age when cognition is most developed. Earlier program effects may reappear at a later period in the life cycle when cognitive decline commences.

Disaggregating by sex shows effects are similar between the sexes for height and short-stature but not for education, and there is no impact for either sex on the cognition. For height, males experienced a 1.05 cm increase (significant at the 10 percent level) and females a 1 cm increase (significant at the 5 percent level). The standard errors for males may be larger than for females due to measurement error in self-reported height in the phone survey.[23] For education, effects are driven solely by males

22. For comparison, results of the impact of the MCH-FP program on the cohort that were born when only family planning services were provided (born 1977-1981), who are age 31-34 during the MHSS2 survey are presented in Table A2. This cohort did not experience any impacts as teens and there are still no impacts on this cohort.

23. About fifteen percent of the male height observations were collected in the phone survey. Results are similar including a control for self-reported height and restricting the sample to only those whose height is measured Table A3.

who experienced 0.96 (14 percent) additional years for grades attained (significant at the 5 percent level). The lack of effect on education for females is consistent with the presence of a large nationwide stipend program to improve girls' education operating at the time in both the treatment and comparison area, that is estiamte to have increased girls' educational attainment by 1.2 years (Shamsuddin, 2015) (see Appendix C for more details).[24]

6.2 Distributional Effects by the Pre-Program Health Endowment

To explore distributional effects of the program on human capital, we examine the interaction between the pre-program maternal human capital endowment (θ_{jp}) and the improved childhood health environment resulting from the MCH-FP program. As noted in the conceptual framework, the MCH-FP program increased public investment, I_1, for the first generation. This improvement may impact adult human capital formation differently depending on the level of human capital of an individual's parents. We proxy the pre-program matrilineal health endowment with terciles of height of the first generation's mother, generation zero, since it should not have been affected by the MCH-FP program.[25] We use adult height as it is a stock measure of health that is known to reflect genetics as well as the early life health environment. We present results on the interaction of pre-program health endowment terciles with the treatment variable and the p-values of the difference between the first tercile and the two higher terciles in Table 3.

Interaction of the ITT effects with terciles of the pre-program health endowment reveal that the MCH-FP program effects are concentrated in those with the lowest maternal pre-program health endowment, indicating a reduction in the human capital inequality gap for all measures except cognition for which there was no main program effect. However, differences between the bottom and top terciles of the pre-program health endowment are only statistically significant for short stature and education. For the shortest tercile, the ITT effect on height for the pooled sample was 2.03cm (significant at the 1 percent level) and was accompanied by a 10 percentage point (50 percent) reduction in short stature (significant at the 1 percent level) and one extra year of schooling (significant at the 1 percent level). These impacts lead to reductions in the gap between the bottom and top terciles of 0.5cm for height, 7.7 percentage points for short stature, and 0.74 years of schooling. These results indicate that the program lead to sizable reductions in the human capital inequality for measures of height, short stature and education for the first generation.

Disaggregating effects by sex reveal that effects are concentrated in the lowest terciles for both males and females. However, only a few of the gaps between the lowest and two higher terciles of the pre-program health endowment are statistically significant, with females experiencing more reductions in gaps than males. In some cases, the gaps are meaningful, but not statistically significant due to the smaller

24. The Bangladesh Female Secondary Education Stipend Program started in Matlab in 1984 and continued to operate when the 1982-1988 cohort was in secondary school. This program provided a stipend for females in both the treatment and comparison areas to attend secondary school, as well as tuition vouchers.

25. To construct the measure of first generation's mother's height we use height information from MHSS1 collected in 1996. Approximately 18 percent of observations are missing height information in MHSS1 of which 13 percent are filled in using height from MHSS2. Because it is possible people lose height between MHSS1 and MHSS2 and the likely presence of measurement error in height in both survey waves, as a robustness check we remove any observations for which the pre-program human capital tercile differed if MHSS2 rather than MHSS1 data was used. Table A13 shows the results are qualitatively the same.

sample sizes. For males, the only statistically significant gap is between the first and middle terciles for cognition (0.21 SD vs -0.05 SD), and shows that cognition may not have faded since adolescence for everyone. While grades attained in the lowest tercile is almost double the other two terciles (1.59 compared to 0.79 and 0.81), the gaps are not statistically significant. For females, gaps between the lowest tercile and an upper tercile are statistically different for short-stature, but are large for height even if not statistically significant. Specifically, the effect on female height in the lowest tercile is almost double the effect for those in the middle tercile and more than twice that of the upper tercile (2.31cm compared to 1.31cm and 0.96cm). The height gains lead to reductions in short stature of 14 percentage points (50 percent) in the lowest tercile compared to approximately 2 percentage points in the middle and upper terciles.[26].

7 Second Generation Results

7.1 ITT Effects on Human Capital

In this section we examine the intergenerational effects of the program on the human capital of the second generation. These children range from ages 0 to 14 and are the firstborn children of first generation females. We present single-difference ITT effects on the second generation in Table 4. We report results on height-for-age (WHO z-score) and stunting for all ages, but split the sample into ages 0-6 and 7-14 for the cognition index because there are not suitable tests for the entire age range. In Columns 1-4, we pool both sexes and separate males and females in Columns 5-12.

Results pooled by sex show there are no intergenerational effects. However, the pooled results mask important differences by sex. Effects for males are generally small, negative and statistically insignificant. However, firstborn females in the treatment relative to the comparison area experienced a 0.33 SD increase in height-for-age (significant at the 5 percent level), a 14 percentage point (50 percent) decrease in the likelihood of stunting (significant at the 1 percent level), and 0.27 SD increase in the cognitive index (significant at the 10 percent level) for the 7-14 year-olds. able A5 indicates that there are sizable effects for each component of the cognitive index for the 7-14 year-olds, except those associated with memory (digit spans). Female 0-6 year-old cognition was unaffected. These results demonstrate that the MCH-FP program there were sizable effects on the human capital of females in the second generation.

7.2 Selection: ITT Effects on Mortality, Fertility, and Mother Characteristics

To understand if differential selection into or maternal investments in the second-generation could be driving human capital effects for this generation, we examine the program's impact on first-generation female fertility, mortality of their first-born children, and key maternal characteristics that can affect child investments in Table 5. These are important mechanisms that may be linked to biological or behavioral pathways which could lead to selection in the second generation. We estimate single-difference models on all first-generation females in columns 1-5, and a subset of those who have had at least one live birth

26. One reason for the differential reduction in short stature is its higher prevalence in the lowest tercile compared to the middle and upper terciles (26 percent compared to 7 and 2 percent)

in columns 6-12. By restricting the sample to mothers, we are conditioning on an endogenous variable, though the vast majority of women have at least one child (over 83 percent) providing the sample of 677 firstborn children who make up the second generation. We further present results for all second-generation children in Panel A and stratified by sex in Panels B and C to investigate whether selection could underlie the differential pattern of effects for the second generation by sex.

To investigate fertility and mortality selection, we examine multiple dimensions of fertility and a measure of mortality, death of the firstborn child. Measures of fertility include marital status, age at menarche, age at first birth, number of live births, presence of a second child, sex of firstborn child, and lost pregnancies prior to first child born live. ITT effects are close to zero for all measures and not statistically significant for the pooled sample indicating that second-generation results are not likely driven by fertility or mortality selection, and that sex composition of firstborn children is balanced. Panels B and C further show that these indicators of fertility and child mortality are similar between treatment and comparison areas by the sex of the firstborn child. This demonstrates that that fertility and mortality selection is also unlikely to be the driving factors behind the differential pattern on the human capital of the second generation by sex.

We further consider effects on mother level characteristics the may affect investments in children in columns 9-12. We limited the sample to mothers and examine ITT effects on human capital outcomes presented earlier and women's empowerment. We measure empowerment through an index of survey questions related to women's decision making power, attitudes toward gender equivalence in social issues, attitudes toward husband violence and women's mobility (see Appendix 4.1 for more details). We find no effect on the female empowerment index and similar effects on human capital for the sample restricted to females having at least one birth as we did for the unrestricted sample. Specifically, ITT effects on height demonstrated earlier are still present for the sample of first-generation females who become mothers, but effects are small for other outcomes. In addition, mother characteristics do not substantially differ for first born males and females, indicating there is little selection based on these potential mechanisms.

7.3 Mechanisms: Investment in Early Childhood Health

The conceptual framework outlined in Section 3 identifies a number of potential pathways through which the MCH-FP program could affect the human capital accumulation in the second generation. In particular, it highlights that the impact of the program on parental human capital (i.e. on the first generation) may lead to improved human capital of their children (second generation), through changes in behavior that result in additional investments, I_1, in their child's human capital. We focus on program impacts early childhood health investments that are available in the data and which may have been affected by changes in parental behavior. Although there could be biological mechanisms also contributing to our findings, biological markers of this process are still an active area of research and not typically available.

We follow equation 3 and estimate the ITT effects using single difference models. We present results in Figure 3 separately by sex. The graph reports the ITT effects, the 95 percent confidence intervals, and the comparison group means in parentheses next to the point estimates.

We consider three prenatal care investments which includes the number of prenatal checkups, the choice of giving birth at a skilled delivery location, and opting for delivery by a trained birth attendant, in addition to two postnatal indicators, namely the number of vaccinations received and participation in preschool. Health promoting behaviors were close to zero and similar for both sexes with the exception that second-generation firstborns, irrespective of their gender, exhibited a higher likelihood of being delivered in a skilled location and under the care of a trained delivery assistant. To the extent that male and female human capital development react differently to the delivery location and skill of the birth attendant, with females' human capital benefiting more than males, it could explain the differential pattern of results by sex for the second generation.

To further investigate if the mechanisms could explain the underlying pattern of main results, we control for the mechanisms and decompose the primary ITT effects by the mechanisms following (Gelbach, 2016).[27] The mechanisms are potentially endogenous so results are purely descriptive but their inclusion does not lead to a reduction in the treatment effect for either sex, nor provide substantive information when decomposed, further suggesting that any differences in the investments between second generation males and females mechanisms are not driving the results (results not reported).

8 Robustness

We examine the robustness of the effects of MCH-FP on the human capital of both generations, including bounding attrition and different weighting schemes, and show the validity of the research design is supported. We present a key subset of robustness tests in Figures 4 and 5 and detailed results for all tests in Tables A6-A11. Given the treatment and comparison areas are contiguous geographic areas, a key threat to identification is a shock occurring in one area but not the other, such as a disease outbreak or placement of another program. Fortunately, this concern is limited because the study area is small and homogeneous (almost everyone in the comparison area are located within 5 km of the treatment area) and programs and services available in both areas are usually similar because they are in the same district.

First, we test if there are pre-program differences between the treatment and comparison areas. For the first generation, the main results are based on a double-difference model that uses an older cohort to control for pre-existing trends. We further test if there could be pre-program differences for this cohort by examining effects on a similarly-aged sample (24-30 years-old) in 1996 using MHSS1 (Table A6). This cohort was born more than five years prior to the MCH-FP intervention and we find no effects on height, cognition (as measured by the MMSE), or grades attained. Effect sizes are on the whole small, negative and statistically insignificant, indicating there are not pre-existing trends. In addition, we include village fixed-effects with the first generation double difference model to control for time-invariant village unobservables and find results are unchanged (Figure 4).

Next, we estimate double-difference models to test for pre-existing trends among the second generation (Figure 5).[28] Identifying a cohort at a similar point in their development (and hence the same age) but

27. The decomposition method in Gelbach (2016) is similar to stepping in each of the endogenous control variables one at a time to determine how their inclusion reduced the treatment effect, except the decomposition accounts for the fact that the order in which you step in the endogenous controls can matter.

28. We cannot examine effects for similarly aged cohorts in 1996 data for the second generation because this group is born

who is not affected by the program is challenging, as females born prior to the first generation have older children. Instead, we use a less-affected cohort of children who are the same ages (0-14 years-old) but are any birth order children of females born before 1982, the first birth year of the first generation cohort. We focus on these women because their human capital was largely unaffected by MCH-FP program. [29] Since birth order can affect human capital, we would like to restrict the less-affected cohort to only firstborns but there is not a large enough sample. The double-difference results are qualitatively similar to the single-difference results, though, based on effects for height, the single difference estimator provides a slightly more conservative estimate of the program effects.

Second, we use two methods to provide evidence that results are not likely a result of confounding shocks in either the treatment or comparison area. First, there is no evidence of spatially correlated errors across villages (results not reported). This could arise, for example, if there was a health shock such as a disease outbreak in a given year in several villages in the treatment area but not any of the other villages. We also have no knowledge of a disease outbreak that affected just one area from the decades of demographic surveillance data on mortality and disease even in the early years (Fauveau, 1994). Next, we take advantage of the two distinct geographic areas that form the comparison area—one north and one west (see Figure 1)—to show that it is unlikely there was a confounding shock in the comparison area or that geography is driving the results. We repeat the analysis using each comparison block separately and results remain largely unchanged, though there is some loss of significance due to a smaller sample size.

Third, we include an extended set of controls to account for prominent changes that took place in Matlab during 1982-2012 that could potentially bias the results. Details on available programs and control variable construction are in Appendix C. For the first generation, this includes the introduction of a river embankment for irrigation and flood protection in 1987, micro-credit and other programs through BRAC in the 1990s, increased access to primary school, construction of welfare clinics, access to modern medical providers, trained midwives, and discovery of arsenic in some deep tubewells. For the second generation, the set of extended controls accounts for changes in the health and education supply by the time of birth of each child. We also control for the Maternal, Neonatal and Child Health program (MNCH) which began in 2007 and encouraged delivery at a facility with skilled obstetric care and treatment placement overlapped with the treatment area for MCH-FP.[30] Again, results are qualitatively the same for both generations.

Fourth, we control for a limited set of baseline variables that only includes variables that are unbalanced for any sex for the first or second generations. For both generations, these controls include religion, tubewell water, household head and spouse age and education and the results using the smaller control set are qualitatively and statistically similar.

Fifth, given the imbalance of religion between the treatment and comparison areas prior to MCH-FP, we present the main results for only those that identify as Muslim. Again results are similar, though point

---

between 1982-1996, which includes the first generation cohort, as well as those born after similar programs became available in the comparison area but when there may still have been some difference.

29. For the effects on the MCH-FP program on the human capital of women born 1977-1981 at ages 15-19 see (Barham, 2012) and, at ages 31-35 using MHSS2, Table A2

30. Another well-known program in the study area is the Maternal and Infant Nutritional Intervention (Minimat). Minimat was an RCT that enrolled children born prior to the second generation cohort in this study, so it should not affect the single-difference results, but could bias the double-difference results (see Appendix C for more details).

estimates on height for females are slightly smaller for both generations (0.82 cm and 0.278 SD). Standard errors are also higher due to the smaller sample sizes.

Finally, we present results without weights (Panel H of Tables A7 and A10) and adjusting for attrition using Kling-Leibman attrition bounds (Tables A8 and A11).[31] Unweighted results are qualitative and statistically similar, and the bounds show significant results are bounded away from zero and, in most cases, are close to base point estimates. These exercises demonstrate that neither the weighting method nor attrition are driving the results, which is not surprising given the low rates of attrition.

9 Conclusion

Low vaccination rates, high rates of fertility, and under-nutrition are common and related obstacles faced by low-income individuals in different regions of the world today. These obstacles have been exacerbated by the COVID-19 pandemic, conflict and climate change. Existing research shows the combination of these obstacles limit individuals in achieving their full human capital potential (Prendergast and Humphrey, 2014; Grantham-McGregor et al., 2007). Mounting correlational evidence shows that disadvantage in human capital experienced in one generation can persist to the next. It is imperative to examine the causal effects of interventions designed to improve the health environment in childhood on adult human capital and on the next generation to understand the potential role of policy in reducing the spread of inequalities over time.

To address these issues, we leverage unique data that connects three generations, along with a robust causal identification approach to provide, long-term and intergenerational evidence regarding the enduring effects of a widely-implemented public health interventions on human capital in the Matlab area of Bangladesh. Estimates are the combined effects of all program interventions in the MCH-FP program including early child health interventions, such as vaccinations, maternal health, family planning, the service delivery model, and any associated externalities. It is important to recognize that our findings may not generalize to other contexts. However, Bangladesh has been a leader among low- and middle-income countries in decreasing fertility rates and improving human capital, and many of the program interventions continue to form the backbone of preventative health policy worldwide. Moreover, malnutrition rates in Bangladesh are comparable to lower-income countries today, though Matlab's mortality and fertility rates are low and closer rates in middle-income countries.

