Thirty-Five Years Later: 
Evaluating Effects of a Quasi-Random Child Health and Family Planning 
Program on Human Capital Development in Bangladesh 

Tania Barham 
Department of Economics and Institute of Behavioral Science 
University of Colorado, Boulder 

***Preliminary Draft not for Citation **** 

June 17, 2015 

Research discussed in this report has been funded by the International Initiative for Impact Evaluation, Inc. (3ie) through the Global Development Network (GDN) (OW3.1060); the National Institute on Aging (R01 AG033713), and the Population Reference Bureau (35 Years Later: Effects of the Matlab Maternal and Child Health and Family Planning Program on Women’s Economic Empowerment). The views expressed in this report are not necessarily those of 3ie or its members. We are grateful to various members of icddr,b (including but not limited to Abudur Razzaque, Jena Hamadani, Sajal Saha) for their myriad of help with the data including providing demographic surveillance data for historic information and drawing samples, help with design of survey and monitoring of the survey team, and excellent help and attention to tracking migrants. We are indebted to the survey team, Mitra and Associates, for excellent data collection and persistence in tracking migrants. We thank Elisabeth Root, Chris Jochem, Gisella Kagy, Noboku Mizoguchi, Patrick Turner, Sveta Milusheva, Emily Steiner for help with preparing survey instruments, field protocols, training, monitoring of field work, data preparation, and numerous contributions to this research project.
ABSTRACT:

Improving the health and nutrition of young children is important not only for immediate well-being, but also because it is believed to reduce poverty in the long-run through improved human capital. Many programs such as Head Start and Conditional Cash Transfer (CCT) programs rely on this postulated link. Little, however, is known about the long-term effects of programs targeted to improve health and nutrition in early childhood on human capital in adulthood. A growing literature suggests that large negative health or nutrition shocks early in life, lead to worse outcomes later in life, but there is little long-run evidence on the effects of interventions designed to improve the health and nutrition of young children. Understanding the longer-run effects of early childhood interventions is important as there is growing interest in investing resources in disadvantaged children at an early stage in life, for example through the spread of poverty reduction programs like CCTs. It is crucial to investigate these questions since evidence on early childhood nutrition and health interventions is mixed as to whether their benefits continue (Pollitt et al. 1993) or fade out (Garces et al. 2002).

This project examines the effects of the Matlab Maternal and Child Health and Family Planning (MCH-FP) program that started in 1977, 35 years later. Treatment and comparison areas were built into the design of the program. The program was phased-in over time starting with family planning and maternal health. Measles vaccination began in 1982 and other child health interventions were included in 1986 (e.g. other vaccinations and vitamin A supplementation). Similar interventions were introduced in the comparison area in the late 1980s providing approximately a 10-year experiment period to evaluate the program.

The study takes advantage of the quasi-experimental design and the phasing-out of the program over time to examine the effect of the program on those who were born during the experimental period on their cognitive functioning and height adulthood. To limit selection bias that is common in panel studies, the study design paid special attention to reducing panel attrition by extensive tracking of migrants.

Previous research shows the MCH-FP program led to important improvement in human capital in early and late childhood (ages 8-14). This study examines if these effects persist when these same people are aged 22-29. We find that while the effect on height persists the effects on the MMSE fade out and there are no statistically different effects between the treatment and comparison areas for the other measures of cognition that were measured only at age 22-29. The difference in results between height and MMSE highlight that physical growth and cognitive development are affected differently by one’s environment. While family planning and childhood vaccination are clearly important interventions for improving the health of children, earlier investment in childhood health and nutrition may be needed to lead to sustained improvements in cognitive functioning.
1. INTRODUCTION

Improving health and nutrition of mothers and young children is important not only for their immediate well-being, but also because it is believed to reduce poverty in the long-run through improved human capital and labor market opportunities for the child (Strauss & Thomas 2008, Heckman 2006). Economic theories of human capital development rely on this postulated link and are part of the rationale for important programs in the US (e.g. Head Start) and in developing countries (e.g. Conditional Cash Transfer (CCT) programs). Despite the global spread of programs such as CCTs, there is limited longer-run evidence of effects of early childhood health, nutrition, or family planning programs, for example on cognitive functioning, health status, and labor market outcomes in adulthood (Hoddinott et al. 2008, Maluccio et al. 2009, Barham 2012). It is important to investigate longer-term effects since evidence from some early childhood nutrition and health interventions is mixed as to whether their benefits continue (Pollitt et al. 1993) or fade out (Garces et al. 2002). The issue of fade-out is particularly pertinent in developing countries, where many competing health and environmental risks and a greater frequency of shocks are coupled with limited ability to smooth consumption.

Programs intended to improve health and nutrition of young children are common in the developing world. Few, however, have been introduced in designs that permit full assessment of their impacts. Even when evaluation is built in, long-term follow-up is rare, so most existing evaluations are limited to the short or medium term. Specifically, opportunities to study longer-term consequences of disease prevention and family planning interventions are extremely unusual. Matlab, Bangladesh, is one of the few settings in the world that combines quasi-randomization of interventions, availability of pre-intervention data, and long duration of follow-up (over 30 years).

In 1977, icddr,b (formerly the International Center for Diarrhoeal Disease Research, Bangladesh) introduced the Matlab Maternal and Child Health and Family Planning (MCH-FP) program in half of a study area, the treatment area, leaving the other half as a comparison area. Treatment and comparison areas were chosen to be economically and socially similar. Mortality and contraceptive rates and most household, household head, and individual characteristics were similar prior to the program. Program interventions were phased-in starting with family planning and tetanus toxoid vaccines for pregnant women. Intensive child health interventions began in 1982 with measles vaccination. Starting in 1985, the project expanded to include child health interventions such as vaccinations against tetanus, pertussis, polio, and tuberculosis. Preventing these diseases improves health and nutritional status of these children. It also reduces chances of cognitive impairment from these diseases either directly or indirectly, for example through nutrition. Similar interventions were introduced in the comparison area in the late 1980s, thus providing approximately a 10-year experiment period to help evaluate the program.

This paper takes advantage of the MCH-FP program quasi-random design and phasing-in of interventions to address shortcomings in the literature. Our overall program is intended to investigate if a well-implemented program that led to improvements in health status in the short-term led to improved health, wealth, and life chances for those who born during the experimental period and in particular on the cognitive functioning and height.

Given the long-term nature of this study and the relationship between better health in childhood and cognitive development, education, health, and migration, it was paramount to our study to minimize attrition. Consequently, great efforts were made to track migrants throughout Bangladesh, to interview international migrants when they returned for holidays, and to
implement a short phone survey for international migrants that we could not interview in person in Bangladesh. This extensive tracking was a key component of the design of this study and was highly successful. 35 years after the start of the program we were able to interview more than 90 percent of men age were born during the experimental period and are the group with the highest migration rates in the study.\footnote{Such tracking was the primary objective and use of the funding provided by 3ie, with the main data collection supported by NIA.} Response rates for females and other age groups were even higher. These rates of attrition are remarkably low compared to other impact evaluation studies covering similar populations, as well as to many longitudinal studies that cover much shorter periods.

