

Tania Barham
Randall Kuhn
Jane Menken
Abudur Razzaque

Thirty-five years later

Evaluating the impacts of a child health and family planning programme in Bangladesh

January 2016

Impact
Evaluation
Report 39

Health



International
Initiative for
Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding and producing high-quality evidence of what works, how, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people's lives.

3ie Impact Evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of this report, *Thirty-five years later: evaluating effects of a quasi-random child health and family planning programme in Bangladesh*, in partial fulfilment of the requirements of grant OW3.1060 issued under Open Window 3. This study is now being published as *Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh*, 3ie Impact Evaluation Report 39. The content has been copy-edited and formatted for publication by 3ie. Due to unavoidable constraints at the time of publication, a few of the tables or figures may be less than optimal. All of the 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 also 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 at <http://www.3ieimpact.org/en/about/3ie-affiliates/3ie-members/>

Suggested citation: Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016. *Thirty-five years later: evaluating effects of a quasi-random child health and family planning programme in Bangladesh*, 3ie Impact Evaluation Report 39. New Delhi: International Initiative for Impact Evaluation (3ie).

3ie Impact Evaluation Report Series executive editors: Jyotsna Puri and Beryl Leach

Managing editor: Omita Goyal and Deepthy Menon

Assistant managing editor: Paromita Mukhopadhyay and Tara Kaul

Production manager: Pradeep Singh

Copy editor: Aruna Ramachandran

Proofreader: Mathew PJ

Cover design: John F McGill

Cover photo: Laura Elizabeth Pohl/Bread for the World

Printer: Via Interactive

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh

Tania Barham

Department of Economics and Institute of Behavioral Science, University of Colorado

Randall Kuhn

Josef Korbel School of International Studies, University of Denver

Jane Menken

Institute of Behavioral Science, University of Colorado

Abudur Razzaque

icddr,b

3ie Impact Evaluation Report 39
January 2016



Acknowledgments

Research discussed in this report has been funded by the International Initiative for Impact Evaluation (3ie) through the Global Development Network (GDN); the National Institute on Aging (R01 AG033713), and the Population Reference Bureau (*Thirty-five 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 help with the data including providing demographic surveillance data for historic information and drawing samples, help with design of the 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 fieldwork, data preparation, and numerous contributions to this research project.

Abstract

Improving the health and nutrition of young children is important not only for immediate well-being, but also because it is believed to reduce poverty in the long-run through improved human capital. Many programs such as Head Start and conditional cash transfer 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, e.g., through the spread of poverty reduction programs like conditional cash transfers. 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 which were implemented in 1977. 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 in terms of 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 that the MCH-FP program led to important improvement in human capital in early and late childhood (ages 8–14). This study examines whether these effects persist when the same sample of people reach the age group 22–29. We find that while the effect on height persists into adulthood, the effects on cognition do not. The difference in results between height and cognition highlight that physical growth and cognitive development may be affected differently by the environment and do not necessarily follow each other.

Contents

Acknowledgments	i
Abstract	ii
List of figures and tables	iv
Abbreviations and acronyms	v
1. Introduction	1
2. Context	2
2.1 Background of intervention.....	2
2.2 Related literature	3
3. Description of intervention and theory of change	5
3.1 The MCH-FP program.....	5
3.2 Program targeting and take-up.....	7
3.3 Data and evaluation surveys	8
3.4 Theory of change mechanisms linking the MCH-FP program and human capital accumulation	9
3.5 Outcomes.....	10
4. Migrant tracking protocols and attrition	11
4.1 Tracking of outmigrant respondents proceeded in four distinct phases.....	12
5. Impact results	18
5.1 Treatment variable	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	22
5.6 Internal and external validity.....	22
6. Conclusions and policy recommendations	23
Appendix A: Sample design	31
Appendix C: Power calculations	38
Appendix D: Study design and methods	39
Appendix E: Description of MMSE	41
References	42

List of figures and tables

Figure 1: Map of Matlab study area	3
Figure 2: Trends in CPR and measles vaccination rate for children 12–59 months by calendar year	8
Figure 3: Location at MHSS2 of surviving MHSS1 respondents/descendants by sex and study cohort.....	13
Figure 4: Phase of MHSS2 interview completion for surviving MHSS1 respondents, by sex and study cohort	18
Table 1a: Treatment/comparison group balance–1974 baseline characteristics	25
Table 1b: Attrition balance–difference in 1974 baseline characteristics	26
Table 2: ITT program effects by age group	27
Table 3: ITT effects disaggregated in the treatment area, 2015.....	28
Table 4: ITT program effect heterogeneity 2015	29
Table 5: ITT program effect by subcomponent of MMSE (z-scores), 2015	30
Table A1: Distribution of cases traced forward from MHSS1 to MHSS2, by MHSS1 sample descent group, study cohort, and survivorship	34

Abbreviations and acronyms

BAMSE	Bangla Adaptation of the MMSE
BRAC	Bangladesh Rural Advancement Committee
CCT	conditional cash transfer
CPR	contraceptive prevalence rate
DD	double difference
GPS	global positioning system
ICC	intracluster correlation
ITT	intent-to-treat
MCH-FP	Matlab Maternal and Child Health and Family Planning
MHDSS	Matlab Health and Demographic Surveillance System
MHSS	Matlab Health and Social Survey
MMSE	mini mental state exam
sd	standard deviation
SD	single difference

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 ultimately through labor market opportunities for the child (Heckman 2006; Strauss and Thomas 2008). 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, Thomas, and Currie 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 the 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 preintervention data, and long duration of follow-up (over 35 years).

In 1977, icddr,b (formerly the International Centre 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 the 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 its effect on those who were born during the experimental period in terms of human capital in adulthood as measured by height and

cognitive functioning. In particular, we use single difference (SD) and double difference (DD) models to examine whether improvements in human capital at ages 8-14 for those born during the experimental period persist when these same people reach the age range 22-29.

Given the long-term nature of this study, it was of paramount importance 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 who 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. Thirty-five 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 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 literature, and 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

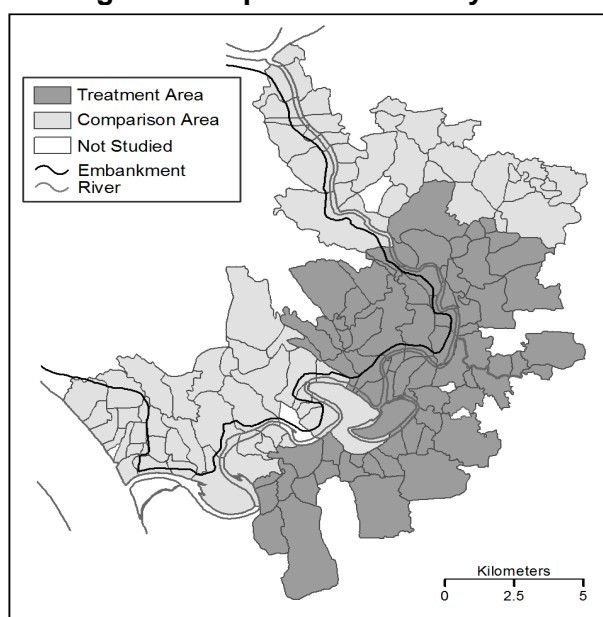
In the early 1960s, icddr,b began the Matlab Health and Demographic Surveillance System (MHDSS) that provided monthly data on fertility, mortality, marriage and divorce, change of residence within the study site, and destination of outmigrants from the area for a population of approximately 200,000 people residing in a geographically defined area (Fauveau 1994). After the disruption caused by 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. Data collection under MHDSS 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 *pardah* 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 diarrheal and acute respiratory illness (see Section 3.1 for more details

¹ Such tracking was the primary objective and mode of use of the funding provided by 3ie, with the main data collection supported by the National Institutes of Aging.

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 were not available regularly in government clinics before 1988.

In 1996, a group of researchers (including many members on this 3ie-funded study) led the collection of data for Matlab Health and Social Survey (MHSS) 1. The survey led to a number of important findings on medium-term impacts of the MCH-FP interventions, discussed later. In 2012, a follow-up survey (support by this 3ie grant, the 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 years after the start of the program.