We find that improving the child health environment increased adult height in the first generation by approximately one centimeter, reducing short-stature, for both sexes, and grades attained for males by almost a year, though there was no effect on cognitive functioning. Intergenerational effects reveal daughters experienced increased height, with ITT effects larger than the first generation (1.6 cm), and a reduction in stunting of 50 percent, essentially eliminate stunting within this cohort. Cognitive effects of 0.26 SD also resurfaced among the 7-14 year olds. The resurgence of program impacts on cognition in the second generation underscores the importance of considering various dimensions of human capital

31. Kling-Leibman attrition bounds assigns attritors the sample mean of the outcome plus or minus 0.10 and 0.25 standard deviations for continuous variables. For binary variables, we assign attritors the worst case scenarios of either all zeros or all ones (Kling, Liebman and Katz, 2007).

development. Human capital measures respond differently to one's environment, and their sensitivity varies across different life stages, making it important to assess affects and at different points in the life cycle and across generations. Additionally, the finding that effects on height and cognition are even larger for the second generation than the first, while not statistically different between generations, suggests the potential for impacts for early childhood health investments to grow over generations.

A unique strength of this paper is the ability to link data across generations to explore how pre-program health endowments (proxied by height of the first generation's mothers) interact with a policy to affect human capital over generations. The finding that the largest improvements in adult human capital for the first generation are experienced by those that have the lowest levels of pre-program health endowment in Section 6.2, demonstrates that, even though the MCH-FP program was not targeted to low endowment individuals, the program was able to break the positive child-parent human capital correlations, reducing human capital inequality.

Behavioral and biological mechanisms link human capital between the generations, but are not well understood. In this study, we are able to rule out that effects are driven by child mortality or fertility selection, and find no effects on typical early childhood health investments or maternal empowerment that explain the pattern of second-generation results by sex. Gender effects are also not a result of second generation females initially having worse human capital than males, leaving more room for improvement. There are numerous potential mechanisms, many not collected in datasets, and further research is needed to understand the key mechanisms and sex-specific dimensions of the intergenerational transmission of human capital.

Together the findings demonstrate that investments in the early health environment can generate substantial and enduring effects on human capital. Policymakers who fail to consider these intergenerational and distributional effects when allocating resources for children may inadvertently underinvest, missing an opportunity to mitigate human capital disparities and potentially reshape economic outcomes of future generations.

References

Alam, Nurul, M Zahirul Haq, and Peter Kim Streatfield. 2010. "Spatio-temporal patterns of under-five mortality in Matlab HDSS in rural Bangladesh." *Global Health Action*, 3(1): 5252.

Almond, Douglas, and Bhashkar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy." *American Economic Journal: Applied Economics*, 3(4): 56–85.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Athey, Susan, and Guido W Imbens. 2017. "The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*. Vol. 1, 73–140. Elsevier.

Attanasio, Orazio, Costas Meghir, and Emily Nix. 2020. "Human Capital Development and Parental Investment in India." *The Review of Economic Studies*, 87(6): 2511–2541.

Attanasio, Orazio P. 2015. "The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Policies." *Journal of the European Economic Association*, 13(6): 949–997.

Bailey, Martha J. 2013. "Fifty years of family planning: new evidence on the long-run effects of increasing access to contraception." National Bureau of Economic Research.

Barham, Tania. 2012. "Enhancing Cognitive Functioning: Medium-Term Effects of a Health and Family Planning Program in Matlab." *American Economic Journal: Applied Economics*, 4(1): 245–273.

Barham, Tania, and Randall Kuhn. 2014. "Staying for Benefits The Effect of a Health and Family Planning Program on Out-Migration Patterns in Bangladesh." *Journal of Human Resources*, 49(4): 982–1013. Publisher: University of Wisconsin Press.

Barham, Tania, Brachel Champion, Andrew D Foster, Jena D Hamadani, Warren C Jochem, Gisella Kagy, Randall Kuhn, Jane Menken, Abdur Razzaque, Elisabeth Dowling Root, et al. 2021. "Thirty-five Years Later: Long-term Effects of the Matlab Maternal and Child Health/Family Planning Program on Older Women's Well-being." *Proceedings of the National Academy of Sciences*, 118(28).

Barham, Tania, Randall Kuhn, and Patrick Turner. 2019. "No Place Like Home: Long-Run Impacts of Early Child Health and Family Planning on Labor and Migration Outcomes." *Working Paper*.

Barr, Andrew, and Chloe Gibbs. 2017. "Breaking the cycle? Intergenerational effects of an anti-poverty program in early childhood."

Behrman, Jere R, Maria C Calderon, Samuel H Preston, John Hoddinott, Reynaldo Martorell, and Aryeh D Stein. 2009. "Nutritional Supplementation in Girls Influences the Growth of Their Children: Prospective Study in Guatemala." *The American Journal of Clinical Nutrition*, 90(5): 1372–1379.

Bhalotra, Sonia, and Samantha Rawlings. 2013. "Gradients of the Intergenerational Transmission of Health in Developing Countries." *Review of Economics and Statistics*, 95(02): 660–672.

Bhatia, Shushum, W. H. Mosley, A. S. G. Faruque, and J. Chakraborty. 1980. "The Matlab Family Planning-Health Services Project." *Studies in Family Planning*, 11(6).

Bhuiya, Abbas, Bogdan Wojtyniak, Stan D'Souza, Lutfun Nahar, and Kashem Shaikh. 1987. "Measles case fatality among the under-fives: a multivariate analysis of risk factors in a rural area of Bangladesh." *Social science & medicine*, 24(5): 439–443.

Black, Sandra, and Paul Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*. 1487–1541.

Chowdhury, Uttam K, Bhajan K Biswas, T Roy Chowdhury, Gautam Samanta, Badal K Mandal, Gautam C Basu, Chitta R Chanda, Dilip Lodh, Khitish C Saha, Subhas K Mukherjee, et al. 2000. "Groundwater Arsenic Contamination in Bangladesh and West Bengal, India." *Environmental Health Perspectives*, 108(5): 393–397.

Cleland, John, James F Phillips, Sajeda Amin, and Gholam Mostafa Kamal. 1994. "The determinants of reproductive change in Bangladesh." *Washington, DC: The World Bank*, 72: 1–55.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review*, 97(2): 31–47.

Das, Sumon Kumar, Mohammod Jobayer Chisti, Mohammad Abdul Malek, Jui Das, Mohammed Abdus Salam, Tahmeed Ahmed, Abdullah Al Mamun, and Abu Syed Golam Faruque. 2015. "Changing childhood malnutrition in Bangladesh: trends over the last two decades in urban–rural differentials (1993–2012)." *Public Health Nutrition*, 18(10): 1718–1727.

Deming, David. 2009. "Early Childhood Intervention and Life-cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics*, 1(3): 111–34.

Drake, Amanda J, and Lincoln Liu. 2010. "Intergenerational transmission of programmed effects: public health consequences." *Trends in Endocrinology & Metabolism*, 21(4): 206–213.

Driessen, J, A Razzaque, D Walker, and David Canning. 2015. "The effect of childhood measles vaccination on school enrolment in Matlab, Bangladesh." *Applied Economics*, 47(55): 6019–6040.

East, Chloe N., Sarah Miller, Marianne Page, and Laura R. Wherry. 2023. "Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health." *American Economic Review*, 113(1): 98–135.

El Arifeen, Shams, Eva-Charlotte Ekström, Edward A Frongillo, Jena Hamadani, Ashraful I Khan, Ruchira T Naved, Anisur Rahman, Rubhana Raqib, Kathleen M Rasmussen, Katarina Ekholm Selling, et al. 2018. "Cohort Profile: The Maternal and Infant Nutrition Interventions in Matlab (MINIMat) Cohort in Bangladesh." *International Journal of Epidemiology*, 47(6): 1737.

Fauveau, Vincent. 1994. *Matlab: Women, Children and Health*. Dhaka; icddr,b.

Fitzgerald, Eamon, Kahyee Hor, and Amanda J Drake. 2020. "Maternal influences on fetal brain development: The role of nutrition, infection and stress, and the potential for intergenerational consequences." *Early human development*, 150: 105190.

Frankenburg, W, J Dodds, AW Fandal, E Kauk, and M Cohrs. 1975. "The Denver Developmental Screening Test–Revised (DDST)." *Denver, CO: University of Colorado Medical Center*.

Gadroen, Kartini, Caitlin N Dodd, Gwen MC Masclee, Maria AJ De Ridder, Daniel Weibel, Michael J Mina, Bryan T Grenfell, Miriam CJM Sturkenboom, David AMC Van De Vijver, and Rik L De Swart. 2018. "Impact and longevity of measles-associated immune suppression: a matched cohort study using data from the THIN general practice database in the UK." *BMJ Open*, 8(11): e021465.

Garcia, Jorge Luis, James J. Heckman, and Victor Ronda. 2021. "The Lasting Effects of early Childhood Education and Promoting the skills and Social Mobility of Disadvantaged African Americans." *NBER Working Paper No 202183*.

Gelbach, Jonah B. 2016. "When do covariates matter? And which ones, and how much?" *Journal of Labor Economics*, 34(2): 509–543.

Gluckman, Peter D, Mark A Hanson, and Alan S Beedle. 2007. "Early Life Events and Their Consequences for Later Disease: A Life History and Evolutionary Perspective." *American Journal of Human Biology*, 19(1): 1–19.

Gluckman, Peter D, Tatjana Buklijas, and Mark A Hanson. 2016. "The Developmental Origins of Health and Disease (DOHaD) Concept: Past, Present, and Future." In *The Epigenome and Developmental Origins of Health and Disease*. 1–15. Elsevier.

Grantham-McGregor, Sally M, Lia C Fernald, and Kavita Sethuraman. 1999*a*. "Effects of Health and Nutrition on Cognitive and Behavioural Development in Children in the First Three Years of Life: Part 1: Low Birthweight, Breastfeeding, and Protein-Energy Malnutrition." *Food and Nutrition Bulletin*, 20(1): 53–75.

Grantham-McGregor, Sally M, Lia C Fernald, and Kavita Sethuraman. 1999*b*. "Effects of Health and Nutrition on Cognitive and Behavioural Development in Children in the First Three Years of Life: Part 2: Infections and Micronutrient Deficiencies: Iodine, Iron, and Zinc." *Food and Nutrition Bulletin*, 20(1): 76–99.

Grantham-McGregor, Sally, Yin Bun Cheung, Santiago Cueto, Paul Glewwe, Linda Richter, Barbara Strupp, International Child Development Steering Group, et al. 2007. "Developmental Potential in the First 5 Years For Children in Developing Countries." *The Lancet*, 369(9555): 60–70.

Grönqvist, Erik, Björn Öckert, and Jonas Vlachos. 2017. "The Intergenerational Transmission of Cognitive and Noncognitive Abilities." *Journal of Human Resources*, 52(4): 887–918.

Heckman, James J. 2007*a*. "The economics, technology, and neuroscience of human capability formation." *Proceedings of the National Academy of Sciences*, 104(33): 13250.

Heckman, James J. 2007*b*. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences*, 104(33): 13250–13255.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina. 2008. "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends." *The B.E. Journal of Economic Analysis & Policy*, 7(2).

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review*, 106(4): 903–34.

Huber, Douglas H, and Atiqur Rahman Khan. 1979. "Contraceptive Distribution in Bangladesh Villages: the Initial Impact." *Studies in Family Planning*, 10(8/9): 246–253.

icddrb. 1974. "Matlab 1974 Census."

icddrb. 1982. "Matlab 1982 Census."

Icddr, B. 2007. "Health and demographic surveillance system-Matlab." *2005 Socioeconomic Census*.

Imbens, Guido W, and Jeffrey M Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1): 5–86.

Ismail, Zahinoor, Tarek K Rajji, and Kenneth I Shulman. 2010. "Brief Cognitive Screening Instruments: An Update." *International Journal of Geriatric Psychiatry*, 25(2): 111–120.

Joshi, Shareen, and T Paul Schultz. 2013. "Family Planning and Women's and Children's Health: Long-term Consequences of an Outreach Program in Matlab, Bangladesh." *Demography*, 50(1): 149–180.

Kaplan, RM, and DP Saccuzzo. 2009. "Standardized Tests in Education, Civil Service, and the Military." *Psychological testing: Principles, Applications, and Issues*, 7: 325–327.

Khan, M Mahmud, and Richard Yoder. 1998. *Expanded Program on Immunization in Bangladesh: Cost, Cost-effectiveness, and Financing Estimates.* Bethesda, MD: Partnerships for Health Reform Project, Abt Associates Inc., 1998.

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

Koenig, M. A., M. A. Khan, B. Wojtyniak, J. D. Clemens, J. Chakraborty, V. Fauveau, J. F. Phillips, J. Akbar, and U. S. Barua. 1990. "Impact of Measles Vaccination on Childhood Mortality in Rural Bangladesh." *Bulletin of the World Health Organization*, 68(4): 441–447.

Koenig, Michael A., Vincent Fauveau, and Bogdan Wojtyniak. 1991. "Mortality Reductions from Health Interventions: The Case of Immunization in Bangladesh." *Population and Development Review*, 17(1).

Maluccio, John A, John Hoddinott, Jere R Behrman, Reynaldo Martorell, Agnes R Quisumbing, and Aryeh D Stein. 2009. "The Impact of Improving Nutrition During Early Childhood on Education Among Guatemalan Adults." *The Economic Journal*, 119(537): 734–763.

Menken, Jane, and James F Phillips. 1990. "Population Change in a Rural Area of Bangladesh, 1967-87." *The Annals of the American Academy of Political and Social Science*, 510(1): 87–101.

Miller, DL. 1964. "Frequency of Complications of Measles, 1963." *British Medical journal*, 2(5401): 75.

Miller, Sarah, and Laura R Wherry. 2019. "The Long-Term Effects of Early Life Medicaid Coverage." *Journal of Human Resources*, 54(3): 785–824.

Mina, Michael J, C Jessica E Metcalf, Rik L De Swart, ADME Osterhaus, and Bryan T Grenfell. 2015. "Long-Term Measles-Induced Immunomodulation Increases Overall Childhood Infectious Disease Mortality." *Science*, 348(6235): 694–699.

Perry, Robert T, and Neal A Halsey. 2004. "The Clinical Significance of Measles: a Review." *The Journal of Infectious Diseases*, 189(Supplement_1): S4–S16.

Phillips, James F., Ruth Simmons, J. Chakraborty, and A. I. Chowdhury. 1984. "Integrating Health Services into an MCH-FP Program: Lessons from Matlab, Bangladesh." *Studies in Family Planning*, 15(4).

Phillips, James F, Wayne S Stinson, Shushum Bhatia, Makhlisur Rahman, and Jyotsnamoy Chakraborty. 1982. "The Demographic Impact of the Family Planning–Health Services Project in Matlab, Bangladesh." *Studies in Family Planning*, 131–140.

Prendergast, Andrew J, and Jean H Humphrey. 2014. "The Stunting Syndrome in Developing Countries." *Paediatrics and International Child Health*, 34(4): 250–265.

Proos, LA. 1993. "Anthropometry in adolescence-secular trends, adoption, ethnic and environmental differences." *Hormone Research in Paediatrics*, 39(Suppl. 3): 18–24.

Rahman, Anisur, Allisyn Moran, Jesmin Pervin, Aminur Rahman, Monjur Rahman, Sharifa Yeasmin, Hosneara Begum, Harunor Rashid, Mohammad Yunus, Daniel Hruschka, et al. 2011. "Effectiveness of an Integrated Approach To Reduce Perinatal Mortality: Recent Experiences From Matlab, Bangladesh." *BMC Public Health*, 11(1): 1–14.

Rahman, O, J Menken, A Foster, C Peterson, MN Khan, R Kuhn, and P Gertler. 1999. "The Matlab Health and Socio-economic Survey: Overview and User's Guide, 1999."

Reddy, Vinodini. 1987. "Interaction Between Nutrition and Measles." *The Indian Journal of Pediatrics*, 54(1): 53–57.

Richter, André, and Per Olof Robling. 2013. "Multigenerational Effects of the 1918-19 Influenza Pandemic in Sweden." *Working paper / Swedish Institute for Social Research (SOFI)*.

Sandovici, Ionel, Denise S Fernandez-Twinn, Antonia Hufnagel, Miguel Constância, and Susan E Ozanne. 2022. "Sex differences in the intergenerational inheritance of metabolic traits." *Nature Metabolism*, 4(5): 507–523.

Shamsuddin, Mrittika. 2015. "Labour Market Effects of a Female Stipend Programme in Bangladesh." *Oxford Development Studies*, 43(4): 425–447.

