Thirty-five years later: evaluating effects of a quasi-random child health and family planning programme in Bangladesh

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

Randall Kuhn
Josef Korbel School of International Studies, University of Denver

Jane Menken
Institute of Behavioral Science, University of Colorado, Boulder

Abudur Razzaque
icddr,b

Grantee Final Report
Accepted by 3ie: August 2015
**Note to readers**

This impact evaluation has been submitted in partial fulfilment of the requirements of grant OW3.1060 issued under Open Window 3. 3ie is making it available to the public in this final report version. All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Any comments or queries should be directed to the corresponding author, Randall Kuhn at rkuhn@du.edu.


Funding for this impact evaluation was provided by 3ie’s donors, which include UKaid, the Bill & Melinda Gates Foundation, Hewlett Foundation and 12 other 3ie members that provide institutional support. A complete listing is provided on the [3ie website](http://www.3ieimpact.org).
Acknowledgements

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.
# TABLE OF CONTENTS

ABSTRACT ............................................................................................................ 2

1. INTRODUCTION ............................................................................................... 3

2. CONTEXT ............................................................................................................. 4
   2.1 Background of Intervention .............................................................................. 4
   2.2 Related Literature .............................................................................................. 5

3. DESCRIPTION OF INTERVENTION AND THEORY OF CHANGE .............. 7
   3.1 The MCH-FP Program ..................................................................................... 7
   3.2 Program Targeting and Take-Up ...................................................................... 8
   3.3 Data and Evaluation Surveys ........................................................................... 10
   3.4 Theory of Change Mechanisms Linking the MCH-FP Program and Human
       Capital Accumulation ....................................................................................... 10
   3.5 Outcomes ........................................................................................................ 11

4. MIGRANT TRACKING PROTOCOLS AND ATTRITION .............................. 12

5. IMPACT RESULTS ............................................................................................ 18
   5.1 Definition of Treatment ................................................................................. 18
   5.2 Treatment Group Balance and Attrition Balance ......................................... 18
   5.3 Main Results ................................................................................................... 19
   5.4 Program Effects Using the Phasing-in of Interventions in Treatment Area...... 21
   5.5 Program Effect Heterogeneity ........................................................................ 21
   5.6 Internal and external validity .......................................................................... 22

6. CONCLUSIONS AND POLICY RECOMMENDATIONS ............................ 22

7. REFERENCES ..................................................................................................... 25

8. TABLES ............................................................................................................... 29

APPENDICES ....................................................................................................... 35
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 or fade out.

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 in 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 to be surveyed.

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 persist into adulthood the effects on the cognition do not. The difference in results between height and cognition highlight that physical growth and cognitive development may be affected differently by one’s environment and do not necessarily follow each other.
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 then ultimately 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, on human capital measures such as health and cognitive functioning 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 35 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 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 human capital in adulthood as measured by height and cognitive functioning. In particular, we use single and double difference models to examine if improvements in human capital at ages 8-14 for those born during the experimental period, persist when these same people are ages 22-29.

Given the long-term nature of this study it was paramount 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 could not be interviewed 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 born during the experimental period; the group with the highest migration rates in the study.\(^1\) 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 of time.

2. CONTEXT

Prior to discussing the long-term results, in this section we provide background on the interventions, discuss related relevant literature, review the short-term evidence of the MCH-FP program. The short-term impacts provide key contextual factors for the interpretation of the longer-term benefits.

2.1 Background of Intervention

In the early 1960s, icddr,b began the Matlab Health and Demographic Surveillance System (MHDSS) that provided at least monthly data on fertility, mortality, marriage and divorce, change of residence within the study site, and destination of out-migrants from the area for a population of approximately 200,000 people residing in a geographically defined area (Fauveau 1994). After the disruption of a devastating liberation war, cyclone, flood and famine, Bangladesh embarked on a systematic effort to develop an evidence-based approach to improve the health of mothers and children and promote family planning. In 1977, icddr,b instituted the MCH-FP program in a treatment area comprising about half the MHDSS population. MHDSS data collection continued in this and the remaining comparison area that was geographically contiguous (see Figure 1). The intention of the MCH-FP program was to develop cost-effective interventions that could be scaled up to the national level. Because of limitations on women’s mobility imposed by the *purdah* system, women of childbearing age received doorstep delivery of contraceptives and antenatal care. Children received in-home vaccination delivery, with increasing coverage of diseases phased in over time, as well as services directed toward prevention, management, and medical referral for childhood diarrhoeal and acute respiratory illness (see Section 3.1 for more details on the program). Similar in-home child health interventions did not begin to be rolled out in the comparison area until around 1989 and vaccinations where not available regularly in government clinics before 1988.

In 1996, a group of researchers (including many members on this 3ie grant) led the collection of the Matlab Health and Social Survey 1 (MHSS1). MHSS1 led to a number of important findings on medium-term impacts of the MCH-FP interventions discussed below. In 2012, a follow-up survey (support by this 3ie grant, National Institutes of Aging and the Population Reference Bureau) started in order to determine the effect of these important health and family planning interventions 35 year after the start of the program.

\(^1\) Such tracking was the primary objective and use of the funding provided by 3ie, with the main data collection supported by NIA.
2.2 Related Literature

Each year more than 7 million children die from preventable causes (Hill et al. 2012). Even more children suffer from illnesses such as vaccine-preventable and diarrheal diseases and malnutrition that affect immediate welfare and may lead to conditions that impair learning and cognition. The latter may have consequences later in life such as on completed education, marriage, and labor market outcomes (Case et al. 2005, Fogel 1994, Haas et al. 1996, Martorell 1995 and 1999, Martorell et al. 1994). In fact, it is estimated that over 200 million children under age five are failing to reach their cognitive potential owing to deprivation early in life (Grantham-McGregor et al. 2007). Research has also demonstrated the critical role of fertility reduction in improving women’s health and survival and in promoting parental investments in children’s human capital (NRC 1986, Ahlburg 1998). As a result, programs that promote the health of mothers and children through access to key health interventions and family planning are typically seen as a critical path not only to alleviating short-term effects of poverty, but also to reducing its long-term and intergenerational effects (Almond & Currie 2011, Ahlburg 1998, Bloom et al. 2004, Bloom & Canning 2001, Sachs & Malaney 2002, Strauss & Thomas 1998). Yet few studies investigate long-term effects of targeted health interventions, and fewer still address these issues outside the historical context of today’s more developed countries.