Figure 1: Map of Matlab study area



2.2 Related literature

Each year, more than 7 million children die from preventable diseases (Hill, You, and Inoue 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. This latter may have consequences later in life, such as on completing education, marriage, and labor market outcomes (Fogel 1994; Martorell, Khan, and Schroeder 1994; Martorell 1995, 1999; Haas *et al.* 1996; Case, Fertig, and Paxson 2005). In fact, it is estimated that over 200 million children under age 5 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 (Ahlburg 1998; Strauss and Thomas 1998; Bloom and Canning 2001; Sachs and Malaney 2002; Bloom, Canning, and Jamison 2004; Almond and Currie 2011). Yet, few studies investigate the 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, lead 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 place long enough in the past for children to have reached adulthood; (ii) for which panel data, including preintervention 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 returns to 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 aged 7-8 and 17-18, and no effect on wages (Grantham-McGregor *et al.* 1997; Walker *et al.* 2005; Gertler *et al.* 2012). In contrast, results from an early childhood supplementation program run by the Guatemalan Institute for Nutrition in Central America and Panama that 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 (sd) 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).

² See Glewwe and Miguel (2008), Strauss and Thomas (2008), and Currie (2009) for recent reviews of literature.

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, Hamory, and Miguel 2008), either the interventions started later than early childhood or they have not been going on long enough to examine the longer-term effects on adult human capital.

2.2.2 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; Koenig, Fauveau, and Wojtyniak 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) find an increase in schooling for boys aged 9–14, but no effect for girls. Chaudhuri (2005) reports that girls younger than 14 experienced improved weight-for-age and boys were significantly less stunted. Barham (2012) finds that the MCH-FP program led to a 0.39 sd increase in cognitive functioning and a 0.2–0.25 sd increase in height and years of education attained for children aged 8–14. Trapp and 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 its 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 later in this report. 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 and 1988 to evaluate the program. Evaluation of the program is also aided by the rolling out of the program services over two main periods: 1977–1981 and 1982–1988.

Program services were rolled out over time. Box 1 provides a general overview of the key time periods of rollout and the birth cohorts affected by the interventions. The program started in October 1977, and prior to 1982 focused on family planning and

Box 1: MCH-FP program interventions by birth year

Age 2012 Program Eligibility		
Oct. 1977– Feb. 1982	30–34	<i>Intensive family planning and maternal health interventions</i> Mother eligible for family planning, tetanus toxoid vaccine, folic acid, and iron in last trimester of pregnancy
Mar. 1982– Dec. 1988	22–29	<i>Child health interventions added</i>
Mar. 1982– Oct. 1985	27–29	<i>Interventions added in half the treatment area</i> Children under age 5 eligible for measles vaccination in half the treatment area
Nov. 1985– Dec. 1988	22–26	<i>Intervention extended to entire treatment area</i> Children under age 5 eligible in entire treatment area for vaccination (measles, DPT, polio, tuberculosis), vitamin A supplementation, nutrition rehabilitation for children at risk (starting 1987)

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).

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 1994).

Between 1982 and 1988, the types of interventions provided were expanded, especially for children under age 5. These interventions were rolled out over time in the treatment area, starting with the measles vaccine in half the treatment area in 1982. Starting in 1985, preventive services were provided to children under the age of 5 in the entire treatment area. These services included vaccines for measles, DPT, polio, 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 (icddr,b 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 (Cleland *et al.* 1994; Phillips *et al.* 2003). 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 8,000 in 1987–1988 to 1 per 5,000 in 1989–1990 (Cleland *et al.* 1994). The ratio was still lower in the treatment area at 1 per 1,300 in 1990.³ 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

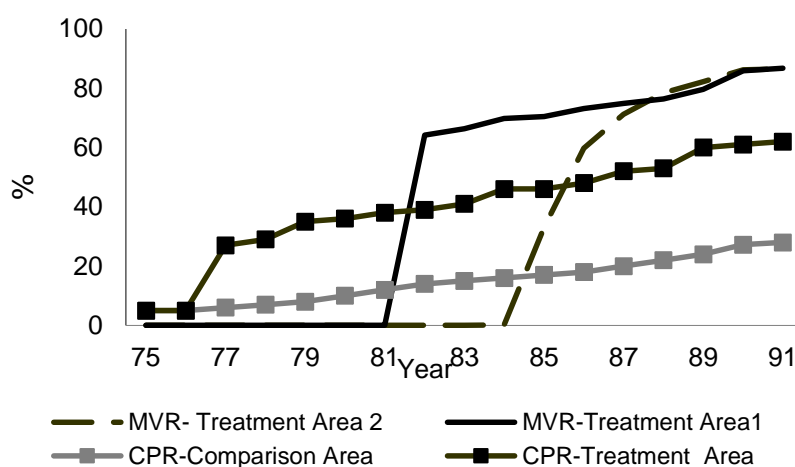
The comparison group was built into the design of MCH-FP (Fauveau 1994), but the treatment and comparison villages were not assigned randomly. Instead, the treatment and comparison areas are contiguous geographic areas (Figure 1) that were viewed as socially and economically similar and geographically insulated from outside influences at the time (Phillips *et al.* 1982). Treatment status was assigned to geographically contiguous areas to reduce potential contamination of the comparison area from the family planning interventions (Huber and Khan 1979). This 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 and Phillips 1990; Joshi and 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-intervention household and household head characteristics were also similar (Barham 2012; Barham and Kuhn 2014). Two exceptions were religion and access to tubewell water. The difference in access to tubewell water is a result of a government program, and 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 effects by pre-intervention access to tubewell water, and the results are similar if one drops the 10 percent 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.

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.

Program implementation followed the planned timeline and uptake was rapid. For example, Figure 2 shows that the contraceptive prevalence rate (CPR) for married women aged 15–49 was similarly low prior to the program (< 6 percent) in both areas. There was a large increase in the CPR to 30 percent in the treatment area during the first year of the project. After that, there was a steady increase in the CPR, which reached almost 50 percent 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 percent in 1988. There was still a 20 percent 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 percent 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). Figures for vaccination rates in the comparison area do not exist, but are believed to be near 0. 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, Fauveau, and Wojtyniak 1991). Nationally, measles vaccination for children under the age of 5 was less than 2 percent in 1986 (Khan and Yoder 1998), and was below 40 percent in the comparison area in 1990 (Fauveau 1994). As the national program scaled up, these differences narrowed substantially.

Figure 2: Trends in CPR and measles vaccination rate for children 12–59 months by calendar year



Source: Contraceptive use data from van Ginneken *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 the ongoing MHDSS.

The more recent survey, the 2012–2015 Matlab Health and Social Survey 2 (MHSS2), was conducted 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, a follow-up study to the 1996 Matlab Health and Socioeconomic Survey (MHSS1) is publicly available from the Rand website (<http://www.rand.org/labor/FLS/MHSS.html>) and Interuniversity Consortium for Political and Social Research (<http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/2705>). Both MHSS1 and MHSS2. 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. However, MHSS2 has a richer array of outcomes, including enhanced cognitive and health tests, more detailed consumption and employment modules, 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 iccdr,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 iccdr,b on the entire study population of Matlab, and is used to help create an intent-to-treat (ITT) 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 analyzed in this report — children born during the experiment period (October 1977 to December 1988). All the mothers of the 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 varied. 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 (October 1977–February 1982) and those born after the rollout (March 1982–December 1988).

The effects of the MCH-FP interventions on health and nutrition are understood better than the effects 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, von König, and Heininger 2005; Reingold and 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 (Reddy 1987; Greenberg, von König, and Heininger 2005).

Second, the child health interventions may also have indirect effects due to 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 healthcare). 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 5 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 tradeoff, 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 and 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 5 (Grantham-McGregor *et al.* 2007; Barham, Macours, and Maluccio 2013). For this reason, we hypothesize that those who were born during the experimental period in the treatment area compared to the comparison area would 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 improved health, nutrition, and cognitive functioning, they would also do better in school, and thus attain higher grades. Based on these mechanisms, we also hypothesize that the MCH-FP program would lead to better human capital outcomes for those born when the child health interventions were available (i.e., March 1982–December 1988) than those born when these 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 measured using the mini mental state exam (MMSE), digit span forward and backwards, and Raven's colored progressive matrices (Raven). Data on height and MMSE were collected both in the 1996 MHSS1 and in the 2012–2015 MHSS2, while the remaining outcomes were collected only in MHSS2.