2. BACKGROUND

2.1 Past Evidence on the MCH-FP Program in Matlab, Bangladesh

Previous research on the short-run effects of the MCH-FP program used the demographic surveillance data and program receipt data to show that the MCH-FP program was effective in reducing child mortality and fertility (Phillips et al. 1982, 1984; Koenig et al. 1990, 1991). Studies using the 1996 MHSS\textsuperscript{1} show improvements in cognitive functioning, anthropometrics, and education in late childhood and adolescence for children exposed early in life to the MCH-FP program. Joshi and Schultz (2007) found an increase in schooling for boys aged 9-14, but no effect for girls. Chaudhuri (2005) reported that girls younger than 14 experienced improved weight-for-age and boys were significantly less stunted. Barham (2012) found that the MCH-FP program led to a 0.39 standard deviation increase in cognitive functioning and a .02-0.25 standard deviation increase in height and years of education attained for children ages 8-14. Trapp & Menken (2005) found that improvements in anthropometric outcomes in children resulted from a combination of child health services and reduced competition for resources from siblings, particularly younger ones.

2.2 The MCH-FP Program

The MCH-FP program was initiated in a rural area of Bangladesh, Matlab, in October 1977 by icddr,b. It started as a demonstration project to help the government design their national family planning program. Treatment and comparison areas were built into the design of the program and covered about 200,000 people in 149 villages with the population split fairly evenly between the two areas (Figure 1). The program included integrated family planning (FP) and maternal and child health (MCH) services. Service delivery was intensive as interventions were administered in the house of the beneficiary during monthly visits made by local female health workers hired and trained by the program, and services were free of charge (Bhatia et al. 1980). Usual government health and family planning services were available in the area, but there was limited or no home delivery of services. In addition, many of the program interventions, such as childhood vaccinations and the array of family planning options, only became readily available from the government after 1988, providing us with an \textit{experimental period between 1978-1988 to evaluate the program}. Evaluation of the program is also aided by the rolling out of the program services over two main periods: 1977-81 and 1982-88.
Program services prior to 1982 focused on family planning and maternal health through the provision of modern contraception, tetanus toxoid vaccinations for pregnant women, and iron and folic acid tablets for women in the last trimester of pregnancy (Bhatia et al. 1980). The health workers brought a wide-array of family planning options to the beneficiary’s home including condoms, oral pills, vaginal foam tablets, and injectables. In addition, beneficiaries were informed about fertility control services provided by the project in health clinics, such as intrauterine device insertion, tubectomy, and menstrual regulation. During these visits the female health workers also provided counseling on contraceptives, nutrition, hygiene, and breastfeeding, motivated women to continue using contraceptives, and instructed women how to prepare oral rehydration solution. These services were supported by a well-developed follow-up and referral system to ensure management of side effects and continued use of contraceptives (Phillips et al. 1984, Fauveau et al. 1994).

Between 1982 and 1988, the types of interventions provided were expanded, especially for children under age five. These interventions were rolled out over time in the treatment area starting with measles vaccine in half the treatment area in 1982. Starting in 1985, preventive services were provided to children under the age of five in the entire treatment area. These services included vaccines for measles, DPT, polio, and tuberculosis, and vitamin A supplementation. By 1988, coverage rates for children aged 12-23 months living in the treatment area were 93 percent for BCG, 83 percent for all three doses of DPT and polio, 88 percent for measles, and 77 percent for all three major immunizations (HDSS 2007). Curative care such as nutrition rehabilitation and acute care for respiratory infections was also introduced late in the period. In addition, the tetanus toxoid immunization was expanded to all women of reproductive age, and safe delivery kits were provided to pregnant women.

The program is still running today, but differences between the treatment area and the rest of the country, including the comparison area, diminished after 1988 as the lessons of the Matlab success were incorporated into national plans (Phillips et al. 2003, Cleland et al. 1994). In particular, Bangladesh greatly increased the number of family welfare assistants to deliver in-home contraceptive and immunization services throughout the country. Expanding the number of family welfare assistants reduced the client-worker ratio from 1 per 8000 in 1987/88 to 1 per 5,000 in 1989/90 (Cleland et al. 1994). The ratio was still lower in the treatment area at 1 per 1,300 in 1990. Improvements in supply chains, products, and management were also rolled out in 1988 and 1989 (Cleland et al. 1994).

2.3 Program Targeting and Take-Up
The comparison group was built into the design of MCH-FP (Faveau 1994) but the treatment and comparison villages were not assigned randomly. Instead, the treatment and comparison areas are contiguous geographic areas (Figure 1) that were viewed as socially and economically similar and geographically insulated from outside influences at the time (Phillips et al. 1982). Assigning treatment status to geographic contiguous areas was used to reduce potential contamination of the comparison area from the family planning interventions (Huber & Khan 1979), and was also important for reducing spillovers from the positive externalities generated by vaccination. Past research shows that the treatment and comparison areas are similar with respect

---

2 A direct comparison of the client-worker ratio between the areas is difficult because the health workers in the treatment area had more responsibilities, for example, the collection of regular demographic surveillance data. The national plan was based on cost-effectiveness analysis from other demonstration projects in Bangladesh (Phillips et al. 2003), so it is unclear how much difference there is in access to services based on the ratios alone.
to potential targeting outcomes including rates of mortality and fertility (Koenig et al. 1990; Menken & Phillips 1990; Joshi & Schultz 2013). This is important since it means the program was not placed first in areas that had poor child health or high fertility. Most pre-interventions household and household head characteristics were also similar (Barham 2012, Barham & Kuhn 2014). Two exceptions are religion and access to tubewell water. The difference in access to tubewell water is a result of a government program so does not reflect household income, propensity to drill a tubewell, a household’s concern about child health, or potential unobservable characteristics. In our analysis, we examine heterogenous program effect by pre-intervention access to tubewell water and religion. Lastly, Barham (2012) shows that cognitive functioning was similar between the treatment and comparison area in 1996 for those whose cognitive functioning was not likely to have been affected by the program.

**Figure 1: Map of Matlab Study Area**

Program implementation followed the planned timeline and uptake was rapid. For example, Figure 2 shows that the contraceptive prevalence rate (CPR) for married women 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. After that, there was a steady increase in the CPR with it reaching almost 50% by 1988. Due to availability of contraceptives from government services, the CPR did increase in the comparison area over time, but not as quickly, and rates remained below 20% in 1988. There was still a 20% difference in the CPR rate between the two areas in 1990.

As also shown in Figure 2, the measles vaccination rate rapidly increased to 60% in 1982 after it was introduced in half of the treatment area (treatment area 1), and in 1985 when it was introduced in the other half (treatment area 2). Vaccination rates in the comparison area are not available for this time period, 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 largely an unvaccinated population (Koenig et al. 1991). Nationally, measles vaccination for children under the age of five was less than 2% in 1986 (Khan and Yoder 1998),
and was below 40% in the comparison area in 1990 (Fauveau 1994). As the national program scaled up, these differences narrowed substantially.

**Figure 2:** Trends in Contraceptive Prevalence Rate (CPR) and Measles Vaccination Rates (MVR) for Children 12–59 Months by Calendar Year

![Graph showing trends in CPR and MVR](image)

Source: Contraceptive use data from van Ginniken et al. 1998; Measles vaccination data from iccdr,b Record Keeping System.

### 2.4 Theory of Change Mechanisms Linking the MCH-FP Program and Human Capital Accumulation

The MCH-FP program is a large program that could have affected the lives of many people. In our discussion of the theory of change we focus on two generations that we analyzed in this report. The first is a group of children born during the experimental period (October 1977 to December 1988) who we refer to as the “child” generation. All the mothers of children in this group in the treatment area had access to family planning; however, depending on the year the child was born, access to child health interventions varies. We further split this group into two, based on whether the child was born before or after the rollout of child health interventions in February 1982: those born before the child health interventions rollout (Oct. 1977-Feb. 1982) and those born after the rollout (Mar. 1982-Dec. 1988). The second generation we examine are the children of the “child” group just described, and we refer to them as the “grandchild” generation.