2.2.1 Broader Evidence of Early-Life Health and Nutrition Interventions on Later Life Outcomes

A growing literature suggests that large negative health or nutrition shocks early in life, such as from flu pandemics or famines, led to worse outcomes later in life.2 However, there is little causal evidence on the long-run effects of interventions designed to improve the health and nutrition of young children. This is because there are few well-designed programs that: (i) took

---

2 See Glewwe and Miguel (2008), Strauss and Thomas (2008), and Currie (2009) for recent reviews or literature.
place long enough ago for children to have reached adulthood, (ii) for which panel data, including pre-intervention data and program take-up data, are available, and (iii) track the baseline sample over time to address attrition bias.

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. Drawing on a wide body of evidence from economics, psychology, and neuroscience, Heckman (2006) argues that returns to such investments are much higher than those made later in life. However, the empirical base for these arguments is not as deep as is often presumed. Evidence from developing countries that child health and nutrition or family planning matters for later life outcomes such as cognitive functioning and income, which our study will examine, is scarce and mixed (Walker et al., 2007). For example, a study in Jamaica which randomized 9–24 month olds to receive a nutritional supplement found a significant positive impact of the supplement on child development two years after the program, but no lasting effects when the children were ages 7-8 and 17-18 years old, and no effect on wages (Gertler et al. 2012, Grantham-McGregor et al. 1997, Walker et al. 2005). In contrast, results from an early childhood supplementation program run by the Guatemalan Institute for Nutrition in Central America and Panama (INCAP) which randomized four villages into two groups, one which received a high energy and protein drink and the other a placebo sugar drink, found that the exposed population experienced height and weight gains in childhood that were sustained in adolescence (Rivera et al. 1995), that children who were exposed to the intervention before age 3 experienced a quarter of a standard deviation increase in both reading comprehension and nonverbal cognitive tests when they were young adults (Maluccio et al. 2009), and that men exposed as children had higher hourly wages (Hoddinott et al. 2008).

In addition, while there is growing evidence from randomized interventions on the short-term and now medium-term effects of varied health, education, and CCT interventions on human capital attainment and even labor market outcomes (Baird et al. 2008), either the interventions started later than early childhood or they have not been going long enough to examine the longer-term effects on adult human capital.

2.2.3 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 MHSS1 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 0.2-0.25 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.
3. DESCRIPTION OF INTERVENTION AND THEORY OF CHANGE

3.1 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 and maternal and child health services. Services were rolled out over time starting with family planning. Box 1 gives a brief summary of intervention rollout, and more details are provided below. 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 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.

---

**Box 1: MCH-FP Program Interventions by Birth Year**

<table>
<thead>
<tr>
<th>Year of Birth</th>
<th>Age 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>Mar. 1982–Dec. 1988</td>
<td>22-29</td>
<td><strong>Child Health Interventions Added</strong></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. Services were added over time, so those in later cohorts had access to the earlier interventions. See a more detailed description of table in Barham (2012).

Program services were rolled out over time. Box 1 provides a general overview of the key time periods of roll out and birth cohorts affected by the interventions. The program started in October 1977, and 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).

3.2 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 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

---

3 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.
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 heterogeneous program effect by pre-intervention access to tubewell water, and results are similar if one drops the 10% of the population who are Hindus. 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.

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 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 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**

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

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

This report 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 2012-2015 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 icddr,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.

3.4. 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 the group of people analysed in this report, children born during the experimental period (October 1977 to December 1988). 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 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.

3.5 Outcomes

In this study we focus on the long-term impact of the MCH-FP program 35 years after it began, on two key measures of human capital: height and cognitive functioning. Cognitive functioning is measures using the Mini Mental State Exam (MMSE), Digit Span Forward and Backwards, and Raven’s Colored Progressive Matrices (Raven). The height and MMSE were collected both in the 1996 MHSS1 and in the 2012-15 MHSS2, the remaining outcomes only in MHSS2.

The cognitive tests measure slightly different domains of cognitive functions, with digit span focusing on memory, Ravens matrix reasoning, and MMSE multiple domains (see Appendix E for more details). Memory is thought to increase with intelligence and Ravens is used a non-verbal and simple measure of general intelligence and perhaps the most common and popular test
for people above the age of 5 (Kaplan and Saccuzzo 2009). With the exception of the MMSE, these measures are well known in the literature and were chosen because they are appropriate for both children and adults. None of the cognitive tests have been normed or validated for Bangladesh, so outcomes are internally standardized into z-scores using the comparison group. Measures of cognitive functioning vary with age. The analysis in this report compare the treatment and comparison areas for specific age ranges and all regressions include birth year fixed effects to control for any trends in the outcomes by age.

It was not possible to include all the tests in the phone survey for time and practicality reasons. Only self-report height and digit span forward were collected in the phone survey. 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 testing part of the survey (book 6) and did not do other parts of the survey. There were a few enumerators who were trained on the entire survey including book 6 to interview adult migrants who were living on their own.

An important question is whether catch-up growth in height is possible between MHSS1 and MHSS2 for the panel of people followed over time (children born during the experimental period). The majority of panel members were not yet old enough for their adolescent growth spurt in MHSS1. Research shows that catch-up growth in late childhood and adolescence is possible, but the literature is mixed as to whether interventions or major changes in the environment are needed to obtain this catch-up growth (Martorell, Khan and Schroeder, 1994; Adair, 1999; Prentice et al., 2013). As a result, it is uncertain whether this intervention should or should not affect height between MHSS1 and MHSS2 surveys.