The cognitive tests measure slightly different domains of cognitive function, with digit span focusing on memory, Raven's matrix on reasoning, and MMSE on multiple domains (see Appendix E for more details). Memory is thought to increase with intelligence, and Raven is used as 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 compares 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 conducted on the individuals in their homes 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 responsible only for implementing the testing part of the survey (Book 6) and did not conduct other parts of the survey. There were a few enumerators who were trained in 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 experiment 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 inconclusive 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 did or did not affect height between the 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 of paramount importance 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; see Parker, Rubalcava, and Teruel 2008), the sample that would have been surveyed had we not tracked migrants

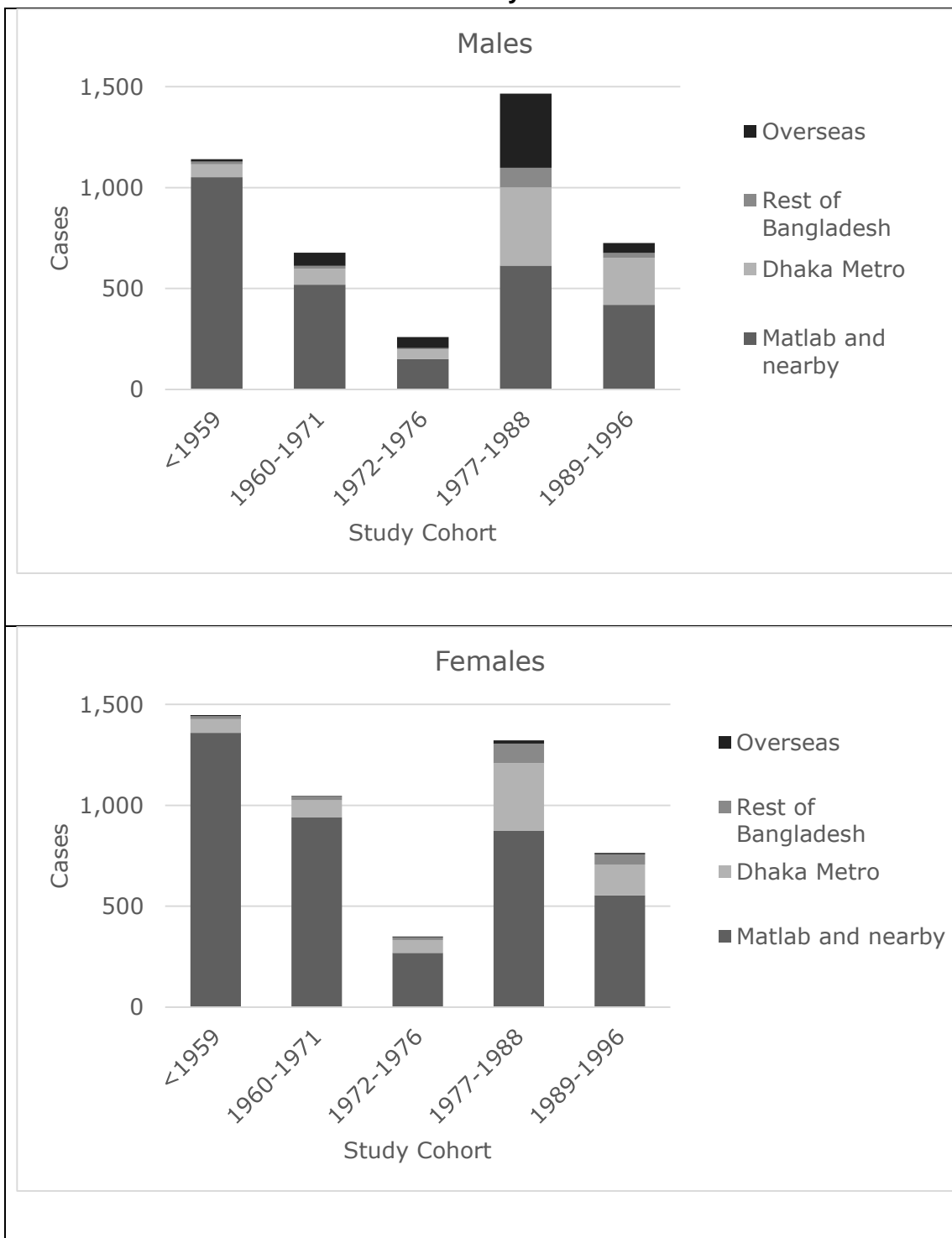
is unlikely to be representative of the entire initial target population. This has clear implications for the external validity. Moreover, migration was itself moderately affected by the treatment (Barham and Kuhn 2014). While these concerns are relevant for all evaluations, they are particularly pertinent 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 outmigration and the level of variation by sex and study cohort, focusing on respondents who were interviewed in MHSS1 and survived to MHSS2. For further details on survivorship, see Appendix Section A.4 and Appendix Table A1. Overall, 73 percent of respondents lived in the MHDSS area or a nearby district, 16 percent were in the Dhaka Metropolitan Area, 4 percent were elsewhere in Bangladesh (including many in the second city of Chittagong, but a large number were scattered throughout the country), and 6 percent overseas. The overall distribution conceals substantial variation by sex and cohort. Among males in the prime study cohort born during 1977–1988, only 42 percent were living in or near Matlab, with 26 percent in Dhaka, 7 percent elsewhere in Bangladesh, and 25 percent living abroad. Among women born during 1977–1988, a relatively high share of 25 percent were living in Dhaka and 7 percent lived elsewhere in Bangladesh, but only 1 percent lived overseas, leaving 66 percent in or near Matlab. A far higher proportion of respondents in older cohorts were living in or near Matlab, including 92 percent of males and 94 percent 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.

4.1 Tracking of outmigrant 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 percent 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 percent 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.

Figure 3: Location at MHSS2 of surviving MHSS1 respondents/descendants by sex and study cohort



These migrant data were checked and transferred to a migrant tracking coordinator. The migrant tracking coordinator assigned cases to teams allocated in 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 case load 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 April 2013 and continued through October 2013 with significant interruptions due to political unrest throughout the country.

Phase 2 – Rapid response, October 2013 – June 2014: In October 2013, after almost all migrants had been through standard migrant tracking procedures, we were not 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 also had lost some of the enthusiasm for the work of icddr, 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, and so on), 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 (i) procedures for in-person tracking of migrants without phones; (ii) assigning supervisors and field teams to specific zones within Dhaka Metro to avoid traveling through areas of high political risk or traffic; and (iii) designating smaller male-only interview teams that could maintain greater mobility 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 households 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 scheduled shortly after they returned. If a migrant was coming only 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 percent of all international migrants and substantially reducing the remaining burden of faraway and hard-to-track migrants.