Thompson, Owen. 2014. "Genetic Mechanisms in the Intergenerational Transmission of Health." *Journal of Health Economics*, 35: 132–146.

van Ginneken, Jeroen, Radheshyam Bairagi, Andres de Francisco, Asdul Mazid Sardar, and Patrick Vaughan. 1998. *Health and demographic surveillance in Matlab: past, present and future.* International Centre for Diarrhoeal Disease Research, Bangladesh.

Venkataramani, Atheendar S. 2011. "The Intergenerational Transmission of Height: Evidence From Rural Vietnam." *Health Economics*, 20(12): 1448–1467.

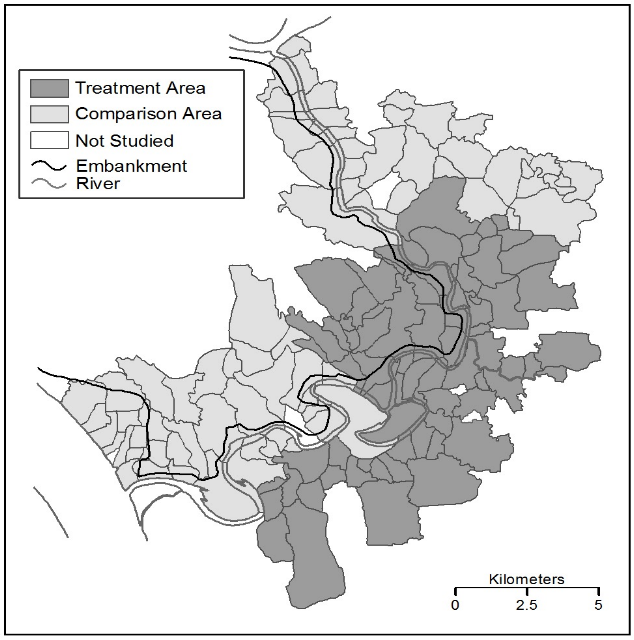
Walker, Susan P, Susan M Chang, Amika S Wright, Rodrigo Pinto, James J Heckman, and Sally M Grantham-McGregor. 2021. "Cognitive, Psychosocial, and Behaviour Gains at Age 31 Years From the Jamaica Early Childhood Stimulation Trial." *Journal of Child Psychology and Psychiatry*.

Walker, Susan P, Theodore D Wachs, Julie Meeks Gardner, Betsy Lozoff, Gail A Wasserman, Ernesto Pollitt, Julie A Carter, International Child Development Steering Group, et al. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries." *The Lancet*, 369(9556): 145–157.

Wasserman, Gail A, Xinhua Liu, Faruque Parvez, Habibul Ahsan, Pam Factor-Litvak, Alexander van Geen, Vesna Slavkovich, Nancy J Lolacono, Zhongqi Cheng, Iftikhar Hussain, et al. 2004. "Water Arsenic Exposure and children's Intellectual Function in Araihazar, Bangladesh." *Environmental Health Perspectives*, 112(13): 1329–1333.

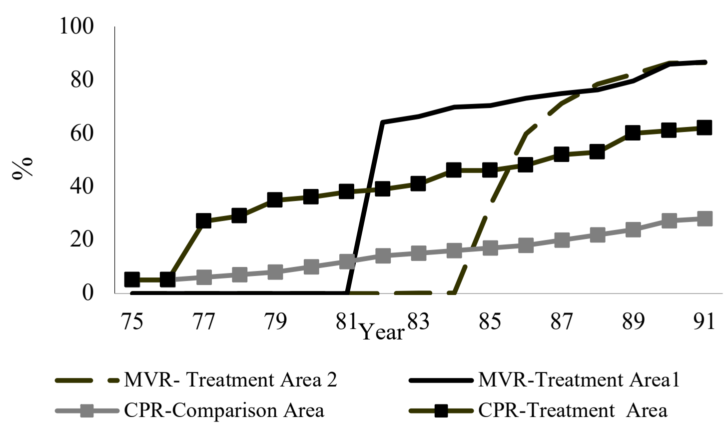
Figures and Tables

FIGURE 1 — MAP OF MATLAB STUDY AREA



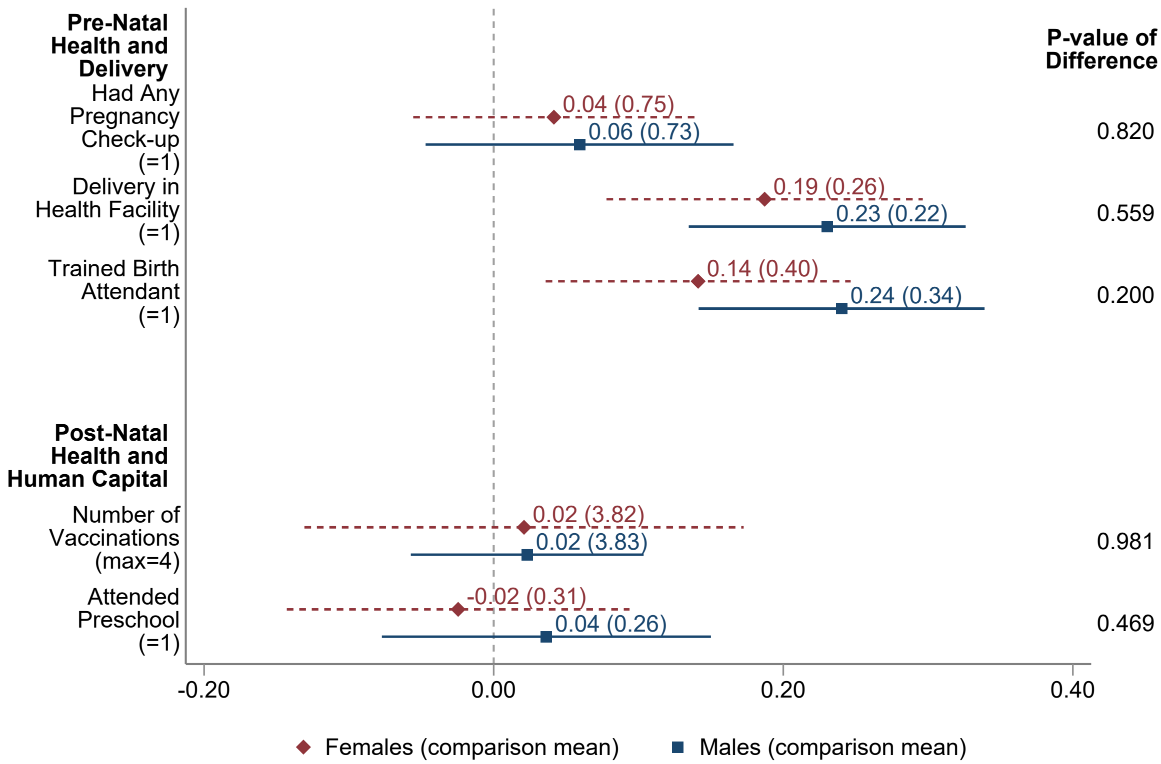
Notes: This figure shows the Matlab study site in Bangladesh where the MCH-FP program was implemented. Subdivisions indicate distinct villages where an individual's household head lived in 1974 used to create the intent-to-treat measure.

FIGURE 2 — TRENDS IN CONTRACEPTIVE PREVALENCE RATE (CPR) AND MEASLES VACCINATION RATES (MVR)



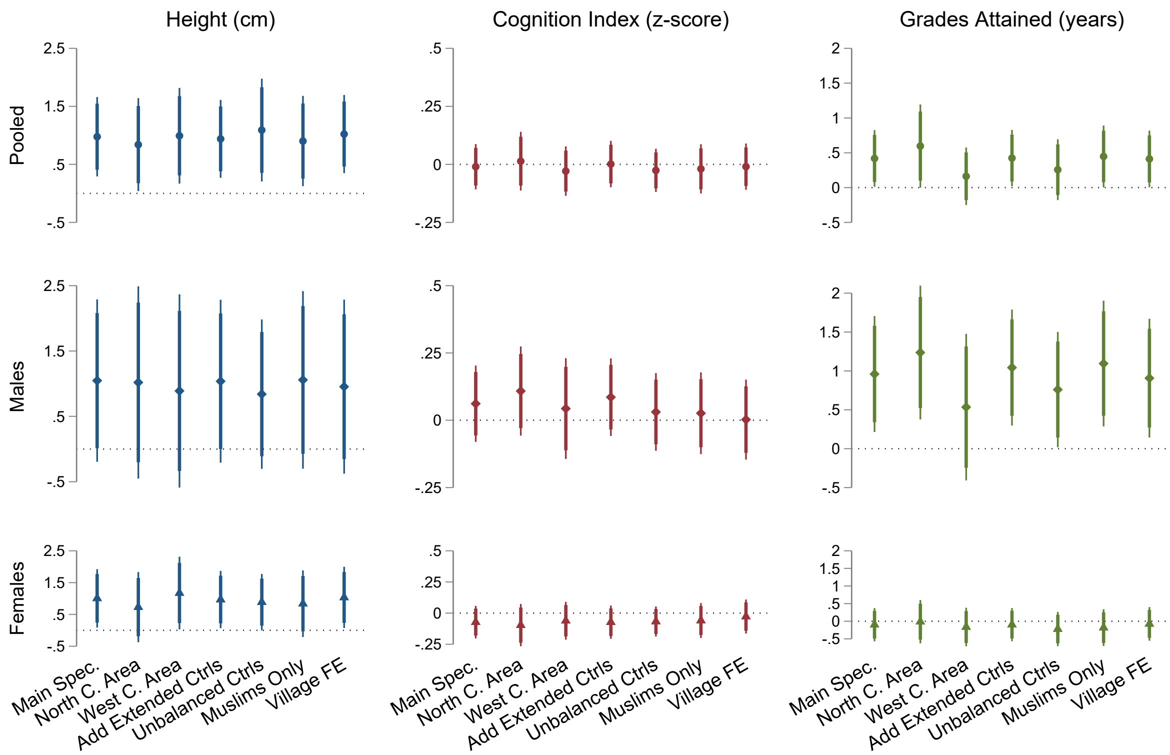
Notes: This figure shows the Contraceptive Prevalence Rate (CPR) and Measles Vaccination Rate (MVR) for children aged 12-59 months by year. Contraceptive use data is from van Ginneken et al. (1998); Measles vaccination data from icddr,b Record Keeping System.

FIGURE 3 — SECOND GENERATION: POTENTIAL MECHANISMS



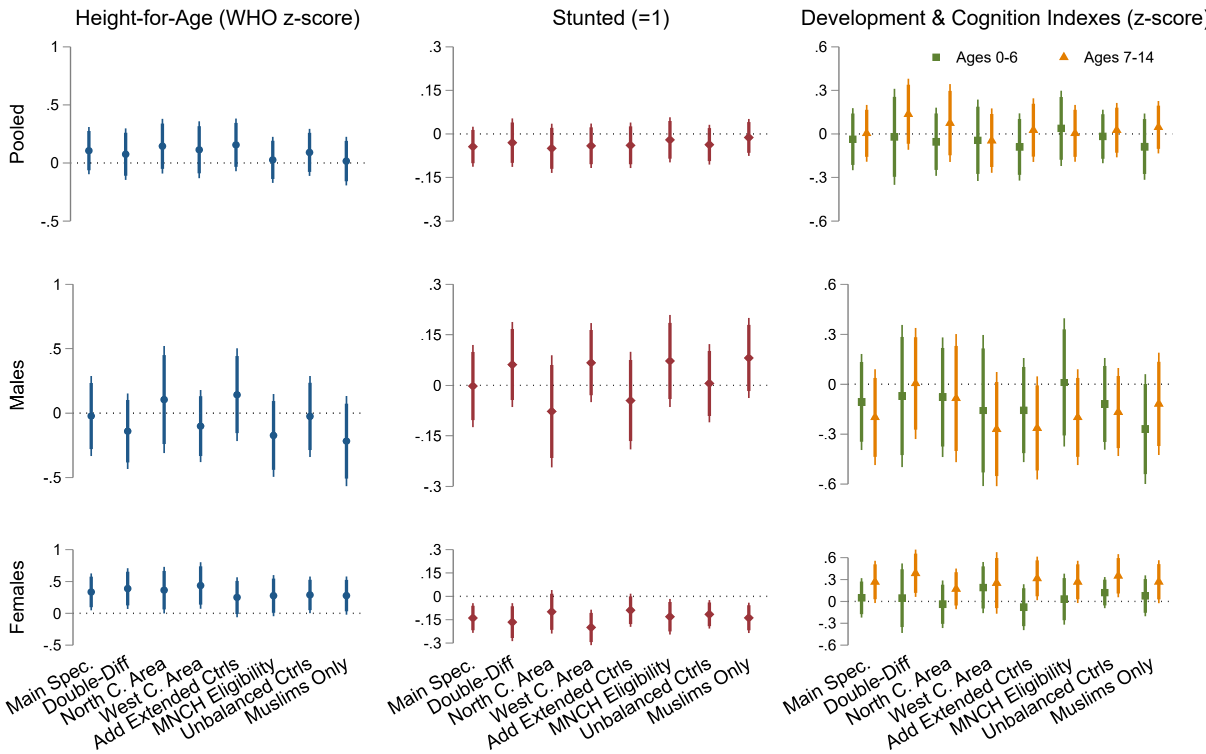
Notes: This figures shows the ITT effects for firstborn children born to 1982-88 mothers from Equation 3, separately by males and females. The mean of the comparison group is displayed in parentheses and the bars represent 95% confidence intervals. *Delivery in Health Facility* equals one if the pregnancy was delivered in a certified medical center. *Trained Birth Attendant* equals one if the delivery was performed by a physician, nurse or other healthcare worker. *Number of Vaccinations* is the total types of vaccinations received out of four (BCG, Polio, DPT and MMR). See Appendix B for a fuller description of each mechanism.

FIGURE 4 — FIRST GENERATION: SELECTED ROBUSTNESS CHECKS



Notes: This figure shows ITT effects on first generation human capital are robust to alternate specifications. Thinner bars represent the 95% confidence interval while thicker bars represent the 90% confidence interval. *Main Spec* shows our preferred model for context. *North C. Area* and *West C. Area* restrict the sample of comparison units to those from the northern and western comparison areas, respectively (see Figure 1). *Add Extended Ctrls* includes additional controls to account for changes over time in the supply of health and education. *Unbalanced Ctrls* limits the pre-program baseline control set to only those that are unbalanced. *Muslims Only* limits the sample to Muslims which are over represented in the comparison area. *Village FE* adds fixed effects for pre-program village to account for village-specific differences. See 8 and Appendix D for a full description of each model and additional robustness checks.

FIGURE 5 — SECOND GENERATION: SELECTED ROBUSTNESS CHECKS



Notes: This figure shows ITT effects on second generation human capital are robust to alternate specifications. Child Development Index (left, green) is available for children aged 0-6 while Cognition Index (right, yellow) is only available for children aged 7-14 (see Appendix B.3 for a description of indexes). Thinner bars represent the 95% confidence interval while thicker bars represent the 90% confidence interval. *Main Spec* shows our preferred model for context. *Double-Diff* shows the ITT estimates from a double-difference model using same-aged children of women born 1977-81 as the unaffected group. *North C. Area* and *West C. Area* restrict the sample of comparison units to those from the northern and western comparison areas, respectively (see Figure 1). *Add Extended Ctrls* includes additional controls to account for changes over time in the supply of health and education. *MNCH Eligibility* adds a control for the 2007 Maternal, Neonatal and Child Health Project (MNCH). *Unbalanced Ctrls* limits the pre-program baseline control set to only those that are unbalanced. *Muslims Only* limits the sample to Muslims which are overrepresented in the comparison area. See Section 8 and Appendix D for a full description of each model and additional robustness checks.