For the child generation, the effect of the MCH-FP interventions on health and nutrition are understood better than the effect on cognitive functioning, one of the main outcomes examined in this report. For this reason, we focus our discussion on mechanisms linking the MCH-FP program to cognitive functioning. The MCH-FP interventions could directly affect cognitive functioning by reducing the incidence of measles and pertussis, since encephalitis, a complication of both these diseases, can lead to long-term brain damage (Greenberg et al. 2005; Reingold & Phares 2006). The program could also affect cognitive functioning indirectly in a number of ways. We provide just a few examples (see Barham 2012 for a longer discussion).
First, vaccine-preventable diseases often lead to high levels of morbidity, especially in developing countries, which affect cognitive development through undernutrition and decreased physical activity and play. Measles, in particular, is known to severely impair a child’s nutritional status through secondary complications such as pneumonia and diarrhea, and prolonged illness from measles can leave a child weakened and at increased risk of illness for a year (Greenberg et al. 2005, Reddy 1987).

Second, the child health interventions may also have indirect effects via sibling competition, with healthier children receiving greater or less parental investment (e.g. in the form of quality time or resources spent on education or health care). Given that the first few years of life are generally believed to be the most important for cognitive development, the effects of sibling competition on cognitive function may be greatest for siblings who are five or fewer years apart in age.

Third, the non-child health components of the MCH-FP program, such as the family planning program, could drive a quality-quantity trade-off, with low-fertility parents having greater resources for their children, such as more or better food and more time. In addition, longer birth spacing resulting from the family planning interventions as well as some of the maternal health inputs (e.g., iron and folate supplementation) may also affect the cognitive development of a child through the improved nutrition and health of the mother while the child is in utero (Walker et al. 2007; Almond & Mazumder 2011). Indeed, Schultz (2009) shows that the MCH-FP led to longer birth spacing and to fewer women with low body mass index.

While cognitive functioning can be affected at all ages, it is believed that the main period when a child’s brain is developing is under the age of five (Grantham-McGregor et al. 2007, Barham et al. 2013a). For this reason we hypothesize that those who were born during the experimental period in the treatment area compared to the comparison area will have better cognitive functioning and be taller (due to better health and nutrition as a child as a result of the program). As a consequence of the improved health, nutrition and cognitive functioning, they will also do better in school, so will attain higher grades. Based on these mechanisms, we also hypothesize that the MCH-FP program will also lead to better human capital outcomes for those born when the child health interventions were available (i.e. Mar. 1982-Dec. 1988) than those born when those program interventions were not available.

The MCH-FP program may, in the long-term, affect the life chances of children of those who themselves received treatment as youngsters (the “grandchild” generation). The key drivers of improved cognitive and nutritional outcomes for the “grandchild” generation are the improved human capital, labor, and marital outcomes of their parents, who were exposed to the MCH-FP program as children. These positive intergenerational effects may be offset by the dual obligations their parents carry - to provide intergenerational support to their own parents (the elder and adult generations) and to their children (the grandchildren). Increased parental support obligations are a result of families having fewer children and there being greater expectations of help from more successful children.

3. DATA

3.1 Data Sources and Sample
This paper draws on the unusually rich data available for the Matlab area, a rural district of Bangladesh, and benefits from the ability to link the various data sources by person and/or household identification number. Three main types of data sources are used to construct the dataset. They include two large socioeconomic surveys (2012-2014 and 1996), several periodic
censuses of the study area, and an ongoing Health and Demographic Surveillance System (HDSS).

The more recent survey, the 2012-2014 Matlab Health and Socioeconomic Survey (MHSS2) was collected especially for this project. The sample design is discussed in Appendix A, and the survey instruments in Appendix B. The main outcome variables for this report are from MHSS2. MHSS2 is a follow-up study to the 1996 Matlab Health and Socioeconomic Survey (MHSS1) which is publicly available from the Rand website (http://rand.org/labor/FLS/MHSS.html) and ICPSR (http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/2705). Both MHSS1 and 2 are multi-purpose surveys. The surveys were designed to be a panel for a subset of MHSS1 respondents, and many of the questions are the same in both surveys. MHSS2, however, has a richer array of outcomes, including enhanced cognitive and health tests, more detailed consumption an employment modules, and greater information on social networks and women’s status, and water samples to test for arsenic.

The second source, periodic censuses (e.g., 1974, 1982) collected by icdr,b on the entire study population (treatment and comparison area), provides information on household location, composition, assets, employment, and education. The 1974 census offers the opportunity to test for preprogram similarity between the treatment and comparison areas, and the 1982 census is used in tracing MHSS respondents back to the 1974 preprogram census.

The third source, demographic surveillance site data on vital events (e.g., births, deaths, migration) was collected by icddr,b on the entire study population of Matlab, and is used to help create an intent-to-treat variable and examine attrition from mortality or outmigration in the study area. The demographic surveillance data have been collected at least quarterly since 1966 on the entire study population (though they are only computerized since about 1976) and are known to be of high quality.

The sample for this paper is a panel of randomly selected participants who were selected to answer the individual books in MHSS1 in 1996. We follow this sample (regardless of if they did or did not respond in MHSS1). Attrition rates in MHSS2 for this sample are remarkably low and approximately 7 percent for MHSS2 as a whole not including those who died. For the sample and outcomes used in this paper the attrition rate including death is approximately 12 percent for those outcomes included in the phone survey of the international migrants and 17% for those outcomes not included in the phone survey.

3.2 Treatment Variable

In this section we describe how we assign treatment status since this is not a randomized intervention. The MCH-FP program used village of residence is used to determine program eligibility. A variable indicating program eligibility based on a person’s 1996 or 2012/15 MHSS village location might be endogenous, since households could have moved to the treatment area to benefit from the MCH-FP program. To avoid this potential endogeneity, individuals are linked to the 1974 census, and village of residence from the 1974 census is used to determine treatment status. Many individuals cannot be linked directly to the 1974 census since they were born after 1974 or moved into the study area between 1975 and 1977. To trace these individuals back to a household in 1974, a dataset is created that indicates each time a person entered or exited the study area and identifies the household head at that time, using census and demographic surveillance data. For those who moved in after 1974, their village of residence the first time they moved into the survey area determines treatment status. For those born after the 1974 census was taken, the head of the household at the time they were born is traced back to the 1974
census and the village of residence of this household head is used to determine treatment status. The intent-to-treat variable, Treatment Area, takes on the value 1 if the individual (or the household head who is traced back to the 1974 census data) resided in a treatment area in 1974, or if the individual (or household head who is traced back) migrated to a treatment village in the study area between 1974 and 1977. The 1974 baseline information is linked to individuals in the MHSS using the same method of tracing back individuals to 1974.

3.3 Primary outcomes to be examined
Given that we study the impact of MCH-FP 35 years after it began, the primary outcomes of this evaluation are as much as possible outcomes related to long-term impacts, in particular cognitive and anthropometric measures. These measures rely on the individual-level survey. The following outcome variable are examined in this report: height, Mini Mental State Exam (MMSE), Digit Span Forward, Digit Span Backwards, and Ravens. With the exception of the MMSE most of these are well known in the development literature. The MMSE is the only cognitive test that was collected both in the 1996 MHSS1 and in the 2012-15 MMSS2 and is described below.