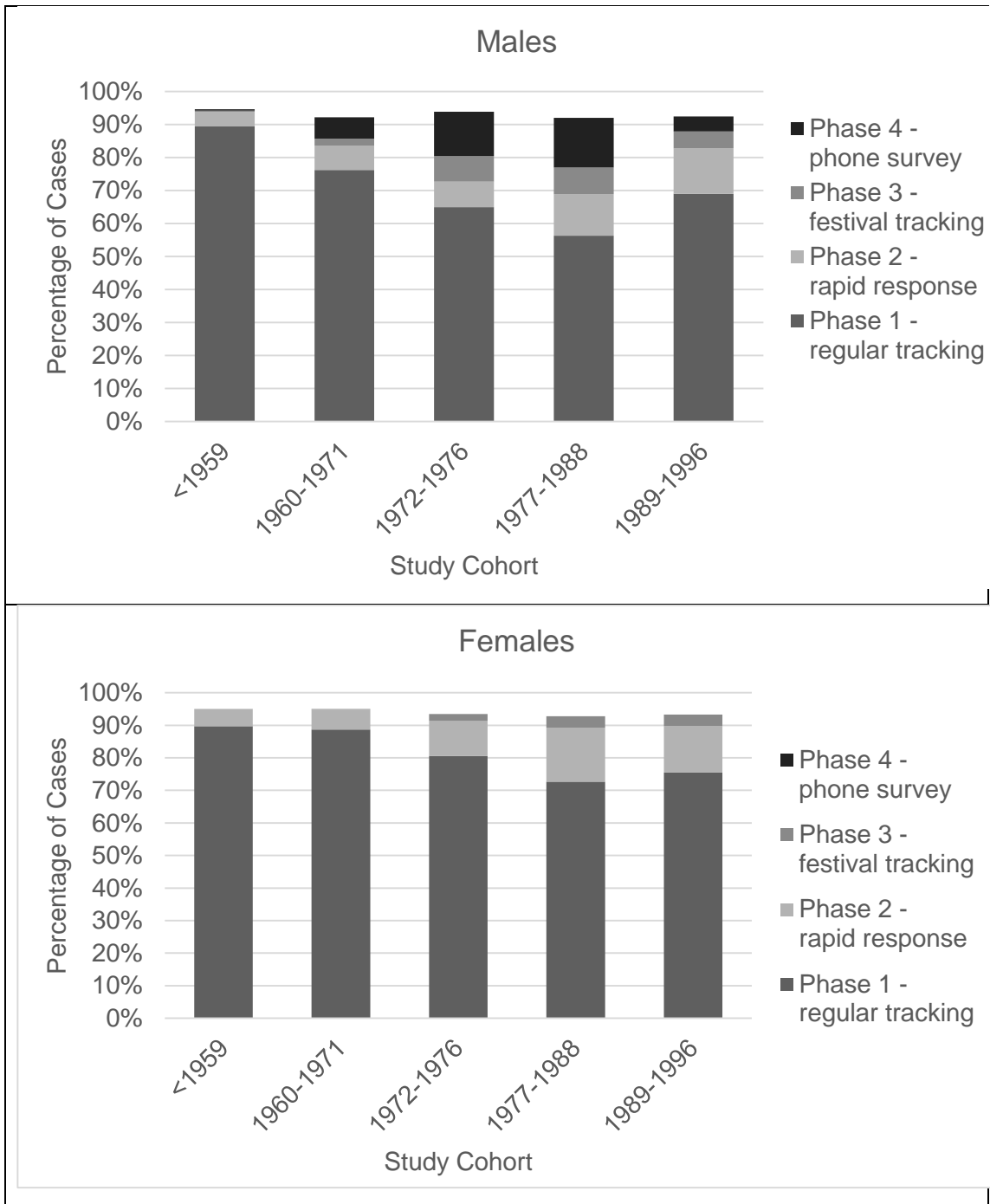
Phase 4 – International migrant phone survey, October 2014–December 2014: As part of the Eid tracking process for international migrants, we also began to collect phone contact information for migrants who would not be returning home. It became clear that a large majority of migrants had phones and were willing to participate in the survey. In October 2014, we piloted a shortened version of the MHSS2 questionnaire for conducting phone surveys, and it was implemented between October and December 2014. 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 percent) were successfully interviewed, with 70 cases who could not be tracked and another 70 who communicated only by means of 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 between 1977 and 1988. We present phase-by-phase completion rates as a percentage of total cases, with any remaining gap short of 100 percent indicating the ultimate rate of attrition. If the survey had been concluded after Phase 1 of migrant tracking, we would have interviewed 82 percent of females, with a range from 90 percent of those born before 1959 to just 73 percent of those born during 1977–1988. For males, we would have interviewed only 71 percent, with a range from 89 percent of those born before 1959 to just 56 percent of those born between 1977 and 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 the 1977–1988 cohort had risen for females from 73 percent in Phase 1 to 90 percent, and for males from 56 percent to 69 percent. 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 percent for females and 77 percent 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 percent. The remaining male–female response gap was entirely explained by the 75 percent 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 during 1977–1988 from 77 percent to 92 percent and for those born during 1972–1976 from 81 percent to 94 percent. After completing all four phases of fieldwork, response rates were tightly clustered across age and sex groups in a range from 92 percent to 95 percent.

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.

Figure 4: Phase of MHSS2 interview completion for surviving MHSS1 respondents, by sex and study cohort



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 the fact that the data have only recently become 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 microcredit program by the Bangladesh Rural Advancement Committee (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 ITT 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 sd of the characteristics for the treatment and comparison areas for the entire sample. The differences in means are statistically insignificant at the 5 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 (2009) argue that 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, a 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 percentage 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.* 2004). 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 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 had no height information. Height was collected in the phone survey. So the first three columns examine the balance between those whom we could not interview and those whom we could interview 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 of whether the person was or was not interviewed in the phone survey, the baseline characteristics are well balanced across those whom 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 or Appendix D for more details on age group): 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 DD. Since differences in the 35–65-year-old group could be due to differential attrition between the treatment and comparison groups, SD results are presented in panel B.

Table 2 presents results for five outcomes: height, MMSE, digit span forward, digit span backwards, and Raven. 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 reported in Barham (2012), but have been remade for this report. In 1996, improvements in height and MMSE were experienced only 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 sd difference between the treatment and comparison area in height, and a 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 age group. For the 30–34 age group (aged 15–19 in 1996), the point estimates are fairly small but significant at the 10 percent level for MMSE in the SD model. Any positive effect that the family planning and maternal health interventions may have had on cognitive development in this age group may have been swamped by sibling competition from younger siblings who received more intensive child health interventions. Alternatively, 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 columns 7 and 8 for MMSE).⁴ 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 for the two rounds of the survey, with point estimates of 0.23 sd in 1996 and 0.19 sd in the 2012/15 survey for the 20–29-year-olds. However, the program effects on the MMSE for the 20–29-year-olds do not persist between the two rounds of the survey, dropping from 0.33 sd 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 the 2012/15 survey regardless of whether 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 Raven. Results are similar for these three measures of cognition. The pattern of results is 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

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.

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 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 irrigation and flood protection in 1987, introduction of microcredit 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 at least it was not placed entirely in the treatment or entirely in the comparison area to start with. Education levels rose in both treatment and comparison areas. So the comparison area will control 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 5 in half the treatment area (treatment area 1) were eligible to receive the measles vaccine starting in March 1982, and those 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 were eligible only past the recommended age. The 22–25-year-olds in both treatment areas 1 and 2 were eligible at the same time, so there are 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 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 age group.

Table 3 disaggregates the 22–29-year-olds into treatment areas 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 areas 1 and 2.

5.5 Program effect heterogeneity

Table 4 examines if the program effect for the 22–29 age group differs by preintervention access to tubewell water, household head education, and whether 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 areas. Interaction effects of the DD estimator for children aged 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 but is relatively small, and all interactions are statistically insignificant. Heterogeneous effects for the 22–29 age group by education of the household head and whether the household head worked in agriculture are also statistically insignificant.⁵ Thus, there do not appear to be heterogeneous treatment effects, at least with respect to these three important variables.

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 that 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 to have remained even after 35 years, so the internal validity is strong.

To understand the external validity of the findings, a key consideration is the comparability of the MCH-FP program, and the similarity of Matlab and of Bangladesh to other countries with poor healthcare 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 al.* 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 that the government service area caught up to the treatment area relatively quickly, this may serve to bias our program treatment effects coefficients downward.

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.

A greater concern with respect to external validity 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 be conducted, among which Matlab is one of the longest-running and thus more valuable studies, 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, BRAC 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, Kuhn, and Peters 2013). We have tracked population-level exposure to these interventions through the MHSS2 community survey and household reports of program participation, and are also collecting program rollout data from the agencies themselves.

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

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.

6. 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 CCTs 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 of 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, Thomas, and Currie 2002).

Previous research shows that the MCH-FP program led to important improvement in human capital in early and late childhood (ages 8–14). This study examines whether these effects persist when these same people are aged 22–29. We find that while the effect on height persists, the effects on MMSE fade 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 ages 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 environment. Difference in the program effects on height and cognition may be due to the nutrition content of the food, though, unfortunately, such analysis is beyond the scope of this report.

It is not clear why the differential effect between treatment and comparison groups in the MMSE at ages 8–14 did not persist into adulthood. We will discuss some possible reasons, but it is 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 age group 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 2 (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 at the in utero stage, 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 have been possible due to the advantages the treatment area had in terms of human capital during childhood and the 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 on 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 other aspects of these people's lives.

Tables

Table 1a: Treatment/comparison group balance—1974 baseline characteristics

	Treatment Area			Comparison Area			Difference in Means		
	Mean	sd	Obs	Mean	sd	Obs	Mean	T-stat	Mean/sd
Family size	6.95	5.98	2,768	6.76	4.54	3,165	0.19	1.39	0.04
Owns a lamp (= 1)	0.64	1.37	2,768	0.61	1.09	3,165	0.03	0.98	0.03
Owns a watch (= 1)	0.16	0.86	2,768	0.16	0.66	3,165	0.00	0.06	0.00
Owns a radio (= 1)	0.08	0.59	2,768	0.08	0.50	3,165	0.00	-0.03	0.00
Wall tin or tinmix (= 1)	0.31	1.04	2,768	0.31	0.93	3,165	0.00	0.03	0.00
Tin roof (= 1)	0.83	0.74	2,768	0.84	0.81	3,165	0.00	-0.13	0.00
Latrine (= 1)	0.82	1.00	2,768	0.86	0.92	3,165	-0.05	-1.88	-0.05
Number of rooms per capita	0.22	0.18	2,768	0.21	0.17	3,165	0.00	0.83	0.02
Number of cows	1.50	3.80	2,768	1.39	3.46	3,165	0.11	1.15	0.03
Number of boats	0.66	1.62	2,768	0.69	1.67	3,165	-0.02	-0.57	-0.01
Drinking water, tubewell (= 1)	0.31	1.45	2,768	0.16	1.06	3,165	0.15	4.52	0.14
Drinking water, tank (= 1)	0.39	1.97	2,768	0.33	1.85	3,165	0.06	1.18	0.03
Drinking water, other (= 1)	0.30	2.50	2,768	0.50	2.05	3,165	-0.21	-3.49	-0.10
HH age	47.86	26.87	2,768	46.20	25.24	3,165	1.66	2.45	0.07
HH years of education (edu.)	2.45	6.52	2,768	2.31	5.49	3,165	0.14	1.02	0.02
HH works in agriculture (= 1)	0.60	1.13	2,768	0.59	1.09	3,165	0.01	0.45	0.01
HH works in fishing (= 1)	0.06	0.66	2,768	0.07	0.52	3,165	-0.01	-0.50	-0.01
HH spouse's age	37.13	21.79	2,768	35.93	19.16	3,165	1.20	2.24	0.06
HH spouse's years of edu.	1.09	3.39	2,768	1.21	2.93	3,165	-0.11	-1.48	-0.04