TABLE 1 — FIRST GENERATION: TREATMENT AND COMPARISON AREA BALANCE

	Treatment Area			Comparison Area			Difference		
	Mean	SD	Obs	Mean	SD	Obs	Mean	T-stat	Mean/SD
Panel A: 1982-88 Cohort Parent Human Capital									
Mother's height (cm)	149.49	9.13	702	150.11	7.81	847	-0.62	-1.41	-0.05
Father's height (cm)	160.96	7.93	616	160.95	8.89	767	0.01	0.02	0.00
Mother's MMSE (z-score)	-0.40	1.54	650	-0.44	1.43	785	0.04	0.50	0.02
Father's MMSE (z-score)	-0.44	1.14	508	-0.55	1.62	605	0.11	1.34	0.06
Mother's education (years)	1.79	3.32	715	1.60	3.73	857	0.19	1.04	0.04
Father's education (years)	3.57	4.27	660	3.20	5.13	793	0.36	1.48	0.05
Panel B: 1947-69 Pre-program Cohort Human Capital in 1996									
Height (cm)	154.53	11.86	1200	154.93	7.18	1373	-0.40	-1.01	-0.03
MMSE (z-score)	-0.01	1.31	1230	0.02	1.07	1404	-0.03	-0.62	-0.02
Education (years)	2.75	5.01	1321	2.46	4.60	1450	0.29	1.60	0.04
Panel C: 1982-88 Cohort Individual and Pre-Program Household Characteristics									
Male (=1)	0.51	0.46	742	0.50	0.46	896	0.01	0.61	0.02
Birth year	1984.97	1.76	742	1984.93	1.89	896	0.03	0.38	0.01
Islamic (=1)	0.84	0.81	742	0.95	0.31	896	-0.12	-3.81	-0.14
HH Bari size	8.63	10.25	742	8.03	10.40	896	0.61	1.18	0.04
HH Family size	6.79	3.51	742	6.54	3.08	896	0.25	1.53	0.05
Latrine (=1)	0.81	0.67	742	0.88	0.73	896	-0.07	-2.03	-0.07
Owns a lamp (=1)	0.62	0.65	742	0.59	0.74	896	0.03	0.92	0.03
Owns a watch (=1)	0.14	0.42	742	0.15	0.45	896	-0.01	-0.41	-0.01
Owns a radio (=1)	0.08	0.34	742	0.07	0.28	896	0.01	0.35	0.01
Number of cows	1.41	2.25	742	1.33	1.98	896	0.08	0.71	0.03
Number of boats	0.65	0.95	742	0.65	0.99	896	-0.01	-0.16	-0.01
Wall tin or tinmix (=1)	0.29	0.56	742	0.29	0.60	896	-0.01	-0.23	-0.01
Tin roof (=1)	0.82	0.50	742	0.83	0.54	896	-0.01	-0.40	-0.01
Number of rooms per capita	0.22	0.11	742	0.22	0.12	896	0.00	0.80	0.03
Drinking water, tubewell (=1)	0.30	0.88	742	0.16	0.74	896	0.15	3.59	0.13
Drinking water, tank (=1)	0.40	1.07	742	0.33	1.31	896	0.07	1.11	0.04
HH age	48.20	16.84	742	45.93	16.35	896	2.27	2.75	0.10
HH <2 years of education (=1)	0.60	0.56	742	0.60	0.77	896	0.00	0.00	0.00
HH works in agriculture (=1)	0.61	0.67	742	0.56	0.73	896	0.04	1.21	0.04
HH works in fishing (=1)	0.07	0.47	742	0.07	0.35	896	0.00	-0.08	0.00
HH spouse's age	37.50	15.62	742	35.62	15.03	896	1.88	2.46	0.09
HH spouse <2 years of education (=1)	0.86	0.43	742	0.87	0.48	896	-0.01	-0.59	-0.02
1982 Land size	10.02	19.28	742	10.99	19.81	896	-0.97	-1.00	-0.04

Notes: This table shows balance of pre-program individual, parental and household characteristics between treatment and comparison areas. Panel A shows the human capital of mothers and fathers of the 1982-88 cohort in 1996, filling in missing values with 2012 values. Panel B includes all respondents born 1947-69 with non-missing 1996 human capital. Panel C includes male and female respondents in the 1982-88 age cohort who have education data in MHSS2 and displays individual characteristics and characteristics of the 1982-88 cohorts' pre-program household. Unless otherwise noted, household characteristics come from the 1974 census. Standard deviations (SD) are clustered at the pre-program village level. Household head and spouse age are reported, but these variables are likely affected by the family planning program which increased birth intervals and decreased family sizes in the treatment area (Barham et al., 2021). Observations are weighted to correct for attrition between birth and the 2012 MHSS2 survey. HH=household head.

TABLE 2 — FIRST GENERATION: ITT EFFECTS ON HUMAN CAPITAL

	Pooled				Males				Females			
	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treat $\times \mathbb{1}$(Born 1982-88)	0.977	-0.041	-0.010	0.419	1.047	-0.037	0.061	0.960	1.005	-0.042	-0.072	-0.098
	(0.346)	(0.022)	(0.049)	(0.206)	(0.629)	(0.029)	(0.072)	(0.377)	(0.464)	(0.034)	(0.066)	(0.239)
Treat	0.022	0.021	0.017	0.029	-0.022	0.023	-0.026	-0.049	-0.025	0.019	0.062	0.084
	(0.283)	(0.018)	(0.042)	(0.159)	(0.421)	(0.022)	(0.057)	(0.262)	(0.343)	(0.025)	(0.046)	(0.164)
Rand. Inf. P-value	0.03	0.03	0.78	0.09	0.25	0.13	0.45	0.11	0.03	0.11	0.24	0.85
Sharpened q-value	0.02	0.08	0.27	0.08	0.17	0.24	0.24	0.05	0.15	0.38	0.38	0.58
1982-88 Mean	157.32	0.09		7.11	163.40	0.08		7.00	151.31	0.11		7.23
Obs	4211	4211	4037	4340	1893	1893	1723	1952	2318	2318	2314	2388

Notes: This table shows double difference ITT effects on the stock of human capital of the first generation. Standard errors are clustered at the pre-program village level and reported in parentheses. Random inference p-values are calculated for 1982-88 cohort effects from a distribution of test statistics constructed by randomly reassigning treatment status to villages over 10,000 permutations while maintaining geographic contiguity. Sharpened q-values are p-values adjusted for the false discovery rate from testing multiple hypotheses. 1982-88 means are for the comparison group. All regressions include individual and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Individual characteristics include year of birth fixed effects, age cohort fixed effects, and religion. Pre-intervention characteristics include all characteristics in Table 1. Short stature is defined as height less than 155cm for males and 145cm for females. The Cognition Index includes MMSE, Digit Spans Forward and Backward, and Ravens Progressive Matrix scores. Male Cognition Index sample sizes are smaller than female because tests were not administered in the phone survey.

TABLE 3 — FIRST GENERATION: DISTRIBUTIONAL ITT EFFECTS BY MOTHER'S HEIGHT

	Pooled				Males				Females			
	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treat × $\mathbb{1}$(Born 1982-88) × Shortest Tercile	2.028	-0.102	0.084	1.007	1.831	-0.056	0.210	1.590	2.327	-0.144	-0.069	0.423
	(0.490)	(0.036)	(0.073)	(0.366)	(0.801)	(0.043)	(0.113)	(0.614)	(0.769)	(0.055)	(0.102)	(0.404)
Treat × $\mathbb{1}$(Born 1982-88) × Middle Tercile	0.890	-0.028	-0.037	0.140	0.709	-0.038	-0.046	0.789	1.313	-0.016	0.011	-0.544
	(0.550)	(0.029)	(0.081)	(0.343)	(0.782)	(0.040)	(0.117)	(0.574)	(0.735)	(0.044)	(0.100)	(0.383)
Treat × $\mathbb{1}$(Born 1982-88) × Tallest Tercile	1.508	-0.025	-0.006	0.262	2.038	-0.024	0.116	0.805	0.963	-0.015	-0.110	-0.206
	(0.548)	(0.023)	(0.069)	(0.289)	(0.936)	(0.032)	(0.092)	(0.499)	(0.677)	(0.036)	(0.090)	(0.377)
Treat	-0.009	0.022	0.017	0.030	-0.055	0.024	-0.025	-0.047	-0.053	0.020	0.061	0.085
	(0.283)	(0.018)	(0.042)	(0.158)	(0.424)	(0.022)	(0.058)	(0.261)	(0.342)	(0.025)	(0.046)	(0.164)
Tercile Differences												
Shortest – Middle	1.138	-0.074	0.121	0.867	1.122	-0.018	0.256	0.801	1.014	-0.128	-0.080	0.966
P(Shortest = Middle)	0.059	0.059	0.225	0.069	0.223	0.727	0.084	0.264	0.280	0.030	0.515	0.096
Shortest – Tallest	0.520	-0.077	0.090	0.745	-0.207	-0.031	0.094	0.785	1.365	-0.129	0.040	0.628
P(Shortest = Tallest)	0.446	0.032	0.310	0.090	0.820	0.469	0.493	0.300	0.175	0.021	0.726	0.225
1982-1988 Tercile Means												
Shortest	154.47	0.19	0.33	6.64	160.81	0.13	0.25	6.61	148.09	0.26	0.41	6.66
Middle	157.38	0.07	0.42	7.02	163.25	0.06	0.40	6.94	151.44	0.07	0.44	7.09
Tallest	160.03	0.02	0.53	7.73	166.15	0.03	0.38	7.37	154.05	0.02	0.68	8.07
Obs	4122	4122	3953	4251	1848	1848	1683	1907	2274	2274	2270	2344

Notes: This table shows double difference ITT effects on the stock of human capital, by terciles of 1992-98 cohort mother's height. Terciles are defined by a mother's height less than 147.5cm (shortest), 147.5-152.1cm (middle) or greater than 152.1cm (tallest). 50.7 percent of respondents in the shortest tercile had a mother of short stature (height less than 145cm). Shortest – Middle is the difference in treatment effects for people in the shortest and middle tercile and P(Shortest = Middle) is the p-value from a two-sided test that the treatment effects are equal. Similarly for Shortest – Tallest. Tercile means are the comparison group. All regressions include individual and pre-intervention characteristics and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE 4 — SECOND GENERATION: ITT EFFECTS ON HUMAN CAPITAL

	Pooled				Males				Females			
	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)
	Ages 0-14		Ages 0-6	Ages 7-14	Ages 0-14		Ages 0-6	Ages 7-14	Ages 0-14		Ages 0-6	Ages 7-14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mother's Treat	0.106	-0.044	-0.036	0.004	-0.022	-0.002	-0.106	-0.198	0.335	-0.139	0.048	0.268
	(0.103)	(0.035)	(0.108)	(0.098)	(0.157)	(0.062)	(0.145)	(0.144)	(0.146)	(0.048)	(0.136)	(0.145)
Rand. Inf. P-value	0.243	0.192	0.910	0.958	0.867	0.974	0.678	0.133	0.008	0.072	0.815	0.048
Sharpened q-value	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.038	0.019	0.202	0.051
Mean	-1.262	0.258			-1.221	0.242			-1.302	0.273		
Obs	629	629	334	262	311	311	159	135	318	318	175	127

Notes: This table shows single-difference ITT effects on the stock of human capital of second generation children of females born 1982-88 using their mother's treatment status. Standard errors in parentheses are clustered at the mother's pre-program village level. Random inference p-values are calculated from a distribution of test statistics constructed by randomly reassigning mother's treatment status to villages over 10,000 permutations while maintaining geographic contiguity. Sharpened q-values are p-values adjusted for the false discovery rate from testing multiple hypotheses. Means are for children of comparison group mothers. All regressions include child's six-month age fixed effects, mother's year of birth fixed effects, mother's individual and pre-intervention characteristics listed in Table A4. All models are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Height-for-age is calculated using WHO standards. Stunted equals one if the child's height-for-age is less than 2 SD below the mean. Child Development and Cognition indexes are internally standardized by sex and age using the comparison group's mean and standard deviation. Child Development Index is the average of the standardized language, fine motor, and gross motor Denver scores. Cognition Index is the average of standardized MMSE, memory, digit spans, Ravens, and matching test scores.

TABLE 5 — FIRST GENERATION FEMALES: ITT EFFECTS ON FERTILITY, CHILD MORTALITY, HUMAN CAPITAL AND EMPOWERMENT BY SEX OF THE FIRSTBORN CHILD

	All Females						Females with ≥1 Live Birth						
	Fertility					Child Mortality	Fertility			Human Capital			Empowerment
	Married (=1)	Age at Menarche (years)	Any Live Birth (=1)	Number of Live Births (=1)	Lost Pregnancy Before Firstborn (=1)	Firstborn Later Died	Male Firstborn (=1)	Age at First Birth (years)	Has Secondborn (=1)	Height (cm)	Cognition Index (z-score)	Grades Attained (years)	(0-1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: All First Generation Females Born 1982-88													
Treat	0.001	-0.055	0.015	-0.039	-0.019	-0.014	-0.006	-0.065	-0.019	1.085	-0.012	-0.071	0.002
	(0.017)	(0.072)	(0.025)	(0.065)	(0.020)	(0.012)	(0.044)	(0.245)	(0.038)	(0.412)	(0.052)	(0.217)	(0.011)
Mean	0.933	13.462	0.838	1.516	0.083	0.045	0.505	21.535	0.559	151.397	0.506	7.276	0.366
Obs	817	805	811	811	811	811	677	677	677	805	804	807	808
Panel B: First Generation Mothers Born 1982-88 With Male Firstborn													
Treat		-0.080		-0.010		-0.012		-0.172	0.041	1.216	-0.036	-0.229	0.011
		(0.101)		(0.078)		(0.026)		(0.343)	(0.056)	(0.602)	(0.093)	(0.304)	(0.017)
Mean		13.428		1.766		0.080		21.612	0.548	151.079	0.432	6.941	0.358
Obs		340		341		341		341	341	340	340	340	340
Panel C: First Generation Mothers Born 1982-88 With Female Firstborn													
Treat		-0.013		-0.114		-0.007		-0.041	-0.059	1.161	0.049	0.240	-0.001
		(0.126)		(0.108)		(0.018)		(0.391)	(0.064)	(0.600)	(0.079)	(0.380)	(0.016)
Mean		13.429		1.853		0.027		21.457	0.571	151.725	0.492	6.973	0.369
Obs		334		336		336		336	336	333	333	335	336
P(Male=Female)		0.673		0.452		0.887		0.797	0.254	0.942	0.501	0.359	0.599

Notes: This table shows single difference ITT effects on marriage, fertility and child mortality outcomes of first generation females in the 1982-88 cohort. Standard errors are clustered at the pre-program village level. Means are for the comparison group. All regressions include individual characteristics and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Individual characteristics include year of birth fixed effects and religion. Pre-intervention characteristics include all 1982-88 cohort individual and household characteristics listed in Table 1.

ALL APPENDICES ARE FOR ONLINE PUBLICATION

Appendix A Appendix Figures and Tables

TABLE A1 — FIRST GENERATION: ITT EFFECTS ON COGNITION INDEX AND ITS COMPONENTS

	Pooled				Males				Females			
	Index (z-score)	MMSE (z-score)	Digit Spans (z-score)	Ravens (z-score)	Index (z-score)	MMSE (z-score)	Digit Spans (z-score)	Ravens (z-score)	Index (z-score)	MMSE (z-score)	Digit Spans (z-score)	Ravens (z-score)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treat × $\mathbb{1}$(Born 1982-88)	-0.011	-0.033	0.003	-0.006	0.061	0.068	0.027	0.067	-0.073	-0.130	-0.019	-0.062
	(0.049)	(0.065)	(0.070)	(0.064)	(0.072)	(0.110)	(0.094)	(0.104)	(0.066)	(0.074)	(0.095)	(0.095)
Treat	0.018	0.026	0.020	0.009	-0.026	-0.015	-0.003	-0.045	0.063	0.074	0.054	0.053
	(0.043)	(0.059)	(0.056)	(0.048)	(0.057)	(0.090)	(0.073)	(0.069)	(0.046)	(0.062)	(0.062)	(0.056)
Obs	4039	4034	4036	4014	1723	1722	1721	1714	2316	2312	2315	2300

Notes: This table shows double difference ITT effects on the components of the Cognition Index of the first generation. Standard errors are clustered at the treatment village level. All regressions include individual characteristics and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Individual characteristics include year of birth fixed effects, age cohort fixed effects and controls for religion. Pre-intervention characteristics include all individual and household characteristics in Table 1. Cognition sample sizes are smaller for men because no cognitive tests, with the exception of digit span forward, were administered in the phone survey. All females in the sample were interviewed in person.