The MMSE measures five areas of cognitive functioning: orientation, attention-concentration, registration, recall, and language (Folstein, Folstein, and McHugh 1975). The test has been widely used to assess higher mental functioning and detect cognitive impairment among adults. Modest to high correlations have been found between the MMSE total score and other tests of intelligence, memory, attention, and executive functioning such as the Wechsler Adult Intelligence Scale (Rush, First, and Blacker 2000). Adaptations of the MMSE are effective at evaluating the cognitive development of children as young as 3 years (Ouvrier et al. 1993; Jain and Passi 2005; Rubial-Álvarez et al. 2007), and it has been shown to correlate fairly well with the Kaufman Brief Intelligence Test for children (Rubial-Álvarez 2007).

The MMSE used in this study is based on the Bangla Adaptation of the Mini Mental State Examination (BAMSE) created by Kabir and Herliz (2000). The BAMSE was designed for an illiterate population and for cultural relevance to Bangladesh, and tests show that the changes made to adapt the instrument do not change the ranking of scores (Kabir and Herliz 2000).

The MMSE asks 33 questions and gives one point for each correct response, for a maximum score of 33. As an example, in the registration section the enumerator reads the respondent a three-sentence story about a house fire and then asks the respondent to repeat the story. The story makes 6 main points (e.g., there are three children in the household, the house is on fire) and the respondent is given a point for each main point he or she repeats. In order to enhance comparison to other studies, the test score for each observation is normalized into a z-score by subtracting the comparison area mean and dividing by the comparison area standard deviation.

The MMSE score increases with age for children (Ouvrier et al. 1993), is on average constant for adults, and then decreases after age 55 or 60 (Strauss, Sherman, and Spreen 2006). The decline in adults is known to persist even when education is controlled. This issue is particularly salient for this paper because of the wide age range being examined. Table 2, panel A, demonstrates that the MMSE score does vary by age in the sample. The MMSE score for adults in this sample is fairly constant until the age of 50, when it starts to decline. Birth year fixed-effects are included in the regression analysis in order to control for this association between age and the MMSE score.
5. EMPIRICAL STRATEGY

For each outcome of interest, we will estimate the intent-to-treat (ITT) or overall program effect of the MCH-FP program on those who were eligible for the program as children during the experimental period. We use interaction models to examine heterogeneous effects. To estimate ITT effects we will use variation in program implementation across locations to compare outcomes in the treatment versus comparison area. We will also exploit phasing-in of interventions over time within the treatment area, which left certain age cohorts differently affected by the program.

This variation across location (treatment and comparison area) and by age group provides an excellent setting for using either single or double difference estimators to determine the ITT effect of the MCH-FP. ITT estimates are derived from the eligible population regardless of whether the treatment was actually received. It is difficult to separate out the effects of the various interventions. However, the combined effect is of great interest since most programs combine these interventions in developing countries and because early childhood vaccination and family planning programs are arguably two of the most important and widespread health programs in developing countries in the latter part of the twentieth century.

<table>
<thead>
<tr>
<th>Year of Birth</th>
<th>Age in 2012</th>
<th>Program Eligibility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mother eligible for family planning, tetanus toxoid vaccine, folic acid and iron in last trimester of pregnancy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Children under age five eligible for measles vaccination in half the treatment area</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Children under age five eligible in entire treatment area for vaccination (measles, DPT, polio, tuberculosis), vitamin A supplementation, nutrition rehabilitation for children at risk (starting 1987).</td>
</tr>
</tbody>
</table>

Notes: The 2012 age groupings are based on age in years rounded to approximate age in December 2012. The exact year and month cutoffs will be used to create groups for the analysis. See a more detailed description of table in Barham (2012).

We will exploit the quasi-experimental research design and use single- and double-difference intent-to-treat models to estimate the effect of the MCH-FP program on adult cognition and anthropometrics for those born during the experimental period. Given that previous research shows that the two areas had similar observable characteristics, single- and double-difference models are appropriate. However, there were more Hindus and better access to tubewell water prior to the program in the treatment area, differences we address in more depth below under
heterogeneous analysis. To account for the differences in baseline characteristics, we include the observables in the regression as a robustness check examine the heterogeneity of the treatment effect by tubewell status in 1974.

Using data on individuals aged 22 to 65, the single difference model for person \( i \) from village \( v \) can be estimated using the following linear equation:

\[
C_{i,v} = \beta_0 + \beta_1(T_{i,v} \ast AG_{i}^{22-29}) + \beta_2(T_{i,v} \ast AG_{i}^{30-34}) + \beta_3(T_{i,v} \ast AG_{i}^{35-65}) + \alpha_{py} + X'Z + \varepsilon_{i,v}
\]

where \( C \) is one of the outcome measures such as height or cognitive functioning. \( T \) is a binary variable that takes on the value 1 if person \( i \), or \( i \)'s household, resided in a treatment village before the MCH-FP program started, and 0 if from a comparison village. \( AG \) is a binary variable used to indicate whether person \( i \) is in age group \( Y \). Thus, the coefficient \( \beta_1 \) is the ITT effect for the 22- to 29-year old age group—the group of children born before the child health interventions were available. \( \beta_2 \) is the ITT estimate for those children eligible to receive less intensive treatment, and \( \beta_3 \) the ITT estimates for the age groups born prior to the program. \( \alpha_{py} \) are birth year fixed-effects to control for differences in the outcome due to age as well as other events that may be correlated with birth year. \( X \) is a vector of individual (gender and religion) and baseline household and household head characteristics. Standard errors are clustered at the village level to account for likely intracluster correlation in the error term.

If the outcome differed between the treatment and comparison areas in the pre-intervention period, a double-difference estimator is more appropriate. Since we do not have pre-intervention data for most of the outcomes, we instead use the 35-65 year old birth cohort since the outcomes of interest are less likely to have been affected by the program. The double-difference estimator can be determined using Equation 2.

\[
C_{i,v} = \beta_0 + \beta_1 T_{i,v} + \beta_2(T_{i,v} \ast AG_{i}^{22-29}) + \beta_3(T_{i,v} \ast AG_{i}^{30-34}) + \alpha_{py} + X'Z + \varepsilon_{i,v}
\]

where variables are defined as above but interpretation differs because the 35-65 year-old group is the group without an interaction term. Therefore, \( \beta_1 \) provides the difference in means between treatment and comparison areas for those 35-65 year olds, or for the “pre-intervention” period. \( \beta_2 - \beta_3 \) are the double difference estimates for the various age groups of interest.

Mechanisms

Understanding which of the program components were most effective is desirable, but, as with many programs, difficult to determine due to the integrated nature of the interventions. However, we can provide some suggestive evidence of the role of the child health interventions. For example, we can take advantage of the fact that the main child health interventions were phased in after 1981 and compare children who were eligible for the program at birth (the 22-29 year olds) to those who would have been eligible in later childhood (the 30-34 year olds). We can also exploit the phasing in of the measles vaccine over time within the treatment area, by comparing the 22-26 year olds to the 27-29 year olds within the treatment area. In addition, we can include a control in equations 1 and 2 indicating if a person’s mother was eligible for family planning, and run the treatment-on-the treated models discussed above.

6. IMPACT RESULTS
In this section, we describe our research findings to date. Given the scope of the data collection, the long-term nature of the data and that the data is only recently available these results are still preliminary. Future versions of this report will try to incorporate additional controls to account for other potential time varying observables, such as a flood embankment, the introduction of the micro-credit program BRAC, and changes in school supply. We will also examine how differential selection from attrition is affecting the results.