Notes: Standard deviations 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

	Full Sample			Full Sample			22–29-year-olds		
	Phone Survey as Non-attrition			Phone Survey as Attrition			Phone Survey as Attrition		
	Diff. in Means	T-Stat	Mean/sd	Diff. in Means	T-Stat	Mean/sd	Diff. in Means	T-Stat	Mean/sd
Male (= 1)	-0.09	-5.49	-0.22	-0.23	-16.08	-0.59	0.54	40.90	1.06
Islamic (= 1)	-0.02	-1.45	-0.02	0.00	-0.26	0.00	0.02	0.69	0.03
Family size	0.08	0.66	0.01	0.07	0.75	0.01	0.44	1.60	0.13
Owens a lamp (= 1)	0.00	-0.17	0.00	0.01	0.39	0.00	0.11	2.64	0.16
Owens a watch (= 1)	0.01	0.63	0.01	0.01	0.65	0.01	0.02	0.51	0.05
Owens a radio (= 1)	-0.02	-2.91	-0.04	-0.01	-1.06	-0.02	0.03	1.01	0.09
Wall tin or tinmix (= 1)	0.00	0.09	0.00	0.01	0.80	0.01	0.03	0.52	0.05
Tin roof (= 1)	-0.02	-1.00	-0.02	0.00	-0.22	0.00	0.00	0.06	0.00
Latrine (= 1)	-0.01	-0.94	-0.01	0.00	0.03	0.00	0.04	1.12	0.06
Number of rooms per capita	0.00	-1.10	-0.02	0.00	-0.66	-0.01	-0.01	-0.67	-0.04
Number of cows	-0.12	-1.43	-0.03	-0.01	-0.08	0.00	0.19	1.07	0.08
Number of boats	-0.02	-0.87	-0.01	-0.01	-0.21	0.00	0.07	1.30	0.07
Drinking water, tubewell (= 1)	0.01	0.73	0.01	0.00	-0.19	0.00	0.00	0.11	0.00
Drinking water, tank (= 1)	-0.02	-0.79	-0.01	0.00	-0.06	0.00	-0.02	-0.41	-0.01
Drinking water, other (= 1)	0.00	0.14	0.00	0.00	0.20	0.00	0.01	0.28	0.01
HH age	-0.70	-1.34	-0.03	-0.75	-1.64	-0.03	-0.12	-0.09	-0.01
HH years of education (edu.)	0.20	1.58	0.03	0.09	0.91	0.01	-0.05	-0.22	-0.01
HH works in agriculture (= 1)	-0.01	-0.44	-0.01	0.01	0.57	0.01	0.05	1.19	0.08
HH works in fishing (= 1)	0.01	0.56	0.01	0.00	-0.25	0.00	-0.02	-0.86	-0.04
HH spouse's age	-0.28	-0.69	-0.01	-0.23	-0.62	-0.01	0.34	0.32	0.02
HH spouse's years of edu.	0.09	1.15	0.03	0.07	1.09	0.02	-0.07	-0.63	-0.04

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: ITT program effects by age group

	Height					MMSE				Digit Span Forward		Digit Span Backwards	Raven
	1996	1996	2015	2015	2015	1996	1996	2015	2015	2015	2015	2015	2015
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: DD results													
Treatment Area (= 1)	-0.05 (0.05)	-0.06 (0.05)	-0.01 (0.06)	-0.02 (0.05)	-0.02 (0.05)	-0.04 (0.04)	-0.03 (0.05)	0.03 (0.05)	0.04 (0.05)	0.02 (0.05)	0.03 (0.05)	0.02 (0.04)	-0.02 (0.05)
Treatment Area* (Age 22–29)	0.22* (0.09)	0.23* (0.09)	0.19+ (0.10)	0.20* (0.09)	0.21* (0.09)	0.38* (0.15)	0.33* (0.16)	-0.03 (0.06)	-0.04 (0.05)	0.02 (0.06)	0.02 (0.06)	-0.00 (0.06)	0.01 (0.07)
Treatment Area* (Age 30–34)	0.06 (0.10)	0.07 (0.10)	-0.06 (0.08)	-0.00 (0.07)	0.00 (0.08)	-0.12 (0.08)	-0.13 (0.09)	-0.13 (0.09)	-0.14+ (0.08)	-0.08 (0.08)	-0.08 (0.09)	-0.09 (0.08)	0.10 (0.10)
Panel B: sd results													
Treatment Area* (Age 35–65)	-0.05 (0.05)	-0.06 (0.05)	-0.01 (0.06)	-0.02 (0.05)	-0.02 (0.05)	-0.04 (0.04)	-0.03 (0.05)	0.03 (0.05)	0.04 (0.05)	0.02 (0.05)	0.03 (0.05)	0.02 (0.04)	-0.02 (0.05)
Treatment Area* (Age 22–29)	0.17** (0.06)	0.17** (0.06)	0.18+ (0.10)	0.18+ (0.09)	0.20* (0.10)	0.35* (0.15)	0.30+ (0.17)	0.00 (0.06)	0.00 (0.03)	0.04 (0.06)	0.05 (0.05)	0.02 (0.06)	-0.01 (0.04)
Treatment Area* (Age 30–34)	0.02 (0.08)	0.01 (0.08)	-0.06 (0.06)	-0.02 (0.06)	-0.01 (0.06)	-0.15+ (0.09)	-0.16+ (0.09)	-0.10 (0.07)	-0.10 (0.07)	-0.06 (0.08)	-0.05 (0.08)	-0.07 (0.08)	0.08 (0.09)
Sample													
1996 Primary Sample	Y	Y	N	N	N	Y	Y	N	N	N	N	N	N
2012/15 Follow-up Sample	N	N	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y
Outcome not missing in 1996	Y	Y	Y	N	N	Y	Y	Y	N	N	N	N	N
Outcome not missing in 2015	N	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N
Phone survey included	N	N	Y	Y	N	N	N	N	N	Y	N	N	N
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 0 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. Pre-intervention characteristics included DD = double difference, SD = single difference. Outcomes are internally standardized z-scores. Age is as of 2012.

Table 3: ITT effects disaggregated in the treatment area, 2015

	Height	MMSE	Digit Span Forward	Digit Span Backwards	Raven
	(1)	(2)	(3)	(4)	(5)
Treatment Area	0.00 (0.05)	0.04 (0.05)	0.03 (0.05)	0.03 (0.04)	-0.01 (0.06)
Treatment Area* (Age 22–26)	0.26* (0.12)	-0.00 (0.06)	0.03 (0.08)	0.03 (0.08)	0.03 (0.07)
Treatment Area 1* (Age 27–29)	0.26** (0.09)	-0.08 (0.06)	0.02 (0.11)	-0.04 (0.11)	-0.00 (0.11)
Treatment Area 2* (Age 27–29)	0.14 (0.11)	-0.10 (0.06)	-0.04 (0.10)	-0.06 (0.08)	-0.01 (0.09)
Treatment Area* (Age 30–34)	0.00 (0.05)	0.04 (0.05)	-0.08 (0.09)	-0.09 (0.08)	0.10 (0.10)
2012/15 Follow-up Sample	Y	Y	Y	Y	Y
Outcome not missing in 1996	N	N	N	N	N
Outcome not missing in 2012/15	Y	Y	Y	Y	Y
Observations	5,630	5,625	5,619	5,623	5,588

Notes: Standard errors are clustered at the village level. "***", "**", or "+" indicates that the difference in the coefficient from 0 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. Pre-intervention characteristics included DD = double difference, SD = single difference. Outcomes are internally standardized z-scores.