TABLE A2 — ITT EFFECTS ON HUMAN CAPITAL FOR COHORT BORN 1977-81

			Cognition Index & Components				
	Height (cm)	Short Stature (=1)	Index (z-score)	MMSE (z-score)	Digit Spans (z-score)	Ravens (z-score)	Grades Attained
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Pooled							
Treat × $\mathbb{1}$(Born 1977-81)	-0.407	-0.028	-0.036	-0.095	-0.095	0.079	0.009
	(0.480)	(0.025)	(0.072)	(0.087)	(0.086)	(0.098)	(0.264)
Treat	-0.007	0.021	0.025	0.030	0.030	0.017	0.043
	(0.280)	(0.018)	(0.043)	(0.059)	(0.056)	(0.049)	(0.159)
1977-81 Mean	155.386	0.151					3.408
Obs	3442	3442	3332	3328	3329	3306	3566
Panel B: Males							
Treat × $\mathbb{1}$(Born 1977-81)	-0.825	0.001	-0.073	-0.121	-0.187	0.076	-0.519
	(0.697)	(0.029)	(0.105)	(0.144)	(0.135)	(0.137)	(0.483)
Treat	-0.006	0.021	-0.012	0.000	0.014	-0.035	-0.058
	(0.415)	(0.022)	(0.058)	(0.090)	(0.073)	(0.070)	(0.265)
1977-81 Mean	162.144	0.105					4.398
Obs	1521	1521	1415	1414	1413	1405	1577
Panel C: Females							
Treat × $\mathbb{1}$(Born 1977-81)	-0.013	-0.058	0.009	-0.076	0.002	0.111	0.487
	(0.691)	(0.043)	(0.089)	(0.101)	(0.107)	(0.135)	(0.351)
Treat	-0.084	0.019	0.067	0.073	0.059	0.060	0.122
	(0.347)	(0.025)	(0.047)	(0.062)	(0.063)	(0.056)	(0.167)
1977-81 Mean	150.049	0.187					2.625
Obs	1921	1921	1917	1914	1916	1901	1989

Notes: This table shows double difference ITT effects on the stock of human capital of first generation 1977-81 cohort using the 1947-69 cohort as the comparison group. Standard errors are clustered at the pre-program village level and reported in parentheses. 1977-81 means are for the comparison group. All regressions include individual and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Individual characteristics include year of birth fixed effects, age cohort fixed effects, and religion. Pre-intervention characteristics include all characteristics in Table 1. Short stature is defined as height less than 155cm for males and 145cm for females. Cognition sample sizes are smaller for men because no cognitive tests, with the exception of digit span forward, were collected in the phone survey. All females in the sample were interviewed in person. Math sample sizes are smaller because the math test was not administered to anyone over age 50.

TABLE A3 — FIRST GENERATION: SENSITIVITY OF MALE HEIGHT TO SURVEY METHOD

	Base Results (1)	Self-Report Control (2)	Drop Self-Report Sample (3)
Treat × $\mathbb{1}$(Born 1982-88)	1.047	1.030	1.251
	(0.629)	(0.627)	(0.662)
Treat	-0.022	-0.015	0.022
	(0.421)	(0.420)	(0.425)
1982-88 Mean	163.40	163.40	163.07
Obs	1893	1893	1729

Notes: This table shows the double difference ITT estimates for male height are robust to self-reported measurements. Column 1 replicates the male height results of Table 2. Column 2 controls for whether the respondent self-reported their height in the phone survey. Column 3 limits the sample to those whose height was measured by an enumerator directly. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A4 — SECOND GENERATION: BALANCE OF 1982-88 COHORT MOTHERS' PRE-PROGRAM CHARACTERISTICS

	Treatment Area		Comparison Area		Difference		
	Mean	SD	Mean	SD	Mean	T-stat	Mean/SD
Mother's Characteristics							
Birth Year	1984	1.97	1984	1.69	-0.16	-1.10	-0.06
Islamic (=1)	0.85	0.59	0.95	0.27	-0.11	-2.79	-0.16
Household Characteristics							
Bari size	9.04	7.13	8.08	6.88	0.97	1.73	0.10
Family size	6.73	2.56	6.36	2.53	0.38	1.87	0.11
Asset Index (max=6)	2.83	1.62	2.77	1.68	0.06	0.47	0.03
Wall tin or tinmix (=1)	0.27	0.47	0.30	0.43	-0.04	-1.02	-0.06
Tin roof (=1)	0.81	0.45	0.82	0.46	-0.01	-0.41	-0.02
Number of rooms per capita	0.22	0.10	0.23	0.11	-0.01	-0.74	-0.04
Drinking water, tubewell (=1)	0.25	0.52	0.14	0.53	0.11	2.72	0.15
Drinking water, tank (=1)	0.44	0.84	0.33	0.90	0.11	1.59	0.09
HH age	48.42	15.30	45.32	15.58	3.11	2.53	0.14
HH <2 years of education	0.64	0.50	0.61	0.55	0.03	0.70	0.04
HH works in agriculture (=1)	0.58	0.44	0.54	0.57	0.03	0.77	0.04
HH works in fishing (=1)	0.07	0.35	0.10	0.33	-0.03	-0.94	-0.05
HH spouse's age	37.31	13.21	35.23	13.24	2.08	1.98	0.11
HH spouse <2 years of education	0.87	0.34	0.88	0.33	0.00	-0.14	-0.01
1982 Land size	9.64	13.83	10.11	15.44	-0.47	-0.40	-0.02
Obs	293		343				

Notes: The sample includes females in the 1982-88 cohort who have a firstborn child in the second generation analytic sample. Unless otherwise noted, household characteristics come from the 1974 census. Standard deviations (SD) are clustered at the treatment village level. Household head and spouse age are reported, but these variables are likely affected by the family planning program increasing birth intervals and decreasing family size in the treatment area. Observations are weighted to correct for attrition between birth and the 2012 MHSS2 survey.

TABLE A5 — SECOND GENERATION: ITT EFFECTS ON INDEX COMPONENTS

	Child Development Index (Ages 0-6)				Cognition Index (Ages 7-14)						
	Index	Language	Fine Motor	Gross Motor	Index	MMSE	Memory	Digits Forwards	Digits Backwards	Ravens	Matching
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Pooled											
Mother's Treat	-0.036	0.042	-0.005	-0.078	0.004	-0.138	0.118	0.000	0.066	0.010	-0.145
	(0.108)	(0.119)	(0.143)	(0.127)	(0.098)	(0.148)	(0.138)	(0.141)	(0.143)	(0.155)	(0.151)
Obs	334	353	348	342	262	266	263	266	266	265	266
Panel B: Males											
Mother's Treat	-0.106	0.096	-0.085	-0.241	-0.198	-0.426	0.053	-0.132	-0.251	-0.143	-0.506
	(0.145)	(0.171)	(0.200)	(0.151)	(0.144)	(0.279)	(0.258)	(0.185)	(0.176)	(0.214)	(0.206)
Obs	159	167	167	164	135	137	135	137	137	137	137
Panel C: Females											
Mother's Treat	0.048	0.023	0.085	0.027	0.268	0.363	0.239	-0.083	0.353	0.313	0.307
	(0.136)	(0.152)	(0.177)	(0.186)	(0.145)	(0.201)	(0.200)	(0.193)	(0.213)	(0.284)	(0.282)
Obs	175	186	181	178	127	129	128	129	129	128	129

Notes: This table shows single difference ITT effects on components of the Child Development and Cognition indexes of second generation children of 1982-88 cohort females using their mother's treatment assignment. Standard errors in parentheses are clustered at the mother's pre-program village level. All variables are internally standardized by sex and age (6 month bins for ages 0-6, 12 months for ages 7-14) using the comparison group's mean and standard deviation. Indexes are the average of the standardized components in subsequent columns. All models include six-month age fixed effects, mother's age fixed effects, mother's religion and mother's pre-intervention characteristics listed in Table A4.

TABLE A6 — FIRST GENERATION: ITT EFFECTS ON HUMAN CAPITAL OF ADULTS AGED 24-30 IN 1996

	Pooled			Males			Females		
	Height (cm)	MMSE (z-score)	Grades Attained (years)	Height (cm)	MMSE (z-score)	Grades Attained (years)	Height (cm)	MMSE (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treat $\times \mathbb{1}$(Age 24-30 in 1996)	-0.589	-0.107	0.206	-0.999	0.074	-0.001	-0.225	-0.254	0.080
	(0.707)	(0.099)	(0.462)	(1.254)	(0.159)	(0.760)	(0.983)	(0.143)	(0.353)
Treat	0.298	0.021	-0.024	0.864	-0.160	-0.218	-0.263	0.175	0.239
	(0.635)	(0.086)	(0.340)	(0.916)	(0.122)	(0.572)	(0.788)	(0.128)	(0.257)
24-30 Mean	154.616		2.830	161.593		3.869	150.117		2.122
Obs	1219	1229	1351	465	448	518	754	781	833

Notes: This table tests for differential pre-trends in the treatment area by estimating double difference ITT effects on the height and internally standardized MMSE of people aged 24-30 in 1996 using people aged 44-65 in 1996 as a comparison group. Standard errors are clustered at the pre-program village level and reported in parentheses. Age 24-30 means are for the comparison group. All regressions include individual and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 1996 MHSS1 survey from which the measures of human capital were gathered. Individual characteristics include year of birth fixed effects, age cohort fixed effects, and religion. Pre-intervention characteristics include all characteristics in Table 1.

TABLE A7 — FIRST GENERATION: ROBUSTNESS CHECKS

	Pooled				Males				Females			
	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Base Results												
Treat × $\mathbb{1}$(Born 1982-88)	0.977	-0.041	-0.010	0.419	1.047	-0.037	0.061	0.960	1.005	-0.042	-0.072	-0.098
	(0.346)	(0.022)	(0.049)	(0.206)	(0.629)	(0.029)	(0.072)	(0.377)	(0.464)	(0.034)	(0.066)	(0.239)
Treat	0.022	0.021	0.017	0.029	-0.022	0.023	-0.026	-0.049	-0.025	0.019	0.062	0.084
	(0.283)	(0.018)	(0.042)	(0.159)	(0.421)	(0.022)	(0.057)	(0.262)	(0.343)	(0.025)	(0.046)	(0.164)
Obs	4211	4211	4037	4340	1893	1893	1723	1952	2318	2318	2314	2388
Panel B: North Comparison Area												
Treat × $\mathbb{1}$(Born 1982-88)	0.841	-0.043	0.014	0.597	1.020	-0.031	0.108	1.237	0.729	-0.050	-0.096	-0.012
	(0.405)	(0.026)	(0.064)	(0.302)	(0.743)	(0.035)	(0.084)	(0.435)	(0.556)	(0.040)	(0.086)	(0.311)
Treat	0.224	0.012	0.010	-0.211	0.013	0.010	-0.039	-0.459	0.308	0.015	0.067	0.002
	(0.358)	(0.023)	(0.050)	(0.186)	(0.523)	(0.028)	(0.067)	(0.291)	(0.422)	(0.031)	(0.053)	(0.169)
Obs	3109	3109	2987	3208	1396	1396	1278	1440	1713	1713	1709	1768
Panel C: West Comparison Area												
Treat × $\mathbb{1}$(Born 1982-88)	0.993	-0.034	-0.029	0.163	0.890	-0.040	0.043	0.534	1.174	-0.027	-0.061	-0.164
	(0.415)	(0.027)	(0.054)	(0.208)	(0.744)	(0.031)	(0.094)	(0.474)	(0.575)	(0.042)	(0.076)	(0.275)
Treat	-0.157	0.034	0.019	0.333	-0.056	0.041	-0.035	0.434	-0.305	0.024	0.062	0.209
	(0.307)	(0.020)	(0.048)	(0.204)	(0.457)	(0.023)	(0.068)	(0.318)	(0.394)	(0.031)	(0.057)	(0.248)
Obs	3073	3073	2947	3174	1381	1381	1258	1430	1692	1692	1689	1744
Panel D: Add Extended Controls												
Treat × $\mathbb{1}$(Born 1982-88)	0.940	-0.037	0.001	0.424	1.038	-0.037	0.086	1.043	0.968	-0.038	-0.073	-0.096
	(0.340)	(0.022)	(0.051)	(0.204)	(0.630)	(0.030)	(0.073)	(0.377)	(0.455)	(0.034)	(0.068)	(0.238)
Treat	-0.040	0.019	-0.005	0.005	-0.075	0.025	-0.053	-0.021	-0.093	0.014	0.050	0.032
	(0.267)	(0.019)	(0.042)	(0.157)	(0.409)	(0.023)	(0.056)	(0.260)	(0.342)	(0.027)	(0.047)	(0.167)
Obs	4208	4208	4034	4337	1891	1891	1721	1950	2317	2317	2313	2387
	⋮											

Notes: This table shows double difference ITT estimates for the first generation are robust to a variety of potential threats. Panel A replicates the results of Table 2. Panels B and C restrict the sample of comparison units to those living in the North and West comparison areas, respectively (see Figure 1). Panel D adds controls for arsenic exposure above 100ppb, living in a village that was fully or partially eroded, access to the BRAC microcredit experiment, access to primary school and village access to healthcare [family welfare clinic, family welfare assistant, midwife, and alopathic] as of the individual's year of birth. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A7 — FIRST GENERATION: ROBUSTNESS CHECKS (CONT.)

	Pooled				Males				Females			
	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	⋮											
Panel E: Unbalanced Controls												
Treat × $\mathbb{1}$(Born 1982-88)	1.091	-0.032	-0.026	0.257	0.840	-0.025	0.031	0.760	0.886	-0.035	-0.067	-0.221
	(0.448)	(0.021)	(0.047)	(0.221)	(0.578)	(0.028)	(0.073)	(0.375)	(0.449)	(0.033)	(0.061)	(0.246)
Treat	-0.083	0.014	0.026	0.080	0.170	0.014	-0.006	0.020	0.049	0.009	0.060	0.128
	(0.360)	(0.019)	(0.041)	(0.163)	(0.403)	(0.021)	(0.057)	(0.252)	(0.324)	(0.026)	(0.043)	(0.174)
Obs	4211	4211	4037	4340	1893	1893	1723	1952	2318	2318	2314	2388
Panel F: Only Muslims												
Treat × $\mathbb{1}$(Born 1982-88)	0.902	-0.035	-0.019	0.447	1.058	-0.049	0.026	1.095	0.837	-0.019	-0.059	-0.181
	(0.393)	(0.024)	(0.054)	(0.224)	(0.687)	(0.030)	(0.077)	(0.409)	(0.529)	(0.036)	(0.071)	(0.264)
Treat	0.147	0.015	0.022	0.022	-0.008	0.030	-0.028	-0.148	0.196	0.002	0.069	0.148
	(0.305)	(0.019)	(0.044)	(0.170)	(0.466)	(0.024)	(0.059)	(0.290)	(0.355)	(0.026)	(0.050)	(0.167)
Obs	3781	3781	3619	3896	1695	1695	1537	1748	2086	2086	2082	2148
Panel G: Village Fixed Effects												
Treat × $\mathbb{1}$(Born 1982-88)	1.023	-0.047	-0.013	0.412	0.955	-0.047	-0.010	0.907	1.035	-0.032	-0.025	-0.075
	(0.341)	(0.023)	(0.049)	(0.206)	(0.673)	(0.032)	(0.072)	(0.386)	(0.487)	(0.035)	(0.070)	(0.242)
Obs	4211	4211	4052	4340	1893	1893	1729	1952	2318	2318	2323	2388
Panel H: Unweighted												
Treat × $\mathbb{1}$(Born 1982-88)	0.901	-0.033	0.008	0.366	0.758	-0.030	0.062	0.864	1.060	-0.036	-0.031	-0.116
	(0.335)	(0.021)	(0.048)	(0.199)	(0.622)	(0.028)	(0.070)	(0.359)	(0.448)	(0.031)	(0.063)	(0.221)
Treat	0.086	0.012	0.002	0.090	0.209	0.019	-0.031	0.020	-0.050	0.008	0.033	0.144
	(0.264)	(0.016)	(0.039)	(0.141)	(0.411)	(0.022)	(0.051)	(0.234)	(0.316)	(0.022)	(0.043)	(0.140)
Obs	4211	4211	4037	4340	1893	1893	1723	1952	2318	2318	2314	2388

Notes: This table is continued from the previous. Panel E includes year of birth fixed effects, cohort fixed effects and controls for religion but limits the pre-intervention characteristics to only those that were imbalanced (family size, tubewell access, household head individual and spouse characteristics [ages in 1974 and an indicators for less than two years of education]). Panel F estimates the original models using the subsample of Muslim respondents. Panel G includes fixed effects for pre-program villages which are collinear with *Treat*. Panel H does not reweight for attrition between birth and the 2012 MHSS2 survey. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A8 — FIRST GENERATION: KLING-LEIBMAN ATTRITION BOUNDING