6.1 Treatment Comparison Group Balance

We test whether the areas are similar using preintervention household and household head characteristics from the 1974 census. Table 1 provides the means and standard deviations (SD) of the characteristics for the treatment and comparison areas for the entire sample. The differences in means are statistically insignificant at the five percent level for all variables except religion, drinking water sources, latrine, and age of household head and household head spouse. Since t-statistics are driven by sample size as well as magnitude of the difference, we also examine the normalized differences (difference in the means divided by the SD of the difference) to get a sense of the size of the differences. Imbens and Wooldridge (2008) argue normalized differences greater than 0.25 are substantial. The mean of the differences that do exist are relatively small and less than 0.10 for all the differences except access to tubewell water which is 0.14. These findings, together with previous results on fertility and mortality, strongly suggest that the two areas had very similar observable characteristics. Baseline characteristics for the experimental group (those aged 22-29 and 22-34) are similar.

Before the program, 14 percent greater proportion of households used tubewell water for drinking, which is concerning since tubewell water is often thought to be cleaner than other sources of water. Because a larger percent of treatment area households had access to this water, the program effect might be biased upwards. Unfortunately, there is widespread groundwater arsenic contamination in the tubewells in Bangladesh (Chowdhury et al. 2000; Alam et al. 2002). Arsenic is a serious health concern and has been shown to reduce IQ among school-aged Bangladeshi children (Wasserman et al. 2006). So greater access to tubewell water in the treatment area might actually bias the estimate of program impacts downwards. In section 5.5, we examine the heterogeneity of the treatment effect with the source of drinking water to help determine whether such a bias exists.

6.2 Main Results

The ITT effects are presented in Table 2 for the three main age groups. Those born before the program started, the 35-65 year olds, those born during the experimental period but in the first part of the program when the main interventions were family planning, the 30-34 year olds, and those born when the child health interventions became available, the 22-29 year olds. Double difference results are presented in panel A, and those born prior to the program as used to make the double difference. Single difference results are presented in Panel B.

Table 2 presents results for five outcomes: height, Mini Mental State Exam (MMSE), Digit Span Forward, Digit Span Backwards, and the Ravens. Both height and MMSE were collected in the earlier round of data from 1996. For both height and MMSE column 1 reports results from the 1996 data. Results for this time period are report in Barham (2012) but are remade for this
report. In 1996, improvements in height and MMSE were only experienced by the 22-29 year olds (who were 8–14 in 1996), the age cohort that benefited from the child health interventions. There is a statistically significant 0.22 standard deviation (SD) difference between the treatment and comparison area in height and 0.38 SD for the MMSE score. Importantly, and as expected, the point estimate is small for the variable Treatment Area, which gives the difference in means between the treatment and comparison areas for the 35–65 year old age group. For the 30-34 year old group (15–19 in 1996) the point estimates are fairly small but significant at the 10 percent level for MMSE in the single difference model. Any positive effect the family planning and maternal health interventions may have had on cognitive development of this age group may have been swamped by sibling competition from younger siblings who received more intensive child health interventions.

To compare results between the 1996 and 2012/15 surveys, the sample is restricted to only those people who had height or MMSE information in both waves of the survey (columns 2 and 3 for height and column 7 and 8 for MMSE). While this sample is likely to be biased due to attrition it does allow comparison of results between the two waves of the sample on a panel of people. The effects on height are similar between the two rounds of the survey with a point estimates of 0.23 SD in 1996 and 0.19 SD in 2012/15 survey for the 20-29 year olds. However, the program effects on the MMSE for the 20-29 year do not persist between the two rounds of the survey dropping from 0.33SD in 1996 to -0.03 SD in the 2012/15 survey. The last columns for height and MMSE (columns 5 and 9) show results for the full sample in 2012/15 survey regardless of if 1996 information was available. Results for this larger sample are similar.

Table 2 presents results for the three other measures of cognition that were not included in the 1996 survey, digit span forward and backwards, and the Ravens. Results are similar for these three measures of cognition. The pattern of results are similar to the MMSE for the 20-29 year olds. Results for the other two age groups are similar to 1996, and highlight that there were no statistically significant program effects for any of the groups.

6.3 Program Effects Using the Phasing-in of Interventions in Treatment Area

We exploit the phasing-in of the measles vaccination over time within the treatment area to provide an additional estimate of the ITT effects of the child health vaccinations on cognitive development, and to estimate an effect that better controls for the family planning and maternal health interventions. Children under the age of five in half the treatment area (Treatment Area 1) were eligible to receive the measles vaccine starting in March 1982 and in the other half (Treatment Area 2) in November 1985. As a result, the 27–29 year olds in Treatment Area 1 had been eligible to receive the measles vaccination at the recommended age of 9 months, while those in Treatment Area 2 had been eligible only past the recommended age. The 22-25 year olds in both Treatment Area 1 and 2 were eligible at the same time so there is no longer two treatment areas. Therefore, the 27–29 year olds provide an opportunity to examine whether the program effect differs for children who were eligible to receive the measles vaccination at the recommended age of 9 months versus those were eligible later, better isolating the effect of just the child health interventions. The inclusion of Treatment Area 2 for this age group may even have led to a downward bias on the point estimate for the 22-29 year old age group.

---

3 Height and weight z-scores are made from internal standardizations from the 1996 survey for both the 1996 and 2012/15 results. So effect sizes are comparable across the waves of the survey.
Table 3 disaggregates the 22-29 year olds into Treatment area 1 and 2. Column 1 and 3 provide program effects for 1996 for height and MMSE and show that the program effects were largest and statistically significant at the 1 percent level for Treatment Area 1 at 0.26 SD for height and 0.49 SD for MMSE. Effects sizes in Treatment Area 2 were 0.19 SD for height (statistically significant at the 5 percent level), and 0.26 SD for the MMSE but not statistically significant. For height, by 2012/15, the program effect in treatment area 1 remained the same at 0.26 SD, but fell to 0.10 and was insignificant for Treatment Area 2. For the MMSE, the program effect in 2012/15 was small and statistically insignificant for both Treatment Area 1 and 2.

6.4 Program Effect Heterogeneity

Table 4 examines if the program effect for the 22-29 year old age group differs by preintervention access to tubewell water, household head education, and if the household head works in agriculture. As shown in section 5.1, it is possible that the findings may be biased by differences in access to tubewell water before the program began between treatment and comparison area. Interaction effects of the double difference estimator for children age 22-29 with a binary variable indicating whether the household used tubewell water for drinking in 1974 are presented in first column for each outcome. The interaction between age group and tubewell water was also included but is not reported. Depending on the outcome, the effect size varies in sign, but are relatively small and all interactions are statistically insignificant. Heterogeneous effects for the 22-19 year old age group by education of the household head and if the household head works on agriculture are also statistically insignificant. Thus, there does not appear to be heterogeneous treatment effects at least with respect to these three important variables.