4. Migrant Tracking Protocols and Attrition

Before providing the detailed summary, we outline the importance of tracking migrants for the overall aim of the study. Given the research questions and the potential (but unknown) relationship between cognitive development, nutrition, education and migration, it was paramount to our study to minimize attrition. Minimizing attrition in our setting was not only important to maximize statistical power, but more importantly to enhance both the internal and external validity of our research. As households and individuals that migrate are likely to have different characteristics than those that stay (as seen in the case of Oportunidades, Parker et al. 2008), the sample that would be surveyed had we not tracked migrants is unlikely to be representative of the entire initial target population. This has clear implications for the external validity. Moreover, migration was itself moderately affected by the treatment (Barham and Kuhn 2014). While these concerns are relevant for all evaluations, they are particularly pertinent for our study given the long period between collection of detailed data (16 years for MHSS1 respondents, up to 35 years for pre-MHSS1 migrants) and the focus on young individuals who tend to be particularly mobile (and hence have high potential attrition). Extensive tracking procedures were therefore a key component of our research design. As noted in Appendix A, we
targeted for follow up *all* members of households in the MHSS1 primary sample, referred to as panelists. In addition, we include most descendants of panelists, and a substantial number of spouses.

Figure 3 describes the extent of internal and overseas out-migration and the level of variation by sex and study cohort, focusing on respondents who were interviewed in MHSS1 and survived to MHSS2. For further detail on survivorship, see Appendix Section A.4 and Appendix Table A1. Overall, 73% of respondents lived in the MHDSS area or a nearby district, 16% were in Dhaka Metropolitan area, 4% were elsewhere in Bangladesh (including many in the second city of Chittagong but a large number scattered throughout the country), and 6% overseas. The overall distribution conceals substantial variation by sex and cohort. Among males in the prime study cohort born 1977-1988, only 42% were living in or near Matlab, with 26% in Dhaka, 7% elsewhere in Bangladesh, and 25% living abroad. Among women born 1977-1988, a relatively high share of 25% were living in Dhaka and 7% lived elsewhere in Bangladesh, but only 1% lived overseas, leaving 66% in or near Matlab. A far higher proportion of respondents in older cohorts were living in or near Matlab, including 92% of males and 94% of females born before 1959. Given the focus of our study on the most mobile age cohorts, and given the exceptionally high rate of international migration among males, it was essential to pursue an intensive and continually evolving approach to migrant follow-up.

Tracking of out-migrant respondents proceeded in four distinct phases.

*Phase 1 – Standard migrant tracking, December 2012-October 2013:* For the majority of migrants, migrant tracking was conducted as part of the Matlab survey fieldwork. From the MHDSS, we ascertained that 80% of migrants had a relative (parent, spouse, child, or sibling) living in a Matlab survey household. Information on these migrants was prepopulated to the Migrant Tracking survey module, in which interviewers collected information on migrant location and contact information. The 20% of migrants who did not have relatives in a Matlab household were assigned to specialized tracking coordinators, retired HDSS staff members who were uniquely familiar with the study area and able to enlist the help of MHDSS community welfare assistants to identify hard-to-track respondents such as those who had left many years earlier. These migrant data were checked and transferred to a migrant tracking coordinator. The migrant tracking coordinator assigned cases to teams assigned to particular regions of the country (mostly near Matlab and in Dhaka), and supervisors were responsible for providing regular updates on case progress and identifying cases that required additional tracking.

Migrant fieldwork was phased in gradually to allow the accumulation of a migrant caseload and the optimization of migrant survey field protocols. Migrant fieldwork began in the third month of the survey (December 2012). Operations were expanded significantly in the April 2013 and continued through October 2013 with significant interruptions due to political unrest throughout the country.
Figure 3: Location at MHSS2 of surviving MHSS1 respondents/descendants by sex and study cohort

**Males**
- Cases: 0, 500, 1,000, 1,500
- Overseas
- Rest of Bangladesh
- Dhaka Metro
- Matlab and nearby

**Females**
- Cases: 0, 500, 1,000, 1,500
- Overseas
- Rest of Bangladesh
- Dhaka Metro
- Matlab and nearby
**Phase 2 – Rapid Response, October 2013 – June 2014:** In October 2013, after almost all migrants had been through standard migrant tracking procedures, we had not been able to locate and interview a significant proportion of migrants. We therefore developed a new protocol for rapid response tracking. This protocol paired designated migrant trackers based in Matlab directly with migrant survey supervisors based in the areas where the migrants might be located in order to process cases quickly for interview. While some of the cases that could not be tracked in Phase 1 simply lacked sufficient tracking information, others had shown moderate resistance to being surveyed given their busy schedules. Typically, they would agree to an interview and would either schedule and break appointments, ask the interviewer to call back later, or ask the interviewer to come on any Friday (the typical day of rest) and then be unavailable. Our analysis and fieldworker experience suggested that any interview that had not been completed or at least firmly scheduled within one week of first contact with a migrant carried a high probability of failure. Migrants, especially those in Dhaka, were not only busy but they also had lost some of the enthusiasm for the work of icddr,b shown by Matlab-based respondents.

To address these challenges, the rapid response protocol focused on motivation and rapid mobilization. Instead of simply asking a Matlab-based informant for the migrant’s contact information, the informant was asked to call the migrant directly and explain that they had participated in this survey and that the migrant’s participation was also important. To solidify the connection and to ensure rapid follow-up, the migrant was then asked to expect a call from a migrant supervisor based in the destination area. The migrant tracker quickly assigned the case to an available supervisor. The destination-area supervisor was then expected to immediately phone the migrant, remind them that their kin had participated in the study, collect standard contact information (including additional phone numbers, best times to call, etc.) and schedule a firm time for an interview. The supervisor phoned the tracker to confirm that contact had been made and an interview had been set. Only after contact had been confirmed would the tracker be expected to leave the informant’s household, thereby ensuring a successful handover. If a migrant could not be reached at the time of the Matlab household visit, the tracker would collect contact information on the migrant and any additional Matlab or destination-based informants, and request the opportunity to return in case of further difficulty.

The rapid response protocol was rolled out in November 2013. In January 2014, further modifications were introduced, including 1) procedures for in-person tracking of migrants without phones, 2) assigning supervisors and field teams to specific zones within Dhaka Metro to avoid traveling through areas of high political risk or traffic, and 3) designating smaller male-only interview teams that could be more mobile and conduct interviews at night with male respondents.