Table 4: ITT program effect heterogeneity 2015

	Height			MMSE			Digit Span Forward			Digit Span Backward			Raven		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Panel A: DD Results															
Treatment Area (= 1)	-0.02 (0.05)	-0.02 (0.05)	-0.02 (0.05)	0.04 (0.05)	0.04 (0.05)	0.04 (0.05)	0.03 (0.04)	0.03 (0.04)	0.03 (0.04)	0.03 (0.04)	0.03 (0.04)	0.03 (0.04)	-0.02 (0.06)	-0.02 (0.06)	-0.02 (0.06)
Treatment Area* (Age 22–29)	0.23* (0.11)	0.27+ (0.15)	0.10 (0.14)	-0.00 (0.05)	-0.04 (0.05)	-0.01 (0.07)	0.01 (0.07)	-0.03 (0.10)	0.05 (0.08)	0.01 (0.07)	-0.03 (0.10)	0.05 (0.08)	-0.01 (0.07)	0.02 (0.10)	0.01 (0.11)
Treatment Area* (Age 30–34)	0.01 (0.08)	0.00 (0.08)	0.01 (0.08)	-0.14+ (0.08)	-0.14+ (0.08)	-0.14+ (0.08)	-0.09 (0.08)	-0.09 (0.09)	-0.09 (0.09)	-0.09 (0.08)	-0.09 (0.09)	-0.09 (0.09)	0.10 (0.10)	0.10 (0.10)	0.10 (0.10)
Treatment Area* (Age 22–29)* Tubewell drinking water in 1974	-0.10 (0.17)			-0.10 (0.07)			-0.07 (0.13)			-0.07 (0.13)			0.13 (0.11)		
Treatment Area* (Age 22–29)* HHH works in agriculture 1974		-0.10 (0.17)			-0.00 (0.05)			0.04 (0.11)			0.04 (0.11)			-0.03 (0.11)	
Treatment Area* (Age 22–29)* HHH has some education 1974			0.12 (0.17)			0.07 (0.07)			-0.08 (0.11)			-0.08 (0.11)			-0.01 (0.12)
Observations	5,630	5,630	5,630	5,625	5,625	5,625	5,623	5,623	5,623	5,623	5,623	5,623	5,588	5,588	5,588

Notes: Standard errors are clustered at the village level. "****", "***", or "+" indicates that the difference in the coefficient from 0 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. Pre-intervention characteristics included: DD = double difference, SD = single difference, HH = household head. Uses 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: ITT program effect by subcomponent of MMSE (z-scores), 2015

	Orientation	Attention-Concentration	Recall	Registration	Language
Panel A: DD results					
Treatment Area (= 1)	0.04 (0.04)	-0.02 (0.05)	0.07 (0.05)	0.04 (0.05)	-0.01 (0.04)
Treatment Area* (Age 22–29)	-0.05 (0.05)	0.04 (0.05)	-0.04 (0.06)	-0.03 (0.05)	-0.01 (0.05)
Treatment Area* (Age 30–34)	-0.18** (0.07)	-0.02 (0.08)	-0.08 (0.08)	-0.10 (0.08)	-0.06 (0.07)
Panel B: SD results					
Treatment Area* (Age 35–65)	0.04 (0.04)	-0.02 (0.05)	0.07 (0.05)	0.04 (0.05)	-0.01 (0.04)
Treatment Area* (Age 22–29)	-0.01 (0.03)	0.03 (0.03)	0.03 (0.04)	0.01 (0.04)	-0.02 (0.03)
Treatment Area* (Age 30–34)	-0.13** (0.05)	-0.04 (0.06)	-0.01 (0.06)	-0.06 (0.08)	-0.08 (0.06)
Observations	5,625	5,625	5,625	5,625	5,625

Notes: Standard errors are clustered at the village level. "***", "**", or "+" indicates that the difference in the coefficient from 0 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. Pre-intervention characteristics included: DD = double difference, SD = single difference, HH = household head. Uses 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.

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 program implementation in detail (see Section 3), but rather use this section to detail the careful procedures we used to (i) improve our sample to construct a sample of beneficiaries and descendants suitable to conduct multigenerational analysis; (ii) improve tracking of the large number of outmigrants from the study area; and (iii) use MHDSS and 1974 census data to better account for program exposure, attrition, and household formation.

A.1. Constructing a multigenerational panel

Conducted in 1996, MHSS1 data were collected on a random sample of the 1993 population of the MHDSS area. At the time, no efforts were made to select or weight this sample in accordance with the distribution of the MHDSS area population in 1974, a period prior to the start of the actual program 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 outmigration, and to error-bound existing analysis to account for differential mortality and outmigration between treatment and comparison areas. But this work could not focus directly at outmigrants 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 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 outmigrants.

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 a 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 co-residence.
- (ii) All other children of panelists who co-resided 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, and so on 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 co-residing with a panelist or descendant.
- (ii) For all marriages in which a descendant was in the MHDSS area and the spouse was outside MHDSS, we tracked the spouse as a migrant.
- (iii) For all marriages in which a descendant was a migrant and the spouse was in MHDSS, we interviewed the spouse living in MHDSS.

If the descendant was outside the MHDSS area and the 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 percent of Matlab MHSS2 households at the time of sample prepopulation on December 31, 2011. In the remaining 75 percent 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 interviewed only descendants and spouses.

Finally, to better represent the population of MHDSS in 2012, we included a 7 percent 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 previously, 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) Un-prepopulated descendants: A small number of children born to descendants after 1996 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 were not 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 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 the Matlab sample fieldwork was concluded, residence and descent information was updated in the computerized database. Any descendants who should have been interviewed or tracked as descendants were assigned for re-interview or migrant tracking.

- 2) To ensure tracking of un-prepopulated descendants, we searched the listings of children in the individual control book child (CH) section for any children who should have been followed but were 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 subsequently listed by a migrant parent. This problem is relatively common when a father migrated before marriage and thus was not linked in the 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 percent 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

Traced							
Descent group	<1959	1960-1971	1972-1976	1977-1988	1989-1996	1997-	Total
Panelist	4,011	1,798	622	2,838	1,522	0	10,791
Descendant	254	1,142	896	2,357	2,091	12,915	19,655
Spouse	236	991	950	2,861	737	17	5,792
Total	4,501	3,931	2,468	8,056	4,350	12,932	36,238

Survived							
Descent group	<1959	1960-1971	1972-1976	1977-1988	1989-1996	1997-	Total
Panelist	2,588	1,724	610	2,787	1,489	0	9,198
Descendant	238	1,108	880	2,317	2,019	12,465	19,027
Spouse	214	974	941	2,848	737	17	5,731
Total	3,040	3,806	2,431	7,952	4,245	12,482	33,956

Appendix B: Survey instruments —overview of questionnaires

MHSS2 is an extensive multipurpose 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, global positioning system (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 the 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 and lists 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, and social networks.

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 histories of pregnancy and contraceptive use.

7) Book 5: Child book

Respondent is mother or primary caretaker of a 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 a child aged 0–6 who is a descendant of a member of an 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 MMSE, the Raven, Wechsler Preschool and Primary Scale of Intelligence (WPPSI) matrix reasoning, the 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 the help of 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.

10) Book TN: Toenail collection book

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 could 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 four 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 being 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 four 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 doctor with a private practice 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 five 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 index respondents aged 22–29 and 2,300 aged 30–34. We are fortunate to be able to estimate intracluster correlations (ICCs) 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 and 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 percent for men and 20 percent for women for income changes if we interview close to 85 percent of the sample. Using ICC between 0.02 and 0.05 for cognitive functioning, we can detect effect sizes of between 0.15 and 0.25 sd. 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 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 the 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 SD or DD estimators to determine the ITT effect of the MCH-FP. Intent-to-treat 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 that follows, 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, SD and DD models are appropriate.⁶ 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 SD 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_i AG_i is a binary variable used to indicate whether person i is in age group Y . Thus, the coefficient β_1 is the ITT effect for the 22–29 age group – the group of children eligible to receive the most intensive health interventions. β_2 is the ITT estimate for those children

⁶We can consider matching models; however, the overlap and distribution of the preintervention characteristics between the two experimental areas are so similar that, in previous research (Barham 2012), peer reviewers pointed out that the models were not additive to the papers.

eligible to receive less intensive treatment, and β_3 the ITT estimates for the age groups born prior to the program. α_{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. \mathbf{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 *DD 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 DD should be an underestimate of the program effect. The DD 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 age group is the group without an interaction term. Therefore, β_1 provides the difference in means between treatment and comparison areas for those 35–65-year-olds, or for the “preintervention” period. $\beta_2 - \beta_3$ are the DD estimates for the various age groups of interest.