7. CONCLUSIONS AND POLICY RECOMMENDATIONS

Improving the health and nutrition of young children is important not only for immediate well-being, but also because it is believed to reduce poverty in the long-run through improved human capital. Many program such as Head Start and Conditional Cash Transfer (CCT) programs rely on this postulated link. Little, however, is known about the long-term effects of programs targeted to improve health and nutrition in early childhood on human capital and labor market outcomes in adulthood. Determining the causal effects of early child-health interventions on later human capital formation and labor market outcomes is challenging. A growing literature suggests that large negative health or nutrition shocks early in life, lead to worse outcomes later in life, but there is little long-run evidence on the effects of interventions designed to improve the health and nutrition of young children. Understanding the longer-run effects of early childhood interventions is important as there is growing interest in investing resources in disadvantaged children at an early stage in life, for example through the spread of poverty reduction programs like CCTs. It is crucial to investigate these questions since evidence on early childhood nutrition and health interventions is mixed as to whether their benefits continue (Pollitt et al. 1993) or fade out (Garces et al. 2002).

Previous research shows the MCH-FP program led to important improvement in human capital in early and late childhood (ages 8-14). This study examines if these effects persist when these same people are ages 22-29. We find that while the effect on height persists the effects on the MMSE fade out and there are no statistically different effects between the treatment and
comparison areas for the other measures of cognition that were measured only at age 22-29. No heterogeneous results were found with respect to access to tubewell water, and household head occupation and education.

The difference in results between height and MMSE highlight that physical growth and cognitive development are affected differently by one’s environment. More work is needed to better understand the fading out of the MMSE effects over time. The medical literature highlights the importance of in utero development (Barker 1992) as well as the risk of growth faltering from birth to age two (Victora et al. 2010). While family planning and childhood vaccination are clearly important interventions for improving the health of children, earlier investment in childhood health and nutrition, or continued intervention at later ages may be needed to lead to sustained improvements in cognitive functioning. Difference between height and cognition may also be due to the nutrition content of the food though, unfortunately, such analysis this is beyond the scope of this report due to the lack of information. Regardless, the study highlights that height that physical growth and cognitive functioning develop differently over time.

Overall, the results highlight that the MCH-FP program led to important and sustained effects on a person height though adulthood. There were important effects of cognitive functioning through at least late childhood, that faded out as they reached adulthood. Future research needs to investigate if these effects on cognitive functioning through late childhood still led to gains in completed education and labor market outcomes.
REFERENCES


## TABLE AND FIGURES:

### Table 1: 1974 Baseline Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Treatment Area</th>
<th>Comparison Area</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Family size</td>
<td>6.95</td>
<td>5.98</td>
<td>6.76</td>
</tr>
<tr>
<td>Owns a lamp (=1)</td>
<td>0.64</td>
<td>1.37</td>
<td>0.61</td>
</tr>
<tr>
<td>Owns a watch (=1)</td>
<td>0.16</td>
<td>0.86</td>
<td>0.16</td>
</tr>
<tr>
<td>Owns a radio (=1)</td>
<td>0.08</td>
<td>0.59</td>
<td>0.08</td>
</tr>
<tr>
<td>Wall tin or tinmix (=1)</td>
<td>0.31</td>
<td>1.04</td>
<td>0.31</td>
</tr>
<tr>
<td>Tin roof (=1)</td>
<td>0.83</td>
<td>0.74</td>
<td>0.84</td>
</tr>
<tr>
<td>Latrine (=1)</td>
<td>0.82</td>
<td>1.00</td>
<td>0.86</td>
</tr>
<tr>
<td>Number of rooms per capita</td>
<td>0.22</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>Number of cows</td>
<td>1.50</td>
<td>3.80</td>
<td>1.39</td>
</tr>
<tr>
<td>Number of boats</td>
<td>0.66</td>
<td>1.62</td>
<td>0.69</td>
</tr>
<tr>
<td>Drinking water, tubewell (=1)</td>
<td>0.31</td>
<td>1.45</td>
<td>0.16</td>
</tr>
<tr>
<td>Drinking water, tank (=1)</td>
<td>0.39</td>
<td>1.97</td>
<td>0.33</td>
</tr>
<tr>
<td>Drinking water, other (=1)</td>
<td>0.30</td>
<td>2.50</td>
<td>0.50</td>
</tr>
<tr>
<td>HH age</td>
<td>47.86</td>
<td>26.87</td>
<td>46.20</td>
</tr>
<tr>
<td>HH years of education (edu.)</td>
<td>2.45</td>
<td>6.52</td>
<td>2.31</td>
</tr>
<tr>
<td>HH works in agriculture (=1)</td>
<td>0.60</td>
<td>1.13</td>
<td>0.59</td>
</tr>
<tr>
<td>HH works in fishing (=1)</td>
<td>0.06</td>
<td>0.66</td>
<td>0.07</td>
</tr>
<tr>
<td>HH spouse's age</td>
<td>37.13</td>
<td>21.79</td>
<td>35.93</td>
</tr>
<tr>
<td>HH spouse's years of edu.</td>
<td>1.09</td>
<td>3.39</td>
<td>1.21</td>
</tr>
</tbody>
</table>

Notes: Standard deviations (SD) are clustered at the 1996 village level. Sample is for the primary book 35 respondent from MHSS1 with non-missing height information in 2012-15 sample. Includes phone survey respondents. Results are similar if the phone survey respondents are left out and for only those born in the experimental period 1978-1989.
## Table 2: Intent-To-Treat Program Effects by Age Group

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: DD results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area (=1)</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)</td>
<td>0.22*</td>
<td>0.23*</td>
<td>0.19+</td>
<td>0.20*</td>
<td>0.21*</td>
<td>0.38*</td>
<td>0.33*</td>
<td>-0.03</td>
<td>-0.04</td>
<td>0.02</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.15)</td>
<td>(0.16)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>0.06</td>
<td>0.07</td>
<td>-0.06</td>
<td>-0.00</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.13</td>
<td>0.14+</td>
<td>-0.08</td>
<td>-0.08</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(0.10)</td>
</tr>
<tr>
<td><strong>Panel B: SD results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area*(Age 35–65)</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.04</td>
<td>-0.03</td>
<td>0.03</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)</td>
<td>0.17**</td>
<td>0.17**</td>
<td>0.18+</td>
<td>0.20*</td>
<td>0.35*</td>
<td>0.30+</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.15)</td>
<td>(0.17)</td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>0.02</td>
<td>0.01</td>
<td>-0.06</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.15+</td>
<td>0.16+</td>
<td>-0.10</td>
<td>-0.06</td>
<td>-0.05</td>
<td>-0.07</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
<tr>
<td><strong>Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996 Primary Sample</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>2012-15 Follow-up Sample</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Outcome not missing in 1996</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Outcome not missing in 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Phone surveys included</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Observations</td>
<td>5,724</td>
<td>5,092</td>
<td>5,917</td>
<td>5,630</td>
<td>4,675</td>
<td>4,000</td>
<td>4,000</td>
<td>5,625</td>
<td>5,914</td>
<td>5,619</td>
<td>5,622</td>
</tr>
</tbody>
</table>

Notes: Standard errors are clustered at the village level. ***, ***, or + indicates that the difference in the coefficient from zero is statistically significant at the 1 percent, 5 percent, or 10 percent significance level respectively. Individual characteristics include year of birth fixed-effects and controls for gender and religion. Preintervention characteristics included. DD = double difference, SD = single difference. Outcomes are internally standardized z-scores. Age is as of 2012.
Table 3: Intent-to-Treat Effects Disaggregated in the Treatment Area