	Height (cm)					Cognition Index (z-score)					Grades Attained (years)				
	Base	0.10 SD		0.25 SD		Base	0.10 SD		0.25 SD		Base	0.10 SD		0.25 SD	
		+	-	+	-		+	-	+	-		+	-	+	-
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: Pooled															
Treat $\times \mathbb{1}$(Born 1982-88)	0.968	0.820	0.913	0.750	0.982	-0.016	-0.010	0.001	-0.017	0.008	0.419	0.410	0.420	0.402	0.428
	(0.342)	(0.310)	(0.309)	(0.312)	(0.310)	(0.048)	(0.044)	(0.044)	(0.045)	(0.044)	(0.206)	(0.202)	(0.200)	(0.204)	(0.199)
Mean	155.766	155.861	156.003	155.755	156.109						4.258	4.236	4.312	4.179	4.368
Obs	4213	4764	4764	4764	4764	4048	4584	4584	4584	4584	4340	4752	4752	4752	4752
Panel B: Males															
Treat $\times \mathbb{1}$(Born 1982-88)	1.047	0.802	0.971	0.676	1.097	0.049	0.037	0.055	0.024	0.069	0.960	0.793	0.855	0.747	0.901
	(0.629)	(0.554)	(0.547)	(0.562)	(0.544)	(0.069)	(0.062)	(0.062)	(0.063)	(0.063)	(0.377)	(0.352)	(0.350)	(0.356)	(0.349)
Mean	162.404	162.315	162.492	162.183	162.625						4.985	4.934	5.035	4.859	5.110
Obs	1893	2197	2197	2197	2197	1728	2021	2021	2021	2021	1952	2188	2188	2188	2188
Panel C: Females															
Treat $\times \mathbb{1}$(Born 1982-88)	0.999	0.880	0.923	0.848	0.955	-0.072	-0.048	-0.043	-0.052	-0.039	-0.098	0.020	-0.004	0.039	-0.022
	(0.460)	(0.418)	(0.417)	(0.420)	(0.419)	(0.067)	(0.062)	(0.061)	(0.062)	(0.061)	(0.239)	(0.233)	(0.234)	(0.233)	(0.235)
Mean	150.330	150.275	150.386	150.191	150.469						3.663	3.636	3.690	3.595	3.731
Obs	2320	2567	2567	2567	2567	2320	2563	2563	2563	2563	2388	2564	2564	2564	2564

Notes: This table shows double difference ITT effects across different attrition bounding schemes. Base models replicate the effects shown in Table 2 while subsequent columns impute attritors' human capital measures with the within-sample mean ± 0.10 and 0.25 standard deviations (SD). Treatment effect estimates for the 44-65 year-old comparison group are omitted for brevity. Age 24-30 means are for the comparison group. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A9 — SECOND GENERATION: DOUBLE DIFFERENCE ITT EFFECTS ON HUMAN CAPITAL

	Pooled				Males				Females			
	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Develop-ment Index (z-score)	Cognition Index (z-score)
	Ages 0-14		Ages 0-6	Ages 7-14	Ages 0-14		Ages 0-6	Ages 7-14	Ages 0-14		Ages 0-6	Ages 7-14
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mother's Treat $\times \mathbb{1}$(Mother Born 1982-88)	0.076	-0.030	-0.020	0.136	-0.140	0.061	-0.071	0.004	0.388	-0.166	0.044	0.386
	(0.112)	(0.042)	(0.167)	(0.124)	(0.148)	(0.064)	(0.216)	(0.168)	(0.161)	(0.062)	(0.240)	(0.163)
Mother's Treat	-0.012	-0.004	-0.021	-0.150	0.026	-0.036	-0.055	-0.164	-0.069	0.039	0.020	-0.166
	(0.066)	(0.027)	(0.120)	(0.088)	(0.078)	(0.034)	(0.145)	(0.102)	(0.084)	(0.036)	(0.187)	(0.107)
1982-88 Mean	-1.262	0.258			-1.221	0.242			-1.302	0.273		
Obs	2796	2796	639	740	1433	1433	320	377	1363	1363	319	363

Notes: This table reports estimates from an ITT double-difference specification on human capital outcomes for the second generation using similarly aged children of mothers born prior to 1982 as the comparison group. Standard errors in parentheses are clustered at the pre-program village level. Means are for children of 24-30 year-old comparison group mothers. All regressions include individual characteristics and pre-intervention characteristics interacted with birth cohort and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. See Table 4 for a description of second generation human capital measures. All models include child's six-month age fixed effects, mother's year of birth fixed effects, mother's religion and mother's pre-intervention characteristics listed in Table A4.

TABLE A10 — SECOND GENERATION: ROBUSTNESS CHECKS

	Pooled				Males				Females			
	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Base Results												
Mother's Treat	0.106	-0.044	-0.036	0.004	-0.022	-0.002	-0.106	-0.198	0.335	-0.139	0.048	0.268
	(0.103)	(0.035)	(0.108)	(0.098)	(0.157)	(0.062)	(0.145)	(0.144)	(0.146)	(0.048)	(0.136)	(0.145)
Obs	629	629	334	262	311	311	159	135	318	318	175	127
Panel B: North Comparison Area												
Mother's Treat	0.145	-0.050	-0.054	0.075	0.105	-0.077	-0.078	-0.085	0.364	-0.099	-0.039	0.171
	(0.119)	(0.043)	(0.118)	(0.135)	(0.209)	(0.084)	(0.180)	(0.192)	(0.184)	(0.070)	(0.163)	(0.139)
Obs	455	455	241	192	235	235	123	102	220	220	118	90
Panel C: West Comparison Area												
Mother's Treat	0.113	-0.041	-0.044	-0.046	-0.101	0.067	-0.158	-0.270	0.436	-0.200	0.190	0.253
	(0.123)	(0.039)	(0.141)	(0.111)	(0.141)	(0.059)	(0.226)	(0.171)	(0.182)	(0.057)	(0.175)	(0.209)
Obs	466	466	240	198	222	222	107	101	244	244	133	97
Panel D: Add Extended Controls												
Mother's Treat	0.156	-0.039	-0.089	0.026	0.142	-0.045	-0.157	-0.263	0.250	-0.089	-0.080	0.315
	(0.115)	(0.040)	(0.117)	(0.111)	(0.182)	(0.073)	(0.157)	(0.155)	(0.158)	(0.054)	(0.158)	(0.150)
Obs	626	626	331	262	309	309	157	135	317	317	174	127
Panel E: Control for MNCH Eligibility												
Mother's Treat	0.026	-0.020	0.038	0.004	-0.173	0.072	0.010	-0.198	0.277	-0.131	0.031	0.268
	(0.100)	(0.039)	(0.131)	(0.098)	(0.162)	(0.069)	(0.194)	(0.144)	(0.163)	(0.058)	(0.176)	(0.145)
Obs	629	629	334	262	311	311	159	135	318	318	175	127
	⋮											

Notes: This table shows single difference ITT estimates for the second generation are robust to a variety of potential threats. Panel A replicates the results of Table 4. Panels B and C restrict the sample of comparison units to those living in the North and West comparison areas, respectively (see Figure 1). Panel D includes controls for available healthcare supply in the child's year of birth (family welfare clinic, family welfare assistant, midwife and alopathic) and distance to education (government, private and informal primary school, government secondary school). Panel E controls for mother's eligibility for MNCH based on her 2005 residence interacted with a dummy indicating if the child was born on or after 2007. Standard errors in parentheses are clustered at the pre-program village level.

TABLE A10 — SECOND GENERATION: ROBUSTNESS CHECKS (CONT.)

	Pooled				Males				Females			
	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)	Height-for-Age (WHO z-score)	Stunted (=1)	Child Development Index (z-score)	Cognition Index (z-score)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	⋮											
Panel F: Mother's Unbalanced Controls												
Mother's Treat	0.091	-0.037	-0.018	0.026	-0.025	0.006	-0.118	-0.167	0.289	-0.116	0.121	0.351
	(0.102)	(0.035)	(0.093)	(0.094)	(0.159)	(0.059)	(0.139)	(0.132)	(0.145)	(0.046)	(0.108)	(0.148)
Obs	629	629	334	262	311	311	159	135	318	318	175	127
Panel G: Muslims Only												
Mother's Treat	0.016	-0.012	-0.087	0.046	-0.217	0.081	-0.270	-0.118	0.278	-0.138	0.076	0.269
	(0.106)	(0.032)	(0.115)	(0.091)	(0.177)	(0.060)	(0.165)	(0.154)	(0.151)	(0.049)	(0.141)	(0.147)
Obs	573	573	301	243	276	276	141	121	297	297	160	122
Panel H: Unweighted												
Mother's Treat	0.103	-0.041	-0.042	-0.013	-0.029	0.002	-0.147	-0.209	0.314	-0.133	0.064	0.254
	(0.101)	(0.032)	(0.106)	(0.100)	(0.148)	(0.059)	(0.145)	(0.149)	(0.149)	(0.047)	(0.131)	(0.150)
Obs	629	629	334	262	311	311	159	135	318	318	175	127

Notes: This table is continued from the previous. Panel F includes six month age fixed effects but limits the mother's attributes and preintervention characteristics to only those that were imbalanced (age, religion, family size, tubewell access, household head's age and household head's spouse's age). Panel G limits the sample children of Muslim mothers. Panel H does not reweight for mother's attrition. Standard errors in parentheses are clustered at the pre-program village level.

TABLE A11 — SECOND GENERATION: KLING-LEIBMAN AND WORST CASE ATTRITION BOUNDING

	Height-for-Age (z-score)					Stunting (=1)			Cognition Index 0-6 (z-score)					Cognition Index 7-14 (z-score)				
	Base	0.10 SD		0.25 SD		Base	Worst case		Base	0.10 SD		0.25 SD		Base	0.10 SD		0.25 SD	
		+	-	+	-		+	-		+	-	+	-		+	-	+	-
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Panel A: Pooled																		
Mother's Treat	0.106	0.130	0.100	0.152	0.077	-0.044	0.014	-0.118	-0.027	-0.014	-0.049	0.013	-0.075	0.004	0.005	-0.023	0.027	-0.044
	(0.103)	(0.096)	(0.096)	(0.097)	(0.096)	(0.035)	(0.034)	(0.034)	(0.109)	(0.094)	(0.093)	(0.094)	(0.092)	(0.098)	(0.087)	(0.087)	(0.087)	(0.087)
Observations	629	672	672	672	672	629	672	672	331	374	374	374	374	262	305	305	305	305
Panel B: Males																		
Mother's Treat	-0.022	-0.015	-0.049	0.011	-0.074	-0.002	0.071	-0.083	-0.099	-0.092	-0.135	-0.059	-0.168	-0.198	-0.174	-0.209	-0.147	-0.235
	(0.157)	(0.146)	(0.146)	(0.146)	(0.146)	(0.062)	(0.058)	(0.060)	(0.146)	(0.127)	(0.126)	(0.128)	(0.126)	(0.144)	(0.125)	(0.124)	(0.126)	(0.124)
Observations	311	339	339	339	339	311	339	339	158	186	186	186	186	135	163	163	163	163
Panel C: Females																		
Mother's Treat	0.335	0.340	0.312	0.362	0.291	-0.139	-0.074	-0.192	0.059	0.073	0.049	0.091	0.030	0.268	0.245	0.223	0.262	0.207
	(0.146)	(0.138)	(0.139)	(0.138)	(0.139)	(0.048)	(0.047)	(0.049)	(0.137)	(0.125)	(0.125)	(0.125)	(0.125)	(0.145)	(0.131)	(0.132)	(0.131)	(0.132)
Observations	318	333	333	333	333	318	333	333	173	188	188	188	188	127	142	142	142	142

Notes: This table shows single difference ITT effects across different attrition bounding schemes. Base models replicate the effects shown in Table 4 while subsequent columns impute attritors' human capital measures with the within-sample mean ± 0.10 and 0.25 standard deviations (SD). For the binary outcome stunting, we impute the attritors with ones (+) or zeros (-). Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A12 — SECOND GENERATION: MECHANISM DECOMPOSITION

	Base Model (1)	Model with Mechanisms (2)	Difference (3)	Decomposition of Difference: Prenatal (4)	Postnatal (5)	Mother (6)	Obs
Panel A: Pooled							
Height-for-Age (WHO z-score)	0.106	0.0894	0.0184	0.0280	-0.00733	-0.00227	626
	(0.103)	(0.106)	(0.0268)	(0.0203)	(0.0149)	(0.00770)	
Stunted (=1)	-0.0438	-0.0496	0.00459	0.00166	0.00264	0.000283	626
	(0.0348)	(0.0374)	(0.00937)	(0.00830)	(0.00415)	(0.00108)	
Child Development Index (z-score)	-0.0364	-0.0453	0.0102	0.00998	0.00373	-0.00350	333
	(0.108)	(0.104)	(0.0246)	(0.0190)	(0.0121)	(0.00681)	
Cognition Index (z-score)	0.00425	-0.00894	0.0116	0.0246	-0.0136	0.000604	260
	(0.0984)	(0.107)	(0.0434)	(0.0352)	(0.0252)	(0.00663)	
Panel B: Males							
Height-for-Age (WHO z-score)	-0.0223	-0.0455	0.0308	0.0380	-0.00232	-0.00489	310
	(0.157)	(0.168)	(0.0413)	(0.0372)	(0.0126)	(0.0103)	
Stunted (=1)	-0.00215	-0.0126	0.00622	0.00605	-0.000149	0.000322	310
	(0.0619)	(0.0666)	(0.0156)	(0.0151)	(0.00363)	(0.00204)	
Child Development Index (z-score)	-0.106	-0.106	-0.000597	0.0564	-0.00228	-0.0547	159
	(0.145)	(0.149)	(0.0660)	(0.0409)	(0.0281)	(0.0426)	
Cognition Index (z-score)	-0.198	-0.180	-0.0330	-0.0109	-0.00102	-0.0210	134
	(0.144)	(0.170)	(0.0707)	(0.0749)	(0.0115)	(0.0218)	
Panel C: Females							
Height-for-Age (WHO z-score)	0.335	0.316	0.00980	0.0225	-0.0164	0.00366	316
	(0.146)	(0.147)	(0.0512)	(0.0321)	(0.0328)	(0.00937)	
Stunted (=1)	-0.139	-0.133	-0.00184	-0.00565	0.00390	-0.0000811	316
	(0.0479)	(0.0454)	(0.0147)	(0.0118)	(0.00654)	(0.00134)	
Child Development Index (z-score)	0.0482	0.0665	-0.0183	-0.0134	0.00220	-0.00708	174
	(0.136)	(0.128)	(0.0432)	(0.0264)	(0.0332)	(0.0146)	
Cognition Index (z-score)	0.268	0.239	0.0249	0.0446	0.0156	-0.0352	126
	(0.145)	(0.130)	(0.0723)	(0.0423)	(0.0421)	(0.0383)	

Notes: This table shows our preferred single-difference ITT estimates (Base Model) and their sensitivity to the inclusion of potential mechanisms (Model with Mechanisms) along with the difference between the two. Each row is a separate regression with the row label indicating the dependent variable. Columns 4-7 decompose the difference by contribution of four groups of mechanisms. The Prenatal group includes indicators for if the mother had a prenatal check-up, if the birth occurred at a skilled delivery location, and if a trained individual assisted with the delivery. The Postnatal group controls for the number of vaccinations the individual received and if they attended preschool. Mother is an index of the mother's decision making power, attitudes toward gender equivalence in social issues, attitudes towards husband violence, and women's mobility.