**Phase 3 – Eid festival interviews, July 2014 and October 2014:** By March 2014, migrant survey operations had been reduced in size, but there remained a significant load of hard-to-track respondents and migrants living in remote destinations. At the same time, we became aware of the exceptionally high rate of overseas migration among key study cohorts. To address each of these challenges, we conducted rapid migrant tracking and survey operations oriented around the festivals of Eid-al-Fitr (July 2014) and Eid-al-Azha (October 2014). In the weeks leading up to each Eid, all remaining migrants were tracked, first by phone and then in person if necessary, beginning with international migrants who were on a holiday visit home (and who could not be interviewed at any other time) before moving to faraway migrants and other hard-to-track migrants.
The period leading up to festivals is also a time when migrants permanently return to their home villages or take extended holidays, especially international migrants. There were thus many opportunities to interview some migrants well before Eid, thereby reducing the substantial burden of cases that needed to be coordinated during the festival week itself. To facilitate this effort, the rapid response process, which was previously divided between trackers and interview teams, grew increasingly integrated. Supervisors visiting Matlab informant household would focus on completing an interview as soon as possible. Interviews were conducted immediately if the migrant had already returned. If a migrant was returning soon, an interview was schedule shortly after they returned. If a migrant was only coming to Dhaka and not to Matlab, interviews were reassigned to Dhaka teams. The two Eid operations were highly successful, allowing us to conduct in-person interviews with about 25% of all international migrants and substantially reducing the remaining burden of faraway and hard-to-track migrants.

**Phase 4 – International migrant phone survey, October 2014-December 2014.** As part of the Eid tracking process for international migrants, we also began to collect phone contact information for migrants who would not be returning home. It became clear that a large majority of migrants had phones and were willing to participate in the survey. In October 2014, we piloted a shortened version of the MHSS2 questionnaire for conducting phone surveys. Between October and December 2014 the phone survey was implemented. The phone survey focused exclusively on male migrants living without family abroad, who constituted the vast majority of international migrants. Out of 717 migrants contacted, 597 (83%) were successfully interviewed, with 70 cases who could not be tracked and another 70 who only communicated via internet phones at cyber cafes. While we successfully interviewed some internet phone users by asking them to call our field teams and offering financial compensation for phone costs, some had not phoned back before field operations were discontinued in December 2014.

Figure 4 demonstrates the considerable impact of phases 2 through 4 of the migrant tracking operation for sample attrition, particularly in the prime age cohorts born in 1977 to 1988. We present phase-by-phase completion rates as a percentage of total cases, with any remaining gap short of 100% indicating the ultimate rate of attrition. If the survey had been concluded after Phase 1 of migrant tracking, we would have interviewed 82% of females, with a range from 90% of those born before 1959 to just 73% of those born 1977-1988. For males, we would have interviewed only 71%, with a range from 89% of those born before 1959 to just 56% of those born 1977 to 1988.

Focusing on just the 1977-1988 cohort, the introduction of rapid response was essential to boosting response rates. Rapid response was particularly important for females, given the difficulty of contacting women in a traditional society without first building trust. Without rapid response, it was difficult to contact unmarried women at all, while contacting married women often raised suspicion among husbands, particularly those with no ties to Matlab and icddr,b. After rapid response, response rates in 1977-1988 cohort had risen for females from 73% in Phase 1 to 90% and for males from 56% to 69% for males. The Eid tracking phase was essential for improving response rates for males in particular and for females who were living far away from Matlab, with response rates in the 1977-1988 cohort rising to 93% for females and 77% for males. By the end of Eid tracking, response rates for internal migrants were broadly comparable across all age and sex groups and always significantly above 90%. The remaining male-female response gap was entirely explained by the 75% of overseas migrants who were not captured by Eid tracking. This gap was resolved by the phone survey, which increased the response rate for
males born 1977-1988 from 77% to 92% and for those born 1972-1976 from 81% to 94%. After completing all four phases of fieldwork, response rates were tightly clustered across age and sex groups in a range from 92% to 95%.

**Figure 4:** Phase of MHSS2 interview completion for surviving MHSS1 respondents, by sex and study cohort
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. In contrast to many other longer-term studies, we tracked all households and targeted individuals, rather than a random subsample, to both increase statistical power and better capture heterogeneity that might be related to differential migration by destination.

5. 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. In the future, we will also examine how differential selection from attrition is affecting the results.

5.1 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 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.

5.2 Treatment Group Balance and Attrition Balance

We test whether the areas are similar using preintervention household and household head characteristics from the 1974 census. Table 1a 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.

Table 1b examines attrition balance using 1974 baseline characteristics. We examine the attrition balance for the full sample used in this study as well as for the 22-29 year olds. In addition, we examine attrition balance for respondents with and without phone survey information. The difference in means gives the difference in the baseline characteristics between those who were not interviewed and those who were interviewed. For this report, we indicate that someone was not interviewed if they have no height information. Height was collected in the phone survey. So the first three columns examine the balance between those we could not be interview and those we could for our sample, where those with height reported in the phone survey were coded as being followed. The remaining columns code the phone survey responses as being people who were not followed. Across the two samples and regardless if the person was or was not interviewed in the phone survey the baseline characteristics are well balanced across those who we could or could not interview. Gender is not well balanced due to males being more likely to migrate.

5.3 Main Results

The ITT effects are presented in Table 2 for the three main age groups (see Box 1 for more details on the age group or Appendix D). 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 (35-65 year olds) are used to make the double difference. Since differences in the 35-65 year old group could be due to differential attrition between the treatment and comparison group, 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. Alternative, the small negative effect may be due to selective mortality if frailer children in this age group in the treatment area survived as a result of the program.

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 an exact panel of people. The effects on height are similar between the two rounds of the survey with 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 22-29 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.

Mortality Selection.— Even if the MCH-FP program were truly randomized, the program itself is likely to cause mortality to differ between treatment and comparison areas over time, potentially biasing the results. Mortality selection is likely to bias the results downwards since frailer individuals (or those with lower health endowments) are more likely to survive in the treatment area due to the intervention, leaving a higher probability of observing someone who is shorter or with a lower level of cognitive functioning in the treatment than in the comparison area in the follow-up period. Indeed, the negative effect for the 30-34 year old group may result from mortality selection from late measles vaccination.