<table>
<thead>
<tr>
<th>Height</th>
<th>MMSE</th>
<th>Digit Span Forward</th>
<th>Digit Span Backwards</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Treatment Area*(Age 35–65)</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area 1 *(Age 27–29)</td>
<td>0.26**</td>
<td>0.26*</td>
<td>0.49**</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.11)</td>
<td>(0.17)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area 2 *(Age 27–29)</td>
<td>0.19*</td>
<td>0.10</td>
<td>0.26</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.18)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area *(Age 22–26)</td>
<td>0.11</td>
<td>0.21</td>
<td>0.31</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.17)</td>
<td>(0.21)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.15+</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.06)</td>
<td>(0.09)</td>
<td>(0.07)</td>
</tr>
</tbody>
</table>

1996 Primary Sample | Y | N | Y | N | N | N | N |
2012-15 Follow-up Sample | N | Y | N | Y | Y | Y |
Outcome not missing in 1996 | Y | N | Y | N | N | N |
Outcome not missing in 2012-15 | N | Y | N | Y | Y | Y |

Observations | 5,724 | 5,630 | 4,675 | 5,625 | 5,619 | 5,623 | 5,588 |

Notes: Standard errors are clustered at the village level. ***, **, or * indicates that the difference in the coefficient from zero is statistically significant at the 1 percent, 5 percent, or 10 percent significance level respectively. Individual characteristics include year of birth fixed-effects and controls for gender and religion. Preintervention characteristics included. DD=double differene, SD =single difference. Outcomes are internally standardized z-scores.
Table 4: Intent-to-Treat Program Effect Heterogeneity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
<th>(11)</th>
<th>(12)</th>
<th>(13)</th>
<th>(14)</th>
<th>(15)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Height</strong></td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.02</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td><strong>MMSE</strong></td>
<td>0.21*</td>
<td>0.25+</td>
<td>0.12</td>
<td>-0.00</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0.02</td>
<td>-0.02</td>
<td>0.05</td>
<td>0.01</td>
<td>-0.03</td>
<td>0.05</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(0.13)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(0.11)</td>
</tr>
<tr>
<td><strong>Digit Span Forward</strong></td>
<td>-0.00</td>
<td>-0.00</td>
<td>-0.00</td>
<td>0.14+</td>
<td>0.14+</td>
<td>0.14+</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.09</td>
<td>-0.09</td>
<td>-0.09</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td><strong>Digit Span Backward</strong></td>
<td>0.00</td>
<td>-0.10</td>
<td>-0.10</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.07</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.13)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
</tr>
<tr>
<td><strong>Raven</strong></td>
<td>-0.08</td>
<td>-0.08</td>
<td>-0.08</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.05)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.12)</td>
<td>(0.12)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>5,917</td>
<td>5,917</td>
<td>5,917</td>
<td>5,625</td>
<td>5,625</td>
<td>5,625</td>
<td>5,914</td>
<td>5,914</td>
<td>5,914</td>
<td>5,623</td>
<td>5,623</td>
<td>5,623</td>
<td>5,588</td>
<td>5,588</td>
<td>5,588</td>
</tr>
</tbody>
</table>

Notes: Standard errors are clustered at the village level. "*" and "+" indicates that the difference in the coefficient from zero is statistically significant at the 1 percent, 5 percent, or 10 percent significance level respectively. Individual characteristics include year of birth fixed-effects and controls for gender and religion. Preintervention characteristics included. DD = double difference, SD = single difference, HH = household head. Use 2012-15 followup sample from column 5 of Table 2. Outcomes are internally standardized z-scores.
Table 5: Intent-to-Treat Program Effect by Subcomponent of MMSE (z-scores)

<table>
<thead>
<tr>
<th>Subcomponent</th>
<th>Orientation</th>
<th>Attention-Concentration</th>
<th>Recall</th>
<th>Registration</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: DD results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area (=1)</td>
<td>0.04</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)</td>
<td>-0.05</td>
<td>0.04</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>-0.18**</td>
<td>-0.02</td>
<td>-0.08</td>
<td>-0.10</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
</tr>
<tr>
<td><strong>Panel B: SD results</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area*(Age 35–65)</td>
<td>0.04</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.04</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.03</td>
<td>0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>-0.13**</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.06</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Observations</td>
<td>5,625</td>
<td>5,625</td>
<td>5,625</td>
<td>5,625</td>
<td>5,625</td>
</tr>
</tbody>
</table>
APPENDIX A: Sample Design

In contrast to the majority of evaluations recently funded under 3ie, this research project is not a “evaluation; rather, it is a follow up to an intervention initiated in 1977, 35 years before the present project, and concluded in 1989. Therefore, we do not describe in detail program implementation (but rather use this section to detail the careful procedures we used to 1) improve our sample to contain a sample of beneficiaries and descendants suitable to conduct multigenerational analysis, 2) improve the large number of out-migrants from the study area, and 3) use Matlab Health and Demographic System (MHDSS) and 1974 census data to better account for program exposure, attrition, and hot formation.

A.1. Constructing a multigenerational panel

Conducted in 1996, the first Matlab Health and Socioeconomic Survey (MHSS1) was collected on a probability sample of the 1993 population of the MHDSS area. At the time, no efforts were taken to select or sample in accordance with the distribution of the MHDSS area population in 1974, a period prior to program start when an MHDSS census exists. Migrants who left the MHDSS area prior to 1996 were followed. Our existing research has used the 1974 MHDSS area census to check for baseline differences between treatment and comparison areas (Barham 2012), to analyze treatment-comparison variation in migration, and to error-bound existing analysis to account for differential mortality and out-migration between program initiation in 1977 and MHSS1 in 1996.

To address these potential limitations and to ensure tracking of all descendants of MHSS1 respondents, we situated the MHSS1 study population within a Matlab Linked Database (MLD) of MHDSS residents to the present, allowing us to identify our potential sample in advance, include migrants who departed the MHDSS area prior to 1996 who were followed. The sample for MHSS1 was based on a probability sample of 1993 households (the Primary Sample), each household a related nearby household was included in the Secondary Sample. MHSS2 follows on the Primary Sample households. While all members of the households are listed and basic information provided, only a subsample of household members were interviewed. The tracing forward of the MHSS1 began with the 11,165 individual respondents to MHSS1, defined as those who responded either to the Interview (Book III) or the Child Interview (Book V). (See Appendix B for fuller description of the instruments.) To account for the absence of migrants departing between 1977 and 1996, we used a select an additional sample of pre-MHSS1 migrants from the Primary Sample households who we included in MHSS1. Given budget constraints, we focused our efforts on including all children born to respondents during the program years 1977 to 1988. These “pre-MHSS1 migrants” totaled 120. We included MHSS1 respondents and pre-MHSS1 migrants as “panelists.” They were 11,285 in number. Beginning with this set of panelists, we will describe our followup survey inclusion procedure in principle and then in practice.

To maximize the capacity to evaluate programmatic effects across generations, sampling focused on including the children (and spouses) of panelists, particularly those children not yet born at the time.

Our set of descendant children included the following:

i) All children born since 1996 to panelists, irrespective of coresidence

ii) All other children of panelists who coresided with the panelist in the MHDSS area since

This child tracing rule was applied recursively to the grandchildren and great-grandchildren of panelists. Taken together, the children, grandchildren, etc. constitute the group of descendants. As with panelists, we followed all descendents to any location within Bangladesh and conducted a phone survey for males outside the country, except as noted below.

Because spouses are important both in determining the life changes of beneficiaries and their children, we also included a larger number of spouses of panelists and descendants, including
i) All spouses currently co-residing with a panelist or descendant

ii) For all marriages in which a descendant was in the MHDSS area and the spouse was outside MHDSS, we tracked the spouse as a migrant.

iii) For all marriages in which a descendant was a migrant and the spouse was in MHDSS, we interviewed the spouse living in MHDSS.