Mechanisms

Understanding which of the program components were most effective is desirable, but, as with many programs, difficult to determine due to the integrated nature of the interventions. However, we can provide some suggestive evidence of the role of the child health interventions. For example, we can take advantage of the fact that the main child health interventions were phased in after 1981, and compare children who were eligible for the program at birth (the 22–29-year-olds) to those who would have been eligible in later childhood (the 30–34-year-olds). We can also exploit the phasing in of the measles vaccine over time within the treatment area, by comparing the 22–26-year-olds to the 27–29-year-olds within the treatment area.

Appendix E: Description of MMSE

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 MMSE (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 sd 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.

References

- Alam, M., Allison, G., Stagnitti, F., Tanaka, A., and Westbrooke, W. 2002. Arsenic contamination in Bangladesh groundwater: a major environmental and social disaster. *International Journal of Environmental Health Research*, 12(3), pp. 236–253.
- Adair, LS, 1999. Filipino children exhibit catch-up growth from age 2 to 12 years. *Journal of Nutrition*, 129(6), pp.1,140–48.
- Ahlburg, D, 1998. Intergenerational transmission of health. *American Economic Review*, 88(2), pp. 265–70.
- Almond, D, and Currie, J, 2011. Human capital development before age five. In T Paul Schultz and JA Strauss, eds, *Handbook of labor economics*, Volume 4, Part B. Amsterdam: North Holland, pp.1,315-486.
- Almond, D, and Mazumder, B, 2011. Health capital and the prenatal environment: the effect of maternal fasting during pregnancy. *American Economic Journal: Applied Economics*, 3(4), pp. 56–85.
- Baird, S, Hamory, J, and Miguel, E, 2008. Tracking, attrition, and data quality in the Kenyan life panel survey round 1. Unpublished manuscript, Center for International and Development Economics Research, University of California, Berkeley.
- Barham, T, 2012. Enhancing cognitive functioning: medium-term effects of a health and family planning program in Matlab. *American Economic Journal: Applied Economics*, 4(1), pp. 245–73.
- Barham, T, and Kuhn, R, 2014. Staying for benefits? the effect of a health and family planning program on out-migration patterns in Bangladesh. *Journal of Human Resources*, forthcoming.
- Barham, T, Macours, K, and Maluccio, J, 2013. Boys' cognitive skill formation and physical growth: long-term experimental evidence on critical ages and early childhood interventions. *American Economic Review (Papers and Proceedings)*, 103(3), pp. 467–71.
- Barker, DJP, 1992. The fetal and infant origins of adult disease. *British Medical Journal*, 301 (6761), p.1111.
- Bhatia, S, Mosley, WH, Faruque, ASG, and Chakraborty, J, 1980. The Matlab family planning–health services project. *Studies in Family Planning*, 11(6), pp. 202–12.
- Bloom, DE, and Canning, D, 2001. A new health opportunity. *Development*, 44(1), pp. 36–43.

- Bloom, DE, Canning, D, and Jamison, DT, 2004. Health, wealth, and welfare: new evidence coupled with a wider perspective suggest sizable economic returns to better health. *Finance and Development*, 41(1), pp.10–15.
- Case, A, Fertig, A, and Paxson, C, 2005. The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24(2), pp. 365–89.
- Chaudhuri, A, 2005. Direct and indirect effects of a maternal and child health program in rural Bangladesh. *Journal of Developing Societies*, 21(1–2), pp. 143–73.
- Chowdhury, UK, Biswas BK, Chowdhury TR, Samanta G, Mandal, BK, Basu, GC et al. 2000. Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environmental Health Perspectives*, 108(5), 393–397.
- Cleland, J, Phillips, JF, Amin, S, and Kamal, GM, 1994. The determinants of reproductive change in Bangladesh: success in a challenging environment. Washington, D.C.: International Bank for Reconstruction and Development/World Bank.
- Currie, J, 2009. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47(1), pp. 87–122.
- Duflo, E, and Glennerster, R, 2008. Using randomization in development economics research: a toolkit, in: T Paul Schultz and JA Strauss, eds, *Handbook of development economics*, Volume 4. Amsterdam: North Holland, Ch. 61.
- Fauveau, V, 1994. *Matlab: women, children and health*. Dhaka: icddr,b.
- Fogel, R, 1994. Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy. *American Economic Review*, 84(3), pp. 369–95.
- Folstein MF., Folstein, SE., McHugh, PR. 1975. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12,189–198.
- Garces, E, Thomas, D, and Currie, J, 2002. Longer-term effects of Head Start. *American Economic Review*, 92(4), pp. 999–1,012.
- Gertler, P, Heckman, J, Pinto, R, Zanolini, A, Vermeersch, C, Walker, S, Chang, S, and Grantham-McGregor, S, 2012. Labor market returns to early childhood stimulation: a 20-year follow-up to the Jamaica study. NBER Working Paper No. 19185.
- Glewwe, P, and Miguel, EA, 2008. The impact of child health and nutrition on education in less developed countries, in T Paul Schultz and John Strauss, eds, *Handbook of development economics*, Volume 4. Amsterdam: Elsevier, pp. 3,561–606.

Grantham-McGregor, S, Walker, S, Chang, S, and Powell, C, 1997. Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *American Journal of Clinical Nutrition*, 66, pp. 247–53.

Grantham-McGregor, S, Cheung, Y, Cueto, S, Glewwe, P, Richter, L, and Strupp, B, 2007. Developmental potential in the first 5 years for children in developing countries. *Lancet*, 369(9555), pp. 60–70.

Greenberg, DP, von König, C-HW, and Heininger, U, 2005. Health burden of pertussis in infants and children. *Pediatric Infectious Disease Journal*, 24(5), pp. S39–S43.

Haas, J, Murdoch, S, Rivera, J, and Martorell, R, 1996. Early nutrition and later physical work capacity. *Nutrition Reviews*, 54(2), pp. S41–S48.

Heckman, J, 2006. Skill formation and the economics of investing in disadvantaged children. *Science*, 312, pp. 1,900–1,902.

Hill, K, You, D, and Inoue, M, 2012. Child mortality estimation: accelerated progress in reducing global child mortality, 1990–2010. *PLoS Med*, 9(8), p.e1001303. doi:10.1371/journal.pmed.1001303.

Hoddinott, J, Maluccio, JA, Behrman, JR, Flores, R, and Martorell, R, 2008. Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *Lancet*, 371(9610), pp. 411–16.

Huber, D, and Khan, AR, 1979. Contraceptive distribution in Bangladesh villages: the initial impact. *Studies in Family Planning*, 10(8–9), pp. 246–53.

ICDDR,B. 2007. Health and demographic surveillance system Matlab: registration of health and demographic events 2005 vol 39, ICDDR,B, Dhaka, Bangladesh.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.

Jain, M, and Passi, G. 2005. Assessment of a modified mini-mental scale for cognitive functions in children. *Indian Pediatrics*, 42(9), 907–912.

Joshi, S, and Schultz, TP. 2007. Family planning as an investment in development: evaluation of a program's consequences in Matlab, Bangladesh. Yale University Economic Growth Center Discussion Paper No. 951.

Joshi, S, and Schultz, TP, 2013. Family planning and women's and children's health: long-term consequences of an outreach program in Matlab, Bangladesh. *Demography*, 50(1), pp.149–80.

Kabir, Z, and Herliz, A. 2000. The Bangla adaptation of mini-mental state examination: an instrument to assess cognitive function in illiterate and literate individuals. *International Journal of Geriatric Psychiatry*, 15(5), 441–459.