TABLE A13 — FIRST GENERATION: ROBUSTNESS OF DISTRIBUTIONAL ITT EFFECTS TO TERCILE DEFINITION

	Pooled				Males				Females			
	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)	Height (cm)	Short Stature (=1)	Cognition Index (z-score)	Grades Attained (years)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treat × $\mathbb{1}$(Born 1982-88) × Shortest Tercile	1.966	-0.096	0.066	1.047	1.720	-0.038	0.197	1.563	2.383	-0.154	-0.083	0.534
	(0.527)	(0.041)	(0.077)	(0.376)	(0.829)	(0.048)	(0.117)	(0.663)	(0.814)	(0.062)	(0.112)	(0.418)
Treat × $\mathbb{1}$(Born 1982-88) × Middle Tercile	0.761	-0.001	-0.039	0.101	0.602	-0.002	-0.033	0.968	1.077	0.002	0.002	-0.766
	(0.593)	(0.032)	(0.090)	(0.354)	(0.798)	(0.040)	(0.126)	(0.578)	(0.779)	(0.051)	(0.116)	(0.404)
Treat × $\mathbb{1}$(Born 1982-88) × Tallest Tercile	1.397	-0.017	-0.030	0.187	1.787	-0.022	0.126	0.646	0.870	-0.002	-0.174	-0.248
	(0.612)	(0.024)	(0.077)	(0.345)	(0.946)	(0.034)	(0.108)	(0.601)	(0.743)	(0.038)	(0.094)	(0.442)
Treat	-0.025	0.022	0.016	0.031	-0.065	0.023	-0.027	-0.048	-0.072	0.020	0.061	0.089
	(0.284)	(0.019)	(0.043)	(0.159)	(0.426)	(0.022)	(0.058)	(0.263)	(0.343)	(0.025)	(0.046)	(0.165)
Tercile Differences												
Shortest – Middle	1.206	-0.095	0.105	0.946	1.118	-0.036	0.230	0.596	1.306	-0.156	-0.085	1.300
P(Shortest = Middle)	0.075	0.027	0.354	0.068	0.268	0.499	0.165	0.459	0.192	0.021	0.563	0.029
Shortest – Tallest	0.569	-0.079	0.096	0.860	-0.068	-0.017	0.071	0.918	1.514	-0.152	0.091	0.781
P(Shortest = Tallest)	0.437	0.058	0.348	0.082	0.943	0.727	0.646	0.296	0.152	0.021	0.512	0.163
1982-1988 Tercile Means												
Shortest	154.25	0.20	0.35	6.69	160.67	0.12	0.28	6.63	147.90	0.27	0.43	6.74
Middle	157.12	0.05	0.46	7.14	163.04	0.05	0.41	6.84	151.52	0.06	0.51	7.42
Tallest	160.53	0.02	0.56	7.93	166.92	0.03	0.40	7.74	154.37	0.02	0.71	8.10
Obs	3832	3832	3695	3960	1701	1701	1568	1760	2131	2131	2127	2200

Notes: This table shows distributional double-difference ITT effects for the subsample of people who are in the same tercile regardless of whether we use mother's MHSS1 height filled in with MHSS2 to define the terciles, or vice versa. See Table 3 for definitions of terciles. P(Shortest Tercile = Tallest Tercile) is the p-value from a two-sided test that the treatment effects among the shortest and tallest tercile are equal. Tercile means are the comparison group. All regressions include individual and pre-intervention characteristics and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Standard errors are clustered at the pre-program village level and reported in parentheses.

TABLE A14 — FIRST GENERATION: ITT EFFECTS ON COMPONENTS OF EMPOWERMENT INDEX FOR 24-30 FEMALES

	Gender Equivalence (1=agree/0=disagree)					
	Practice purdah? (1)	Girls equally educated? (2)	Women allowed to visit relatives? (3)	Women can divorce? (4)	Women should support parents? (5)	Husband & wife choose kids? (6)
Treat	-0.020	-0.004	-0.023	-0.001	-0.011	-0.005
	(0.012)	(0.004)	(0.056)	(0.041)	(0.005)	(0.003)
Treat × Male Firstborn	0.011	0.007	0.027	-0.017	0.006	-0.000
	(0.008)	(0.008)	(0.054)	(0.046)	(0.007)	(0.001)
Mean	1.000	0.003	0.566	0.661	0.011	0.005
Obs	673	673	673	673	673	673
	Women's Mobility (1=yes/0=no)					
	Visit other baris? (1)	Vist other villages? (2)	Use public transit? (3)	Make small purchases? (4)	Make large purchases? (5)	
Treat	0.017	-0.007	0.018	-0.051	0.074	
	(0.025)	(0.018)	(0.019)	(0.048)	(0.035)	
Treat × Male Firstborn	-0.031	-0.032	0.017	0.030	-0.023	
	(0.031)	(0.021)	(0.017)	(0.062)	(0.037)	
Mean	0.916	0.965	0.954	0.423	0.822	
Obs	676	676	675	675	676	
	Husband justified in physically harming wife if she... (1=yes/0=no)					
	Goes out w/o telling him? (1)	Doesn't obey elders? (2)	Neglects her children? (3)	Refuses him sex? (4)		
Treat	-0.008	-0.025	-0.005	-0.015		
	(0.030)	(0.033)	(0.031)	(0.014)		
Treat × Male Firstborn	-0.010	0.009	0.007	0.020		
	(0.034)	(0.032)	(0.037)	(0.018)		
Mean	0.114	0.152	0.106	0.027		
Obs	673	673	673	673		
	Do you have the final say about... (1=yes/0=no)					
	Major purchases? (1)	Spouse's healthcare? (2)	Own healthcare? (3)	Visiting relatives? (4)		
Treat	0.018	-0.024	-0.015	0.052		
	(0.048)	(0.045)	(0.035)	(0.040)		
Treat × Male Firstborn	0.005	-0.022	0.028	-0.019		
	(0.047)	(0.045)	(0.042)	(0.046)		
Mean	0.303	0.292	0.176	0.211		
Obs	675	675	675	675		

Notes: This table shows ITT effects for the components of the women's empowerment index for the sample of 24-30 year-old first generation females with at least one live birth. *Male Firstborn* is an indicator for whether the female's firstborn child is male. All regressions include mother's individual and pre-intervention characteristics and are weighted to correct for attrition between birth and the 2012 MHSS2 survey. Standard errors are clustered at the pre-program village level and reported in parentheses.

Appendix B Data and Construction of Selected Variables

This appendix describes the data sources, attrition, and the creation of main outcome variables.

B.1 Data Sources

MHSS1/2.—The main outcomes variables used in this paper are drawn from MHSS2. MHSS2 is a large socioeconomic survey comprised of several instruments including a household survey, a village survey, facility surveys, and market price survey. The household survey contains two modules that is given to individuals: one, the individual module, collected information on individuals in a question format, and the another, the testing module, collected information on individuals from tests including the anthropometric and cognitive tests used in this survey. Educational outcomes were collected in the individual module of the household survey, as well as, by proxy in the household roster. Most of the data were collected during face-to-face interviews in the homes of the respondents, though a subset of data was collected in a phone survey of international migrants who did not return to Bangladesh during the data collection period (about 15 percent of our male sample). The MHSS2 phone survey instrument was shorter than the in-person survey instrument, but includes grades of education attained, self-report height and weight, and digit span forward. The remaining measures of cognition or other health tests, such as blood pressure, examined in this paper are not in the phone survey so sample sizes are smaller for these outcomes.

MHSS2 was conducted between 2012 and 2014 and was designed to be a panel to MHSS1 (icddr,b 1996). MHSS1 is a seven percent random subsample of household compounds (called baris) from the study area and was designed to be representative of the study area's 1996 population. In MHSS1, two households were interviewed in each bari: a primary household, selected randomly, and a secondary household, selected purposively. Within a household, individuals aged six and older were randomly selected to be personally interviewed.

The MHSS2 sample includes all individuals selected for personal interview in MHSS1 primary households creating panel data for these individuals.[32] To limit migration selection for key age groups, the MHSS2 sample included individuals born between 1972 and 1989 to a MHSS1 primary household that had migrated out of Matlab between 1977 and 1996 (referred to as pre-1996 migrants).[33] To the extent that a whole household and lineage migrated out of Matlab between the start of the program and 1996, leaving no one in that lineage available for selection into the MHSS1 sample, the MHSS2 sample could still suffer from migration selection. The annual migration rate of whole households from the study areas prior to 1996 was low, 0.66 percent, and most of the exiting households were Hindu.

With the exception of the phone survey, all tests were collected on the individuals in their home by well-trained testers. Testers were extensively trained to implement protocols in a similar fashion. Retraining took place a couple of times throughout the survey period to restandardize the testers. The testers were generally female, though there were a few male testers who implemented the tests on adult male respondents. For the most part, testers were only responsible for implementing the test module and collected information a few weeks after the household survey. There were a few enumerators who were trained on the entire household survey, including the testing module, to interview adult migrants who were living on their own.

Census Data.—Periodic censuses were collected for all individuals in the study area (treatment and comparison areas) by icddr,b. These data typically include household location, household characteristics and composition, employment, education, and assets. We obtain pre-program individual and household data on the analysis sample from the 1974 census (icddr,b 1974) and use these data to test for differences in baseline characteristics between the treatment and comparison areas. We also use the 1974 and 1982 censuses (icddr,b 1982) to link individuals to the study area before 1977 to construct an individual's intent-to-treat status (see section C below).

32. MHSS2 sample also included all panel member descendants, and their co-resident spouses. However, spouses of panel members who had migrated out of the study areas were not tracked for interview.

33. The pre-1996 migrants were identified by using the detailed demographic surveillance data.

Demographic Surveillance Site (DSS) Data.—Vital registration data provide prospective tracking of every birth, death, marriage, divorce, and in- and out- migration occurring in the study area. As such, we know when someone enters and leaves the study area. Data were collected by icddr,b and are high quality in part because they were collected so frequently: every two weeks until 1997, every month between 1998 and 2006, and every two months between 2007 and MHSS2. These data include pre-program data from 1974 onwards and are used to construct birth dates and an individual's intent-to-treat status. In addition, we use these data to construct pre-program migration network variables for each individual in the analysis sample, as well as, out-migration variables such as whether someone has ever migrated, and out-migration variables for years not covered in the MHSS2 migration history.

B.2 Attrition

Table B1 and Table B2 documents attrition rates for the first and second generations. Relative to other long-term follow-up surveys, attrition in the MHSS2 sample is low. For the first generation, including death and any other type of non-response, the attrition rate at the household level is 7 percent. Attrition rates are slightly higher for variables from the testing module at 10 percent for men and 7 percent for females for height information (Table B1). Attrition rates for men reached 24 percent for outcomes that were not collected in the phone survey. Differences in attrition rates between the treatment and comparison areas for the full sample are less than one percent and are not statistically different from each other. For the second generation, attrition in similarly low. Of the 677 females aged 24-30 with a firstborn child, 629 were surveyed and had a valid height measurement. Only 47 children were lost (6.92% attrition rate), 4 of which had invalid height measurements, 29 died before being surveyed and 14 were born to migrant females whose children were not followed. The attrition rate did not differ between treatment and comparison areas by any definition or among males and females separately.

Even though the attrition rates are low, there could be selection into attrition. Table B3 presents the means of individual and baseline household characteristics by attrition status for the first generation pooled sample. Individual and household characteristics are fairly well balanced by attrition status, though those who were lost to attrition were more likely from larger baris and less likely to own a radio, as well as, to be male and older which is not surprising given men migrate for work and older men were more likely to not be in the survey due to death. Table B3 further shows that there is limited differential attrition between the treatment and comparison areas, however for females selection into attrition was negative based on household head's education which was seven percentage points lower. Results are similar if we split them by sex and if we include the phone survey respondents in the attrition rather than the surveyed group.

The low attrition rates are a result of a carefully designed tracking protocol. Migrants were tracked all over Bangladesh, and a rapid response system was developed that allowed trackers in Matlab to connect enumerators placed in different parts of the country with respondents who had left Matlab. Intensive interviewing took place during all the Eid holidays from 2012–2014. Survey teams targeted international migrants, far away domestic migrants, and hard-to-track migrants returning to Matlab for the holiday. Finally, a phone survey was employed to collect information on a subset of questions from the main survey from male international migrants who did not return to Bangladesh during the survey period. While there is a limited set of variables available for this group in the phone survey, most educational and anthropometric outcomes, including self-report height and weight, used in this study were collected during the phone survey. However, with the exception of digit span forward, no other cognition or health measures used in the study were included in the phone survey.

TABLE B1 — FIRST GENERATION: 2012 ATTRITION RATES OF 1982-1988 COHORT

	Pooled			Males			Females		
		T-C Difference			T-C Difference			T-C Difference	
	Attrition Rate	Mean	(SE)	Attrition Rate	Mean	(SE)	Attrition Rate	Mean	(SE)
Not found or refused	5.61%	-0.007	(0.013)	5.11%	0.008	(0.018)	6.13%	-0.023	(0.016)
Not found, refused, or dead	7.22%	-0.008	(0.014)	6.96%	0.005	(0.019)	7.49%	-0.020	(0.017)
Non-missing height information	9.44%	-0.003	(0.015)	10.01%	0.011	(0.020)	8.85%	-0.017	(0.019)
Non-missing MMSE information	17.00%	-0.012	(0.017)	24.59%	-0.013	(0.029)	9.08%	-0.012	(0.019)
Non-missing height, no phone survey	16.78%	-0.006	(0.017)	24.37%	0.004	(0.028)	8.85%	-0.017	(0.019)

Notes: Sample includes 1982-1988 cohort. The Attrition Rate column displays the percent of the sample who attrited, while mean and SE show the difference and standard error of the difference in attrition rates by treatment and comparison area. The standard error of the difference is clustered at the pre-program village level. There are 919 men and 881 females across the two cohorts in the sample frame. Missing height information indicates the respondent was not able to be measured, including not found, refused, or dead, but outliers are not dropped here as they are in the final analysis.

TABLE B2 — SECOND GENERATION: 2012 ATTRITION RATES

	Pooled				Males				Females			
	Number	Attrition	T-C Difference		Number	Attrition	T-C Difference		Number	Attrition	T-C Difference	
	Attrited	Rate	Mean	(SE)	Attrited	Rate	Mean	(SE)	Attrited	Rate	Mean	(SE)
Child missing height	47	6.92%	-0.037	(0.020)	29	8.48%	-0.060	(0.030)	18	5.34%	-0.014	(0.025)
Child not found	43	6.33%	-0.032	(0.019)	28	8.19%	-0.054	(0.030)	15	4.45%	-0.010	(0.023)
Child died	29	4.27%	-0.024	(0.016)	22	6.43%	-0.034	(0.027)	7	2.08%	-0.014	(0.016)
Mother migrated	14	2.06%	-0.008	(0.011)	6	1.75%	-0.020	(0.014)	8	2.37%	0.004	(0.017)

Notes: Sample includes live firstborn children of 1982-1988 females. Number Attrited is the number of observations in the sample who attrited, while Attrition Rate displays the percent of the sample who attrited, and mean and SE show the difference and standard error of the difference in attrition rates by treatment and comparison area. The standard error of the difference is clustered at the pre-program village level. There are 342 males and 337 females in the sample frame. Missing height information indicates the respondent was not able to be measured, including not found, refused, or dead, but outliers are not dropped here as they are in the final analysis.

TABLE B3 — FIRST GENERATION: ATTRITION BALANCE OF 1982-1988 COHORT'S PRE-PROGRAM CHARACTERISTICS

	Attrited		Surveyed		Difference		
	Mean	SD	Mean	SD	Mean	T-stat	Mean/SD
Individual Characteristics							
Male (=1)	0.54	0.48	0.50	0.46	0.04	0.96	0.06
Birth year	1984	2.20	1984	1.84	-0.12	-0.68	-0.04
Islamic (=1)	0.93	0.28	0.90	0.59	0.02	1.03	0.03
Mother's height (cm)	150.27	5.66	149.36	10.29	0.91	1.63	0.08
Father's height (cm)	160.88	7.03	160.79	8.41	0.10	0.13	0.01
Mother's years of education	1.45	2.31	1.70	3.61	-0.25	-1.14	-0.06
Father's years of education	3.13	4.04	3.48	5.23	-0.35	-0.93	-0.05
Household Characteristics							
HH Bari size	9.27	7.37	8.27	10.21	1.00	2.10	0.08
HH Family size	6.84	3.07	6.65	3.43	0.18	0.76	0.04
Latrine (=1)	0.80	0.40	0.85	0.72	-0.05	-1.55	-0.06
Owns a lamp (=1)	0.57	0.53	0.61	0.71	-0.04	-1.02	-0.04
Owns a watch (=1)	0.13	0.32	0.15	0.45	-0.02	-0.89	-0.04
Owns a radio (=1)	0.04	0.18	0.07	0.30	-0.03	-2.11	-0.09
Number of cows	1.33	1.98	1.37	2.12	-0.04	-0.28	-0.01
Number of boats	0.61	0.66	0.65	0.98	-0.04	-0.81	-0.03
Wall tin or tinmix (=1)	0.26	0.40	0.29	0.59	-0.03	-0.98	-0.04
Tin roof (=1)	0.77	0.41	0.83	0.53	-0.06	-1.71	-0.08
Number of rooms per capita	0.21	0.09	0.22	0.12	-0.01	-0.75	-0.04
Drinking water, tubewell (=1)	0.22	0.51	0.22	0.88	0.00	0.07	0.00
Drinking water, tank (=1)	0.36	0.56	0.36	1.24	0.00	0.05	0.00
HH age	45.79	14.36	46.87	17.58	-1.08	-0.96	-0.05
HH <2 years of education (=1)	0.65	0.49	0.60	0.69	0.05	1.44	0.06
HH works in agriculture (=1)	0.61	0.54	0.58	0.72	0.03	0.65	0.03
HH works in fishing (=1)	0.12	0.44	0.07	0.39	0.05	1.60	0.09
HH spouse's age	36.33	10.90	36.39	15.70	-0.06	-0.07	0.00
HH spouse <2 years of education (=1)	0.89	0.31	0.87	0.46	0.02	0.83	0.03
1982 Land size	9.97	15.54	10.58	19.98	-0.62	-0.56	-0.02
Obs	170		1630				

Notes: The sample includes male and female respondents in the 1982-1988 cohort. Attrition is defined as missing height information prior to dropping outliers. Standard deviations are clustered at the village level. Attrition is also balanced for the 1947-1969 cohort (results not shown).