Spillover Effects.—The ITT effects may also be biased by the program’s indirect effects on nonparticipants: informational spillovers about the program, and the positive externalities of some of the interventions, such as vaccinations, leading to a downward bias of the program effect. In both of these cases, spillovers are more likely to occur in areas that border or are relatively close to the treated villages, since knowledge about the programs is likely to be spread by word-of-mouth, and the positive externalities of vaccination are largely local.

Changes Over Time.— At present the analysis does not control for other changes that took place in the follow-up period. Three important changes were introduction of a river embankment for

---

4 Height and MMSE 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.
irrigation and flood protection in 1987, introduction of micro-credit and other programs through BRAC in the 1990s, and increased levels of education in Bangladesh. The embankment cut through both the treatment and comparison areas, and BRAC was introduced in a cross-over design with the MCH-FP program. The BRAC program eventually covered most villages, but it was at least not placed all in the treatment or all in the comparison area to start. Education levels rose in both the treatment and comparison areas. So the comparison area will be controlling at least partially for many of these changes over time. Future research will include controls for these types of changes over time.

5.4 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.

Table 3 disaggregates the 22-29 year olds into Treatment area 1 and 2. For height, in 2012/15, the program effect in treatment area 1 remained the same at 0.26 SD, but fell to 0.14 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.

5.5 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 the 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 22-29 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.

---

5 Phone survey respondents were not included in the sample for height and digit span forward in order to follow a similar sample across outcomes. Results are similar when phone survey respondents are included.
5.6 Internal and External Validity

The internal validity of the results is based on the quasi-experimental variation in the data discussed in section 3.2. Previous research using data prior to the start of the program shows the quasi-experiment resulted in a balanced sample with the exception of access to tubewell water and religion. With the intensive tracking and the low attrition rates, the balance is expected remained even after 35 years, so the internal validity is strong.

To understand the external validity of the findings, a first key consideration is the comparability of the MCH-FP program and the similarity of Matlab and of Bangladesh to other countries with poor health care systems. In terms of the program itself, the MCH-FP program eventually became the template for the Bangladesh national MCH-FP program, which was rolled out between 1986 and 1988 (Cleland et al. 1994). In modified form, this approach was adopted in a wide range of countries seeing substantial MCH-FP success like Ghana and Ethiopia (Phillips et. 2003). Thus the mix of short-term impacts on family size, child survival and maternal survival is globally relevant. To the extent that Bangladesh was uniquely proactive in incorporating the Matlab MCH-FP service package into the national program and the government service area caught up to the treatment area relatively quickly, this may serve to bias our program treatment effects coefficients downward.

A more external validity concern relates to the context in which long-term programmatic impacts took place. In this respect, every country will have substantial variations in the social, ecological, economic, and institutional conditions governing the pathway from health to long-term change. Beyond stating that more such long-term studies need to conducted, among which Matlab will be among the longest-running and thus more valuable, the study affords us the unique opportunity to model the very pathways of social and economic opportunity that might mediate the pathway from health to development. Most notably, as we set out in our proposals, we have two opportunities to understand how new economic opportunities may have amplified the effects of MCH-FP. In 1991, the Bangladesh Rural Advancement Committee and icddr,b introduced a quasi-randomized microcredit experiment, with treatment and control villages introduced cross-wise into the MCH-FP treatment and comparison area. In 1987, the Bangladesh Water and Power Development Board introduced a river embankment that has been associated with significant changes in cropping, wealth, and marriage (Mobarak et al. 2013). We have tracked population-level exposure to these interventions through the MHSS2 community survey, household reports of program participation, and are also collecting program roll out data from the agencies themselves.

The other notable changes in the economic opportunity structures related to widespread international guest worker migration by males and internal migration by females to a growing readymade garment sectors. The importance of migration for health and well-being (Kuhn 2005, 2006; Kuhn et al. 2011) and the potential for differences in the practice of migration between treatment and comparison area (Barham and Kuhn 2014) motivated us to carry out such extensive follow-up of out-migrants.

Armed with a large sample size and an exhaustive representation of these key confounding factors, we will be able to produce policy-relevant research on the specific secondary factors that impede or accelerate the path from health to development. Thus while the Bangladesh experience may not be externally valid in all respects, we will have an opportunity to show specifically how it may be valid.
5 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 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. Determining the causal effects of early child-health interventions on later human capital formation is challenging due to a lack longitudinal data. 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 MMSE fades out with the comparison group catching-up over time. 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. Difference in the program effects on height and cognition may be due to the nutrition content of the food though, unfortunately, such analysis this is beyond the scope of this report.

It is not clear why differential effect between treatment and comparison group in the MMSE at age ages 8 -14 did not persist into adulthood. Below we discuss some possible reasons, but it will be difficult to determine the exact mechanism. For example, it could be the program itself. While the majority of the comparison group children were too old to really benefit from late vaccination when it became available, they could have benefited from more resources in the family due to increases in family planning in the comparison area. However, the lack of results for the 30-34 who mainly benefited from family planning makes this possibility less likely. 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, sustained gains in cognition over time may require earlier investments, such as in-utero, than were provided by the MCH-FP program interventions. Alternatively, more sustained health or childhood interventions may be required throughout a longer period of the child’s life.

Despite lack or persistence in the differential cognitive effect over time, improvements in other areas, such as labor market outcomes, may be possible due to the advantages the treatment area had in human capital while children and early teens. Future research will investigate this possibility.