- Kaplan, RM, and Saccuzzo, DP, 2009. Standardized tests in education, civil service, and the military. In *Psychological testing: principles, applications, and issues*, 7th edn. Belmont, CA: Wadsworth, pp. 325–27.
- Khan, M, and Yoder, RA, 1998. Expanded program on immunization in Bangladesh: cost, cost-effectiveness and financing estimates. Partnerships for Health Reform Technical Report No. 24. Accessed February 2011. Available at: <http://www.path.org/vaccineresources/files/Abt-PNACH278.pdf>
- Koenig, MA, Fauveau, V, and Wojtyniak, B, 1991. Mortality reductions from health interventions: the case of immunization in Bangladesh. *Population and Development Review*, 17(1), pp. 87–104.
- Koenig, MA, Khan, MA, Wojtyniak, B, Clemens, JD, Chakraborty, J, Fauveau, V, Phillips, JF, Akbar, J, and Barua, US, 1990. Impact of measles vaccination on childhood mortality in rural Bangladesh. *Bulletin of the World Health Organization*, 68(4), pp. 441–47.
- Kuhn, R, 2005. The determinants of family and individual migration: a case-study of rural Bangladesh. University of Colorado Population Center Working Paper 2005–05.
- Kuhn, R, 2006. The effects of fathers' and siblings' migration on children's pace of schooling in a migrant-sending region of Bangladesh. *Asian Population Studies*, 2(1), pp. 69–92.
- Maluccio, JA, Hoddinott, J, Behrman, JR, Martorell, R, Quisumbing, A, and Stein, AD, 2009. The impact of experimental nutritional interventions on education into adulthood in rural Guatemala. *Economic Journal*, 119(April), pp. 734–63.
- Martorell, R, 1995. Results and implications of the INCAP follow-up study. *Journal of Nutrition*, 125(4), pp. S1127–S1138.
- Martorell, R, 1999. The nature of child malnutrition and its long-term implications. Special issue on public nutrition, *Food and Nutrition Bulletin*, 20(3), pp. 288–92.
- Martorell, R, Khan, L, and Schroeder, D, 1994. Reversibility of stunting: epidemiologic findings in children from developing countries. *European Journal of Clinical Nutrition*, 48, pp. S45–S57.
- Menken, J, and Phillips, JF, 1990. Population change in a rural area of Bangladesh, 1967–87. *Annals of the American Academy of Political and Social Science*, 510, pp. 87–101.
- Mobarak, AM, Kuhn, R, and Peters, C, 2013. Consanguinity and other marriage market effects of a wealth shock in rural Bangladesh. *Demography*, 50(5), pp. 1,845–71.

National Research Council (NRC), 1986. *Population growth and economic development: policy questions*. Washington, D.C.: National Academy Press.

Ouvrier, R, . Goldsmith, R et al. 1993. The value of the mini-mental state examination in childhood: a preliminary study. *Journal of Child Neurology*, 8(2), pp. 145–148.

Parker, S, Rubalcava, L, and Teruel, G, 2008. Evaluation of conditional schooling and health programs. In TP Schultz and J Strauss, eds, *Handbook of development economics*, Volume 4. Amsterdam: North-Holland Press, pp. 3,964–80.

Phillips, J, Simmons, R Chakraborty, R, Al, Chowdhury, 1984. Integrating health services into an MCH-FP program: lessons from Matlab, Bangladesh. *Studies in Family Planning*, 15(4), pp. 153–61.

Phillips, JF, Jones, TC, Nyonator, FK, and Ravikumar, S, 2003. Evidence-based development of health and family planning programs in Bangladesh and Ghana. Population Council Working Paper No. 175.

Phillips, JF, Stinson, WS, Bhatia, S, Rahman, M, and Chakraborty, J, 1982. The demographic impact of the Family Planning-Health Services project in Matlab, Bangladesh. *Studies in Family Planning*, 13(5), pp. 131–40.

Pollitt, E, Gorman, KS, Engle, P, Martorell, R, and Rivera, JA, 1993. Early supplementary feeding and cognition: effects over two decades. *Monographs of the Society for Research in Child Development*, Serial no. 235, 58(7).

Prentice, AM, Ward, KA, Goldbert, GR, Jarjou, LM, Moore, SE, Fulford, AJ, and Prentice, A, 2013. Critical windows for nutritional interventions against stunting. *American Journal for Clinical Nutrition*, 97(5), pp. 911–18.

Reddy, V, 1987. Interaction between nutrition and measles. *Indian Journal of Pediatrics*, 54(1), pp. 53–57.

Reingold, AL, and Phares, CR, 2006. Infectious diseases. In MH Merson, RE Black, and A Mills, eds, *International public health: diseases, programs, systems, and policies*. Sudbury, MA: Jones and Bartlett, pp.127–86.

Rivera, J, Sotres-Alvarez, D, Habicht, J-P, Shamah, T, and Villalpando, S, 1995. Nutritional supplementation during the preschool years influences body size and composition of Guatemalan adolescents. *Journal of Nutrition*, 125(4), Supplement, pp. 1068S–1077S. Rubial-Álvarez, S, Machado, M et al. 2007. A preliminary study of the mini-mental state examination in a Spanish child population. *Journal of Child Neurology*, 22(11), pp. 1269–73.

Rush, A, First M. Blacker D. (eds) 2008. *Handbook of Psychiatric Measures*, 2nd Edition, Arlington, VA: American Psychiatric Publishing.

Sachs, J, and Malaney, P, 2002. The economic and social burden of malaria. *Nature*, 415(6872), pp. 680–85.

Schultz, TP, 2009. How does family planning promote development? evidence from a social experiment in Matlab, Bangladesh, 1977–1996. Unpublished [online]. Available at: <http://www.econ.yale.edu/~pschultz/TPS_10_30_QJE.pdf> [Accessed 15 November 2015].

Strauss E, Sherman, E, Spreen, O, 2006. *A Compendium of Neuropsychological tests: Administration, Norms, and Commentary*, 3rd ed., New York: Oxford University Press.

Strauss, J, and Thomas, D, 1998. Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2), pp. 766–817.

Strauss, J, and Thomas, D, 2008. Health over the life course, in TP Schultz and J Strauss, eds, *Handbook of development economics*, Volume 4. New York: North-Holland, pp. 3,375–474.

Trapp, E, and Menken, J, 2005. Adolescent health and sex bias in nutrition in Bangladesh: evaluating the standards used to classify malnutrition. Paper presented at the 2005 Population Association of America, Philadelphia.

Van Ginneken J, Bairagi, R, Francisco, Ad, Sarder, AM, and Vaughan, P, 1998. *Health and demographic surveillance in Matlab: past, present and future*. Dhaka: Public Health Division, International Centre for Diarrhoeal Disease Research.

Victora, Cesar Gomes, Mercedes de Onis, Pedro Curi Hallal, Monika Blossner, and Roger Shrimpton. 2010. Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, 125(3), pp. e473–e480.

Walker, S, Chang, S, Powell, C, and Grantham McGregor, S, 2005. Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *Lancet*, 366(9499), pp. 1804–7.

Walker, S, Wachs, T, Gardner, J, Lozoff, B, Wasserman, G, Pollitt, E, and Carter, J, 2007. Child development in developing countries 2: child development –risk factors for adverse outcomes in developing countries. *Lancet*, 369(9556), pp.145–57.

Wasserman, G, Liu, X, Parvez, F, Ahsan, H, Factor-Litvak, P, van Geen, A, Slavkovich, V, Lolocono, N, Chen, Z, Hussain, I, Momotaj, H, and Graziano, J, 2004. Water arsenic exposure and children’s intellectual function in Araihasar, Bangladesh. *Environmental Health Perspectives*, 114(1), pp. 124–29.

Publications in the 3ie Impact Evaluation Report Series

The following reports are available from <http://www.3ieimpact.org/en/publications/3ie-impact-evaluation-reports/3ie-impact-evaluations/>

Can e-governance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R (2015)

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A (2015)

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C, and Özler, B (2015)

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E (2015)

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G (2015)

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, 3ie Impact Evaluation Report 24. Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L (2014)

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R (2015)

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M (2015)

Shelter from the storm: upgrading housing infrastructure in Latin American slums, 3ie Impact Evaluation Report 21. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R (2015)

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, 3ie Impact Evaluation Report 20. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P (2015)

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E (2014)

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K (2015)

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, 3ie Impact Evaluation Report 17. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME (2014)

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M (2014)

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V (2014)

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S (2014)

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R (2014)

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R (2014)

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B (2014)

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N (2013)

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, 3ie Impact Evaluation Report 9. Ashraf, N, Bandiera, O and Jack, K (2013)

Paying for performance in China's battle against anaemia, 3ie Impact Evaluation Report 8. Zhang, L, Rozelle, S and Shi, Y (2013)

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, 3ie Impact Evaluation Report 7. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P (2013)

The impact of daycare on maternal labour supply and child development in Mexico, 3ie Impact Evaluation Report 6. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E (2014)

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, 3ie Impact Evaluation Report 5. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M (2014)

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A (2013)

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, 3ie Impact Evaluation Report 3. Casey, K, Glennerster, R and Miguel, E (2013)

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, 3ie Impact Evaluation Report 2. Burwen, J and Levine, DI (2012)

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, 3ie Impact Evaluation Report 1. Martinez, S, Naudeau, S and Pereira, V (2012)

Improving the health and nutrition of young children is important for their immediate well-being. Policymakers also believe it helps to reduce poverty in the long run through improved human capital. This study examined the long-term impact of the Matlab Maternal and Child Health and Family Planning Programme in Bangladesh that began in 1977. Thirty-five years after it began, the study looked at two key measures of human capital: height and cognitive functioning. The findings showed that the programme led to important and sustained effects on a person's height through adulthood. There were important effects on cognitive functioning through 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 other aspects of these people's lives.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444



www.3ieimpact.org