If the descendant was outside the MHDSS area and their spouse was living in a different household outside the MHDSS area, we did not track the spouse. We note that the survey included extensive proxy data on spouses who were not interviewed.

Taken together, the panelists, descendants, and selected spouses constitute the target respondents of MHSS2. The universe of households included in MHSS2 includes all households that had at least one target respondent on December 31, 2011, when data for pre-population of questionnaires was extracted.

To better understand the economic context of the households in which target respondents live, we also included all non-target household members in 25% of Matlab MHSS2 households at the time of sample pre-population on December 31, 2011. In the remaining 75% of MHSS2 Matlab households and in all migrant households, we included non-target household members in the household listing and accounts for their activities in the Household Economy Book (Book 2 of the survey), but did not interview them for individual survey and testing books. In migrant households, we only interviewed descendants and spouses.

Finally, to better represent the population of MHDSS in 2012, we included a 7% random sample of households that were new to the MHDSS area since MHSS1 in 1996. Specifically, we identified households from the year 2012 that included no members who were residents or descendants of a 1996 HDSS household, and thus could not have been sampled in 1996.

A.2. Updating the target population in the field

As mentioned above, pre-population of questionnaires was carried out using the information available in MHSS1 and the continuously collected MHDSS. Although pre-population dramatically improved data quality and the ability to track descendants, we nonetheless had to include several field steps to ensure full coverage of the target population of descendants.

1) All new births and migrants entering a Matlab household subsequent to pre-population were added to the household roster. The descent rules described in section A.1 were applied to these new members, and any descendant members were assigned for interview.

2) All descendants pre-populated to a Matlab household who had subsequently migrated had to be shifted into the migrant tracking system.

3) Unpre-populated descendants - A small number of children born after 1996 to descendants had never lived in Matlab and thus would not have been included in the pre-population. As part of the MHSS2 individual interview, adult respondents were asked to report on all children, including those who were or weren’t pre-populated, in the Individual Control Book. Interviewers were trained to add any children born after 1996 who were not pre-populated (and thus had not lived in the HDSS area) to the migrant tracking module.

A.3. Data checks to ensure quality of tracking and descent linkage

To ensure high quality of descendant coverage and linkage, we conducted a number of real-time post-survey checks. Some of these checks were incorporated into our larger system of consistency checking of the data, while others were part of a dedicated sample tracking system.

1) Once Matlab sample fieldwork was concluded, residence and descent information were updated in the computerized database. Any descendants who should have been interviewed or tracked as descendants were assigned for reinterview or migrant tracking.
2) To ensure tracking of unprepuplated descendants, we searched the listings of children in the Individual Control Book Child (CH) section for any children who should have been followed but had not.

3) Parent-spouse-child reconciliation – The post-field checking process included linking all parents, spouses, and children to one another. This process could yield additional respondents for interview/tracking (for instance a parent who should have been coded as a spouse). In addition, it could reveal unknown linkages between a respondent who was prepuplated and subsequent listed by a migrant parent. This problem is relatively common when a father migrated before marriage and thus was not linked in MHDSS to his wife or children.
APPENDIX B: Survey Instruments: Overview of Questionnaires

MHSS2 is an extensive multi-purpose survey comparable to MHSS1, though modules were added and adjusted as necessary (http://www.rand.org/labor/FLS/MHSS/). The household survey, implemented in the household of the respondent, includes household information, GPS coordinates, and individual questionnaires for each household member. Community, school facility, and health facility questionnaires and GPS coordinates were collected for all villages and facilities in Matlab.

A copy of each of the instruments described below is included in a separate file. Since cognitive tests are proprietary, only the anthropometry and physical test modules are included for Books 6A and 6B.

The Household Survey consists of 10 instruments
1. Book HC: The Household Control Book
   The Respondent is Head of Household, Spouse of Head of Household or HH Member 15 years or older who is knowledgeable about characteristics of HH Members. Book HC contains the household listing and basic information on each member. It also records results of arsenic testing of well water.
2. Book IC: The Individual Control Book
   The Respondent is an adult 15 years or older. Book IC is administered to all adults in Matlab households and migrants, spouses, and their children 15 years and older in households outside Matlab. It collects lists and basic information on parents, spouses, children, and migration.
   The Respondent is Head of Household, Spouse of Head of Household, or HH Member 15 years or older who is knowledgeable about characteristics of HH Members. Book 1 collects information on household characteristics, water resources and well switching, and consumption.
   The Respondent is Head of Household, Spouse of Head of Household, or HH Member 15 years or older who is knowledgeable about characteristics of HH Members. Book 2 collects information on household agriculture and non-agriculture income and assets, and borrowing.
5. Book 3: Adult Book
   The Respondent is anyone eligible to complete Book IC. Book 3 covers a myriad of topics including employment, education, health, marriage, migration, social networks etc.
   The Respondent is an ever-married woman aged 15-54 who was eligible to complete Book IC. Book 4 includes pregnancy and contraceptive use histories,
7. Book 5: Child Book
   Respondent is mother or primary caretaker of child 0-14 years of age who is a descendant of an MHSS1 household member. Book 5 covers education and health of the child.
8. Book 6A: Tests for Children 0-6
   Respondent is child aged 0-6 who is a descendant of member of MHSS1 household.
9. Book 6B: Tests for Ages 7 and above
   Respondent is a member of an MHSS1 household, spouse, or descendant aged 7 and above.

Cognitive testing in Books 6 covers all ages and includes: the Mini Mental State Exam, the Raven’s, WPPSI matrix reasoning, Denver Developmental Screenign Test, digit span forward and backwards, a short-term memory test, a processing speed test, a stroop-like test, and a math test. Data was collected on home environment, temperament, depression, and locus of control. When necessary, tests were adjusted modestly to improve local understanding. All tests were extensively pretested and translated with a local Medical Doctor who also has a PhD in Child Development from University College London. The objective health tests include anthropometrics, grip strength, lung function, blood pressure, and objective measures of physical capacity.
   Collected from respondents to any of books 1-6B. Toenail samples were collected for later analysis of arsenic levels.

*The Phone Survey* consists of a subset of questions from the Household Survey instruments that can be answered without a face-to-face interview.
   The respondent is a person aged 15 or older who was eligible for the full Household Survey but is an international migrant.

*The Community/Facility Survey* consists of six instruments:
1. Men’s Community Survey
   Respondents were interviewed in a group of approximately 4 people from the community. The goal was to recruit a knowledgeable farmer, a business person, an elected leader, and an elder, with at least one group member an older person and at least one group member being well-educated.
2. Women’s Community Survey
   Respondents were interviewed in a group of approximately 4 people. The goal was to recruit a teacher or school committee member (someone knowledgeable about schools), an NGO worker, and social worker, with at least one group member being well-educated.
3. Health Facility Survey
   Respondent is the Head of the facility
4. Health Provider Survey
   Respondent is a private practice doctor or village doctor
5. School Survey
   Respondent is the Headmaster or the Principal of the school.
6. Pharmacy Location Survey
   Surveyor entered GPS coordinates of each pharmacy. The owner of the largest pharmacy in each village also completed a Health Facility Survey.

*The Market Prices Survey* was completed in "weekly markets", "regular (ie.- permanent) shops", and in "supermarkets" when supermarkets were available. The market surveys was repeatedly collected in 5 locations in Matlab and then multiple times in Dhaka and in most migrant areas.