Overall, the results highlight that the MCH-FP program led to important and sustained effects on a person’s height through adulthood. There were important effects of cognitive functioning
through at least late childhood, that did not persist through adulthood. Future research needs to investigate if these effects on cognitive functioning through late childhood still led to gains in in other aspects of these people’s lives.
REFERENCES


TABLE AND FIGURES:

Table 1a: Treatment/Comparison Group Balance - 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>Obs</td>
</tr>
<tr>
<td>Family size</td>
<td>6.95</td>
<td>5.98</td>
<td>2768</td>
</tr>
<tr>
<td>Owns a lamp (=1)</td>
<td>0.64</td>
<td>1.37</td>
<td>2768</td>
</tr>
<tr>
<td>Owns a watch (=1)</td>
<td>0.16</td>
<td>0.86</td>
<td>2768</td>
</tr>
<tr>
<td>Owns a radio (=1)</td>
<td>0.08</td>
<td>0.59</td>
<td>2768</td>
</tr>
<tr>
<td>Wall tin or tinmix (=1)</td>
<td>0.31</td>
<td>1.04</td>
<td>2768</td>
</tr>
<tr>
<td>Tin roof (=1)</td>
<td>0.83</td>
<td>0.74</td>
<td>2768</td>
</tr>
<tr>
<td>Latrine (=1)</td>
<td>0.82</td>
<td>1.00</td>
<td>2768</td>
</tr>
<tr>
<td>Number of rooms per capita</td>
<td>0.22</td>
<td>0.18</td>
<td>2768</td>
</tr>
<tr>
<td>Number of cows</td>
<td>1.50</td>
<td>3.80</td>
<td>2768</td>
</tr>
<tr>
<td>Number of boats</td>
<td>0.66</td>
<td>1.62</td>
<td>2768</td>
</tr>
<tr>
<td>Drinking water, tubewell (=1)</td>
<td>0.31</td>
<td>1.45</td>
<td>2768</td>
</tr>
<tr>
<td>Drinking water, tank (=1)</td>
<td>0.39</td>
<td>1.97</td>
<td>2768</td>
</tr>
<tr>
<td>Drinking water, other (=1)</td>
<td>0.30</td>
<td>2.50</td>
<td>2768</td>
</tr>
<tr>
<td>HH age</td>
<td>47.86</td>
<td>26.87</td>
<td>2768</td>
</tr>
<tr>
<td>HH years of education (edu.)</td>
<td>2.45</td>
<td>6.52</td>
<td>2768</td>
</tr>
<tr>
<td>HH works in agriculture (=1)</td>
<td>0.60</td>
<td>1.13</td>
<td>2768</td>
</tr>
<tr>
<td>HH works in fishing (=1)</td>
<td>0.06</td>
<td>0.66</td>
<td>2768</td>
</tr>
<tr>
<td>HH spouse's age</td>
<td>37.13</td>
<td>21.79</td>
<td>2768</td>
</tr>
<tr>
<td>HH spouse's years of edu.</td>
<td>1.09</td>
<td>3.39</td>
<td>2768</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 1b: Attrition Balance - Difference in 1974 Baseline Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Full Sample Phone Survey as Non-Attrition</th>
<th>Full Sample Phone Survey as Attrition</th>
<th>22-29 year olds Phone Survey as Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diff in Means</td>
<td>T-Stat</td>
<td>Mean/SD</td>
</tr>
<tr>
<td>Male (=1)</td>
<td>-0.09</td>
<td>-5.49</td>
<td>-0.22</td>
</tr>
<tr>
<td>Islamic (=1)</td>
<td>-0.02</td>
<td>-1.45</td>
<td>-0.02</td>
</tr>
<tr>
<td>Family size</td>
<td>0.08</td>
<td>0.66</td>
<td>0.01</td>
</tr>
<tr>
<td>Owns a lamp (=1)</td>
<td>0.00</td>
<td>-0.17</td>
<td>0.00</td>
</tr>
<tr>
<td>Owns a watch (=1)</td>
<td>0.01</td>
<td>0.63</td>
<td>0.01</td>
</tr>
<tr>
<td>Owns a radio (=1)</td>
<td>-0.02</td>
<td>-2.91</td>
<td>-0.04</td>
</tr>
<tr>
<td>Wall tin or tinmix (=1)</td>
<td>0.00</td>
<td>0.09</td>
<td>0.00</td>
</tr>
<tr>
<td>Tin roof (=1)</td>
<td>-0.02</td>
<td>-1.00</td>
<td>-0.02</td>
</tr>
<tr>
<td>Latrine (=1)</td>
<td>-0.01</td>
<td>-0.94</td>
<td>-0.01</td>
</tr>
<tr>
<td>Number of rooms per capita</td>
<td>0.00</td>
<td>-1.10</td>
<td>-0.02</td>
</tr>
<tr>
<td>Number of cows</td>
<td>-0.12</td>
<td>-1.43</td>
<td>-0.03</td>
</tr>
<tr>
<td>Number of boats</td>
<td>-0.02</td>
<td>-0.87</td>
<td>-0.01</td>
</tr>
<tr>
<td>Drinking water, tubewell (=1)</td>
<td>0.01</td>
<td>0.73</td>
<td>0.01</td>
</tr>
<tr>
<td>Drinking water, tank (=1)</td>
<td>-0.02</td>
<td>-0.79</td>
<td>-0.01</td>
</tr>
<tr>
<td>Drinking water, other (=1)</td>
<td>0.00</td>
<td>0.14</td>
<td>0.00</td>
</tr>
<tr>
<td>HH age</td>
<td>-0.70</td>
<td>-1.34</td>
<td>-0.03</td>
</tr>
<tr>
<td>HH years of education (edu.)</td>
<td>0.20</td>
<td>1.58</td>
<td>0.03</td>
</tr>
<tr>
<td>HH works in agriculture (=1)</td>
<td>-0.01</td>
<td>-0.44</td>
<td>-0.01</td>
</tr>
<tr>
<td>HH works in fishing (=1)</td>
<td>0.01</td>
<td>0.56</td>
<td>0.01</td>
</tr>
<tr>
<td>HH spouse's age</td>
<td>-0.28</td>
<td>-0.69</td>
<td>-0.01</td>
</tr>
<tr>
<td>HH spouse's years of edu.</td>
<td>0.09</td>
<td>1.15</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Notes: The difference in means is between those who were not-followed and those who were followed. Sample is for the primary book 35 respondent from MHSS1. If a panel member is missing height information they are counted as not followed. Respondents who reported height in the phone survey are counted as being followed in the first 3 columns but as not being followed in the remaining columns.
Table 2: Intent-To-Treat Program Effects by Age Group

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>MMSE</th>
<th>Digit Span</th>
<th>Digit Span</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
</tr>
<tr>
<td></td>
<td>(11)</td>
<td>(12)</td>
<td>(13)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel A: DD results