TABLE B4 — FIRST GENERATION: ATTRITION BALANCE OF 1982-1988 COHORT'S PRE-PROGRAM CHARACTERISTICS BY TREATMENT STATUS

	Pooled		Males		Females	
	Main Effect	Interaction	Main Effect	Interaction	Main Effect	Interaction
Individual Characteristics						
Male (=1)	0.006	0.028				
	(0.020)	(0.026)				
Birth year	-0.002	-0.003	-0.002	-0.003	-0.000	-0.003
	(0.005)	(0.008)	(0.008)	(0.011)	(0.008)	(0.010)
Islamic (=1)	0.004	0.032	0.104	-0.109	-0.077	0.140
	(0.054)	(0.064)	(0.051)	(0.078)	(0.086)	(0.091)
Household Characteristics						
HH Bari size	0.002	0.001	0.001	0.001	0.003	0.002
	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.004)
HH Family size	0.007	-0.005	0.003	0.003	0.009	-0.012
	(0.005)	(0.007)	(0.007)	(0.010)	(0.008)	(0.010)
Latrine (=1)	-0.059	0.055	-0.041	0.033	-0.066	0.065
	(0.037)	(0.050)	(0.047)	(0.067)	(0.053)	(0.064)
Owns a lamp (=1)	-0.010	0.017	-0.010	0.006	-0.005	0.011
	(0.020)	(0.035)	(0.035)	(0.060)	(0.030)	(0.036)
Owns a watch (=1)	-0.009	0.035	-0.009	0.019	0.004	0.030
	(0.025)	(0.043)	(0.032)	(0.058)	(0.036)	(0.061)
Owns a radio (=1)	-0.063	0.042	-0.005	0.030	-0.113	0.059
	(0.034)	(0.052)	(0.051)	(0.081)	(0.040)	(0.062)
Number of cows	0.002	0.000	-0.005	0.004	0.007	-0.003
	(0.007)	(0.012)	(0.009)	(0.015)	(0.010)	(0.015)
Number of boats	-0.031	0.022	-0.018	-0.021	-0.052	0.073
	(0.022)	(0.029)	(0.032)	(0.048)	(0.035)	(0.042)
Wall tin or tinmix (=1)	-0.008	0.029	-0.009	0.022	-0.002	0.047
	(0.025)	(0.034)	(0.038)	(0.054)	(0.034)	(0.049)
Tin roof (=1)	-0.027	0.005	-0.016	-0.006	-0.042	0.010
	(0.028)	(0.042)	(0.039)	(0.067)	(0.045)	(0.060)
Number of rooms per capita	-0.007	-0.086	-0.012	0.004	0.001	-0.200
	(0.123)	(0.176)	(0.158)	(0.220)	(0.157)	(0.262)
Drinking water, tubewell (=1)	-0.017	0.071	-0.027	0.091	-0.009	0.083
	(0.028)	(0.040)	(0.037)	(0.059)	(0.037)	(0.055)
Drinking water, tank (=1)	-0.017	0.074	-0.022	0.076	-0.009	0.076
	(0.022)	(0.037)	(0.026)	(0.046)	(0.034)	(0.057)
HH age	-0.002	0.002	-0.003	0.002	-0.002	0.003
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
HH <2 years of education (=1)	0.001	0.032	0.013	-0.002	-0.013	0.079
	(0.020)	(0.033)	(0.027)	(0.043)	(0.030)	(0.041)
HH works in agriculture (=1)	0.029	-0.018	0.016	0.008	0.045	-0.038
	(0.023)	(0.031)	(0.026)	(0.044)	(0.035)	(0.044)
HH works in fishing (=1)	0.122	-0.072	0.151	-0.213	0.122	0.010
	(0.063)	(0.103)	(0.094)	(0.126)	(0.073)	(0.113)
HH spouse's age	0.003	-0.003	0.002	-0.001	0.004	-0.007
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
HH spouse <2 years of education (=1)	-0.001	-0.000	-0.025	0.024	0.019	-0.026
	(0.029)	(0.041)	(0.047)	(0.064)	(0.038)	(0.058)
1982 Land size	0.001	-0.001	0.001	-0.003	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
F-stat all interactions = 0	1.13		1.10		1.83	
P-value	0.32		0.35		0.02	
Obs	1800		919		881	

Notes:

B.3 Construction of Selected Variables

B.3.1 First Generation

Height & Short Stature.—Adult respondents were measured standing using a Shorr height board. If the respondent was interviewed by phone, their height was self-reported. We dropped height values above 195cm and 182cm for males and females, respectively. We defined an individuals short stature if their height was less than 155cm for males or 145cm for females.

Cognition Index.—This index was created by taking the average of standardized MMSE, Digit Spans Forward and Backward, and Ravens scores. All three components were internally standardized using the comparison group mean and standard deviation by sex.

Metabolic Syndrome.—This index was created by taking the average of standardized underweight, overweight, and stage 1 hypertension. Respondents were considered underweight if their BMI was below 18.5 kg/m^2 and overweight if their BMI was above 23 kg/m^2, following the WHO recommendation that for Asian populations a cutoff of 23 is more appropriate than the 25 used for the US population. If a respondent's BMI was above the 99.5 percentile or below the 0.5 percentile for her age, they were dropped from the sample. Weight was measured using the SECA881 U digital scale (150kg maximum and 0.01kg increments). Three measurements of blood pressure were taken and then averaged. Stage 1 hypertension is defined as a systolic blood pressure above 130 mm Hg or a diastolic blood pressure above 85 mm Hg. For outliers, if systolic blood pressure was less than 60 or more than 250 it was set to missing. Similarly, if diastolic blood pressure was less than 40 or more than 150 it was set to missing. Blood pressure was measured using the Lifesource 767-PV automatic blood pressure device. All three components were internally standardized using the comparison group mean and standard deviation by sex.

Mothers' Height Terciles.—We calculated terciles of mothers' MHSS2 height (using their measurement from MHSS1 if their MHSS2 height was missing) across the entire sample. First generation respondents were in the lowest tercile if their mother's height was below 147.5cm, in the middle tercile if between 147.5 and 152.1cm and in the highest tercile if above 152.1cm. 17.4 percent of first generation respondents in our analytic sample had a mother of short stature (height less than 145cm).

B.3.2 Second Generation

Height-for-age and Stunting.—Second generation children were measured using the Shorr height board lying down if the child was less than two years old or shorter than 83cm and standing up if older than two or taller than 83cm. We calculated standardized height-for-age using WHO growth standards. Any values more than five standard deviations from the mean were set to missing. Stunting was defined as a height-for-age of two standard deviations or more below the mean.[34]

Child Development Index.—For children aged 0-6, we construct a Child Development Index by taking the average of standardized Denver Language, Fine Motor and Gross Motor skills. All components were internally standardized using the comparison group mean and standard deviation by sex and age in years.

34. Guideline: Assessing and Managing Children at Primary Health-Care Facilities to Prevent Overweight and Obesity in the Context of the Double Burden of Malnutrition: Updates for the Integrated Management of Childhood Illness (IMCI). Geneva: World Health Organization; 2017. Table 1, World Health Organization (WHO) classification of nutritional status of infants and children. Available from: https://www.ncbi.nlm.nih.gov/books/NBK487900/table/fm.s1.t1/

Cognition Index.—For children aged 7-14, we construct a Cognition Index by taking the average of standardized MMSE, Memory score, Digit Spans Forward and Backward, Ravens, and Matching score. All components were internally standardized using the comparison group mean and standard deviation by sex and age in years.

Vaccinations.—BCG, DPT/Penta, Polio and MMR vaccine histories for children were recorded by questioning the mother or caretaker. We calculate the number of vaccination types (out of four) the child received.

Mother's Empowerment Index.—We construct an index measuring the mother's relative empowerment from four subindicies: decision making, gender equivalence in social issues, attitudes towards husband violence, and mobility. All measures are scored or recoded so that a larger value represents more empowerment of the woman. Each subindex and its components are described below.

Decision Making.—Married or previously married women were asked who had the final say in four categories: major household purchases, decisions about spousal health and treatment, decisions about their own health and treatment, and visits to family or relatives. We coded each response with 1 if the woman made decisions independently or jointly about the matter and 0 if someone else had the final say. Finally, we created an index by averaging the four responses.

Gender Equivalence in Social Issues.—Women were asked to give an individual opinion on six claims: (1) women should practice purdah (female seclusion), (2) girls should be educated as much as boys, (3) women should be allowed to go out alone, (4) women should have the right to initiate divorce, (5) daughters should support elderly parents and (6) husband and wife should jointly agree to have children. We coded responses to (1) as 0=agree, 0.5=partially agree and 1=do not agree while responses to (2)-(6) were coded in reverse so that a larger value represents a higher opinion of gender equivalence. We then took the average of all six responses to form an index of gender equivalence in social issues.

Attitudes Toward Husband Violence.—Women were asked whether a husband is justified in hurting his wife if he is ever angered or annoyed by his wife's behavior in four situations: (1) going out without telling her him, (2) disobeying family elders, (3) neglecting their children and (4) refusing to have sex. We recoded responses so that yes=0 and no=1 to represent a more egalitarian viewpoint and averaged her four responses to create an index of attitudes toward husband violence.

Women's Mobility.—Women were asked if in the past twelve months they had gone out alone or with someone else for five activities: (1) visiting women in other baris, (2) visiting someone outside their village, (3) visiting any place on public transport, (4) visiting the local store for small purchases (oil, rice, salt, etc.) and (5) visiting a store or market to make larger purchases such as clothing. We coded each response as 1 if they went alone, 0.5 if they went out with someone and 0 if they never went out. Then we averaged responses to create an index that represents more mobility.

Appendix C Potential Confounders

We accounted for several other important and well-documented changes that occurred in Matlab over the 35-year period since program inception that could confound results.

The Bangladesh Female Secondary Education Stipend Program.—One government program that is pertinent for MCH-FP effects on education, and difficult to control for, is the Bangladesh Female Secondary Education Stipend Program. This was a national program that became available in Matlab in 1984 for females attending grades 6-10 who were unmarried, had 75 percent attendance and scored 45 percent on school exams. The program targeted individual females, not schools, and provided a stipend and covered many school costs. This program was available to all the females in the first generation sample in this paper for the entirety of their secondary schooling in both the treatment and comparison areas, providing no variation to test for heterogeneity treatment effects of MCH-FP based on differential access to this program. In addition, we do not have data on who received the stipend program, nor sufficient data to determine who may have been eligible. Shamsuddin (2015) estimates that five years of exposure to the program led to one year gain in education. As a result, any program effects on educational outcomes of the MCH-FP program on females needs to outweigh the already large effects on education from the female stipend program.

Erosion and Flood Risk.— In 1987 the government of Bangladesh completed the Meghna Dhonnogoda Irrigation Project. This project involved constructing a river embankment along the northern bank of the Meghna River where it meets the west bank of the smaller Dhonnogoda River, which runs through Matlab (see Figure 1). The villages near this project were all located in the comparison area, and the embankment had two important consequences for these villages. First, seven villages in this area lining the river were partially or fully inundated between 1984 and 1986 as part of the project. All households in these villages were displaced, with most initially relocating to adjoining villages within the comparison area. To control for potential differences in the Meghna area in general, we include a variable indicating if the person's treatment village was submerged as a result of the project.

1993 BRAC Microcredit Experiment.—In 1993, BRAC introduced an experiment that provided landless females with access to microcredit. The rollout was designed to include villages in both treatment and comparison areas, but the presence of the program could still bias our results. We include an indicator for whether the village ever participated in BRAC during its experimental period.

School Access.—There was a large expansion of education during the 1990s, including construction of primary schools. If education supply was not equal between the two areas, it could bias our results. We control for differential primary school access by including an indicator variable that is one if a primary school was present in the first generation individual's treatment village during the years that they were age-eligible. For the second generation analysis, we control for the distance to nearest government, private and informal primary school and the distance to the nearest government secondary school.

Health Facility Access.—Access to health care supply could have changed over time differentially between the two areas. To control for post-program access to health care, for the first generation, we use information on access to health care form the MHSS1 village survey (taken in 1996) . MHSS1 surveyed village leaders about health facilities used by people from their village including the date of construction. We construct indicators for the presence of a Family Welfare Center (a government clinic, FWC), a Family Welfare Assistant (a government health worker that travels to villages, FWA), a non-MBSS allopathic doctor, and a Trained Traditional Birth Attendant (a midwife) in the individual's birth year to account for differential access to pre- and post-natal care. We use the MHSS1 Village Survey rather than its MHSS2 equivalent to avoid selection from the timing of facility opening and closings. For the second generation, we use a similar questionnaire in MHSS2 to construct dummies for the same facility access in the year a second generation child was born.

Arsenic Exposure.—To account for differential access in tubewell water, which contains arsenic, between the treatment and comparison areas, we control for arsenic exposure in tube well water. We use 2003 measures of arsenic collected by icddr,b . We use the 2003 measure rather than measures collected at later times, including with MHSS2, because it was unknown that there were arsenic in the wells prior to

2003, so this represents a period before families engaged in well switching which could be correlated with treatment status, and since it was measured at a time closer to when the sample of interest were young children. Wells are linked to MHSS1 households using the ID of the person who takes care of the well. For household who do not take care of a well, we take the average arsenic level in the three closest wells. For households that reported not using a tubewell in MHSS1 (which was prior to knowing about arsenic in the well), the value of arsenic is set to zero. The included control is an indicator of an arsenic level above 100 parts per billion (ppb, micrograms per liter). Note a majority of the children in the sample were born after the wells were established, so the age fixed-effects control for the length of time exposed to the well water.

2001-2003 Maternal and Infant Nutritional Intervention (MINIMat).—Between November 2001 and October 2003 pregnant females in the study area were randomized into a nutritional intervention program. The randomization was done at the individual level and was conducted in both the MCH-FP treatment and comparison areas so should be balanced across areas and not drive results (El Arifeen et al., 2018). The children impacted by the MINIMat intervention were largely born before the second generation cohort that is considered in this study so should not impact results. We do drop any children born prior to 2003, and results remain the same.

2007 Maternal, Neonatal and Child Health Project (MNCH).— Beginning in 2007 this program aimed to increase facility-based childbirth and neonatal health. Pregnant women enrolled in the program were visited by a community health worker at 12-14 weeks and again between weeks 32 - 34. Women were encouraged to deliver at a facility with skilled obstetric care rather than at home. Postnatal visits were made by a community health worker on days 3, 7, and 28 to provide counseling on newborn and maternal well-being. See Rahman et al. (2011) for a detailed description of the program. The targeting of this program overlapped with the treatment area of the MCH-FP program, however children born before 2007 were born after the program so did not benefit. As such, the program would have mainly affected children less than age 6. For robustness, we control for mother's exposure to the MNCH program based on her residence in 2005 interacted with a dummy indicating whether the child was born on or after 2007. Results are similar when including this control and available upon request.

Appendix D Weights

The main results are weighted for attrition between birth and MHSS2 using inverse propensity weights. The analysis sample includes respondents from MHSS1 and individuals from MHSS1 households that had migrated out of the DSS area prior to the survey conducted in 1996. The main reason for non-response are migration in early adulthood and death primarily during infancy. Weights are constructed in two steps. First, we estimate weights to account for selection into the MHSS1 sample frame between birth and MHSS1, which is mainly a result of mortality. Second, we estimate weights to account for attrition of MHSS1 respondents in the MHSS2 survey. We estimate these two probabilities separately and then multiple them to obtain a weight to account for attrition between birth and MHSS2.

The weight the account for attrition between birth and MHSS1 estimate the conditional probability that an individual born in the study site was present to be surveyed in MHSS1 using demographic surveillance data. To do this, we assign treatment status to the universe of individuals born in the study site between 1977 and 1988. Separately by cohort and sex, we use a probit model to predict the probability an individual is present in the study site on January 1, 1996 using the set of baseline household and household head characteristics (which includes pre-program migration networks for the household compound), their interactions with the treatment variable, month of birth and year of birth fixed effects, and indicators for whether an individual was from a village that experienced erosion or was exposed to the Meghna Dhonnogoda Irrigation Project.

The weight to account for attrition between MHSS1 and MHSS2 is constructed in a similar manner. We estimate the probability of non-attrition between the two survey waves for each cohort-sex group using a probit model and the same set of covariates. The resulting attrition weight is the inverse of the product of the two probabilities.