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>MMSE</th>
<th>Digit Span</th>
<th>Digit Span</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Area (=1)</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.01</td>
<td>-0.02</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td></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>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.09)</td>
<td>(0.09)</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>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
</tr>
</tbody>
</table>

Panel B: SD results

<table>
<thead>
<tr>
<th></th>
<th>Height</th>
<th>MMSE</th>
<th>Digit Span</th>
<th>Digit Span</th>
<th>Raven</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
</tr>
<tr>
<td></td>
<td>(11)</td>
<td>(12)</td>
<td>(13)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sample

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1996 Primary Sample</td>
<td>2012-15 Follow-up Sample</td>
<td>Outcome not missing in 1996</td>
<td>Outcome not missing in 2015</td>
<td>Phone survey included</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Observations: 5,724 5,092 5,092 5,917 5,630 4,675 4,000 4,000 5,625 5,914 5,619 5,622 5,587

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, 2015

<table>
<thead>
<tr>
<th></th>
<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>
<td>(5)</td>
</tr>
<tr>
<td>Treatment Area</td>
<td>0.00</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Treatment Area *(Age 22–26)</td>
<td>0.26*</td>
<td>-0.00</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Treatment Area 1 *(Age 27–29)</td>
<td>0.26**</td>
<td>-0.08</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Treatment Area 2 *(Age 27–29)</td>
<td>0.14</td>
<td>-0.10</td>
<td>-0.04</td>
<td>-0.06</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.06)</td>
<td>(0.10)</td>
<td>(0.08)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Treatment Area*(Age 30–34)</td>
<td>0.00</td>
<td>0.04</td>
<td>-0.08</td>
<td>-0.09</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>2012-15 Follow-up Sample</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>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Outcome not missing in 2012-15</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>5,630</td>
<td>5,625</td>
<td>5,619</td>
<td>5,623</td>
<td>5,588</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.
Table 4: Intent-to-Treat Program Effect Heterogeneity 2015

<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>
<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>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area (=1)</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>0.03</td>
<td>0.03</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>
<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>(0.04)</td>
<td>(0.04)</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>
<td>(0.06)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)</td>
<td>0.23*</td>
<td>0.27+</td>
<td>0.10</td>
<td>-0.00</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0.01</td>
<td>-0.03</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>(0.15)</td>
<td>(0.14)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.10)</td>
<td>(0.08)</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>Treatment Area*(Age 30–34)</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>-0.14+</td>
<td>-0.14+</td>
<td>-0.14+</td>
<td>-0.09</td>
<td>-0.09</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>
<td>0.10</td>
</tr>
<tr>
<td></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>(0.08)</td>
<td>(0.09)</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>
<td>(0.10)</td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)* Tubewell drinking water in 1974</td>
<td>-0.10</td>
<td>-0.10</td>
<td>-0.10</td>
<td>0.07</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.07)</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>
<td>(0.11)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)* HHH works in agriculture 1974</td>
<td>-0.10</td>
<td>-0.00</td>
<td>0.04</td>
<td>0.11</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.04</td>
<td>-0.03</td>
<td>0.04</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</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>(0.11)</td>
<td>(0.11)</td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment Area*(Age 22–29)* HHH has some education 1974</td>
<td>0.12</td>
<td>0.07</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>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.07)</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>
<td>(0.12)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>5,630</td>
<td>5,630</td>
<td>5,630</td>
<td>5,625</td>
<td>5,625</td>
<td>5,625</td>
<td>5,623</td>
<td>5,623</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>
<td>5,588</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, HH=household head. Use 2012-15 follow-up sample from column 5 of Table 2. Phone survey respondents not included but results are similar when they are included. Outcomes are internally standardized z-scores.
Table 5: Intent-to-Treat Program Effect by Subcomponent of MMSE (z-scores), 2015

<table>
<thead>
<tr>
<th>Panel A: DD results</th>
<th>Orientation</th>
<th>Attention-Concentration</th>
<th>Recall</th>
<th>Registration</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<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>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: SD results</th>
<th>Orientation</th>
<th>Attention-Concentration</th>
<th>Recall</th>
<th>Registration</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<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>
</tbody>
</table>

Observations 5,625 5,625 5,625 5,625 5,625
APPENDICES:

A. Sample design
B. Survey instruments
C. Power calculations
D. Study/design and methods
E. Description of Mini Mental State Exam
APPENDIX A: Sample Design

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

A.1. Constructing a multigenerational panel

Conducted in 1996, the first Matlab Health and Socioeconomic Survey (MHSS1) was collected on a random sample of the 1993 population of the MHDSS area. At the time, no efforts were taken to select or weight this sample in accordance with the distribution of the MHDSS area population in 1974, a period prior to the actual program start when an MHDSS census exists. Migrants who left the MHDSS area prior to 1996 were not 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 out-migration, and to error-bound existing analysis to account for differential mortality and out-migration between treatment and comparison areas. But this work could not look directly at out-migrants who had left the area 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 from 1974 to the present, allowing us to identify our potential sample in advance, include migrants who departed prior to MHSS1, and prepopulate key identifying characteristics to survey books to improve tracking of out-migrants.

The sample for MHSS1 was based on a probability sample of 1993 households (the Primary Sample). For each household a related nearby household was included in the Secondary Sample. MHSS2 follows only Primary Sample households. While all members of the households are listed and basic information provided on them, only a subsample of household members were interviewed. The tracing forward of the MHSS1 sample began with the 11,165 individual respondents to MHSS1, defined as those who responded either to the Adult Interview (Book III) or the Child Interview (Book V). (See Appendix B for fuller description of the survey instruments.) To account for the absence of migrants departing between 1977 and 1996, we used MHDSS to select an additional sample of pre-MHSS1 migrants from the Primary Sample households who were not initially included in MHSS1. Given budget constraints, we focused our efforts on including all children born to MHSS1 respondents during the program years 1977 to 1988. These “pre-MHSS1 migrants” totaled 120. We refer to the MHSS1 respondents and pre-MHSS1 migrants as “panelists.” They were 11,285 in number. Beginning from the set of panelists, we will describe our follow-up 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 of MHSS1. 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 1996

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 descendants to any location within Bangladesh and conducted a phone survey for males living 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 coresiding 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 prepopulation 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 prepopulation on 31 December, 2011. In the remaining 75% of MHSS2 Matlab households and in all migrant households, we include non-target household members in the household listing and account 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, prepopulation of questionnaires was carried out using the information available in MHSS1 and the continuously collected MHDSS. Although prepopulation 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 prepopulation 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 prepopulated to a Matlab household who had subsequently migrated had to be shifted into the migrant tracking system.

3) Unprepopulated 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 prepopulation. As part of the MHSS2 individual interview, adult respondents were asked to report on all children, including those who were or weren’t prepopulated, in the Individual Control Book. Interviewers were trained to add any children born after 1996 who were not prepopulated (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 unprepopulated 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 prepopulated 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.
A.4. Description of the target population

Table A1 describes the distribution of this target population by sample status and study cohort. A total of 36,238 respondents were traced, of whom 2,282 died before MHSS2, yielding a total sample of 33,956. This included 9,198 surviving panelists (out of 10,791 total, a 15% loss to mortality). It also included 19,027 surviving descendants, primarily including 12,465 born since 1996 along with a fair number who were alive but not included in MHSS1. The sample also included 5,731 spouses of panelists of descendants.

Table A1: Distribution of cases traced forward from MHSS1 to MHSS2, by MHSS1 sample descent group, study cohort, and survivorship

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panelist</td>
<td>4,011</td>
<td>1,798</td>
<td>622</td>
<td>2,838</td>
<td>1,522</td>
<td>10,791</td>
</tr>
<tr>
<td>Descendant</td>
<td>254</td>
<td>1,142</td>
<td>896</td>
<td>2,357</td>
<td>2,091</td>
<td>12,915</td>
</tr>
<tr>
<td>Spouse</td>
<td>236</td>
<td>991</td>
<td>950</td>
<td>2,861</td>
<td>737</td>
<td>5,792</td>
</tr>
<tr>
<td>Total</td>
<td>4,501</td>
<td>3,931</td>
<td>2,468</td>
<td>8,056</td>
<td>4,350</td>
<td>36,238</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panelist</td>
<td>2,588</td>
<td>1,724</td>
<td>610</td>
<td>2,787</td>
<td>1,489</td>
<td>9,198</td>
</tr>
<tr>
<td>Descendant</td>
<td>238</td>
<td>1,108</td>
<td>880</td>
<td>2,317</td>
<td>2,019</td>
<td>12,465</td>
</tr>
<tr>
<td>Spouse</td>
<td>214</td>
<td>974</td>
<td>941</td>
<td>2,848</td>
<td>737</td>
<td>5,731</td>
</tr>
<tr>
<td>Total</td>
<td>3,040</td>
<td>3,806</td>
<td>2,431</td>
<td>7,952</td>
<td>4,245</td>
<td>12,482</td>
</tr>
</tbody>
</table>

39
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.

3. **Book 1: Household 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 1 collects information on household characteristics, water resources and well switching, and consumption.

4. **Book 2: Household Economy 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 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.

6. **Book 4: Woman Book**
   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 Screening 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 (i.e.- permanent) shops", and in "supermarkets" when supermarkets were available. The market surveys were repeatedly collected in 5 locations in Matlab and then multiple times in Dhaka and in most migrant areas.
APPENDIX C: Power Calculations

Power calculations were done separately for each age group of interest. For example, the MHSS2 sample includes approximately 4,000 aged 22-29 and 2,300 aged 30-34 index respondents. We are fortunate to be able to estimate intra-cluster correlations (ICC) for sample size calculations based on similar populations within Bangladesh using MHSS1 data, since they are preferable to using estimates from other data sources reported in the literature (Duflo & Glennerster 2008). In addition, for cognitive tests, we use similar data from Nicaragua to help determine appropriate ICCs. We use a Matlab village as the cluster. For the two age groups discussed above, using a power of 0.9 and ICC of 0.5 for income, we can detect effect sizes of 15\% for men and 20\% for women for income changes if we interview close to 85\% of the sample. Using ICC between 0.02 and 0.05 for cognitive functioning, we can detect effect sizes of between 0.15-0.25 standard deviations. Our ICC for cognitive functioning is low compared to many studies of academic test scores. Our cluster is the village rather than the school; the type of information gathered in the cognitive tests is often not “taught” in school, so a lower ICC is consistent with greater variation across people living in the same village.
APPENDIX D: Study Design and Methods

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 as summarized in Box 1. 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. Nevertheless, in the mechanisms section below we describe some analysis that may be suggestive of the role of the child health interventions.

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 height 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. To account for the differences in baseline characteristics, we include the observables in the regression.

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:

\[
(1) \ C_{iv} = \beta_0 + \beta_1(T_v * AG_{i}^{22-29}) + \beta_2(T_v * AG_{i}^{30-34}) + \beta_3(T_v * AG_{i}^{35-65}) + \alpha_{by} + X'Z + \varepsilon_{iv}
\]

where \( C \) is one of the outcome measures such as height or cognitive functioning, \( T_v \) 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. \( AG^V_i \) 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 eligible to receive the most intensive health interventions. \( \beta_2 \) 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. \( \alpha_{by} \) 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. We investigate if this older cohort benefited, say due to improved health environment or remittances; if they did the double difference should be an under-estimate of the program effect. The double-difference estimator can be determined using Equation 2.

\[
(2) \ C_{iv} = \beta_0 + \beta_1 T_v + \beta_2(T_v * AG_{i}^{22-29}) + \beta_3(T_v * AG_{i}^{30-34}) + \alpha_{by} + X'Z + \varepsilon_{iv}
\]

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.

---

\( p \) We can consider matching models, however, the overlap and distribution of the pre-intervention characteristics between the two experimental areas are so similar that, in previous research (Barham, 2012), peer reviewers pointed out the models were not additive to the papers.
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.

APPENDIX E: Description of Mini Mental State Exam

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 from MHSS1. 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. Birth year fixed-effects are included in the regression analysis in order to control for this association between age and the MMSE score.