Supplementary information for

Thirty-five years later: Long-term effects of the Matlab Maternal and Child Health/Family Planning Program on older women's well-being

Tania Barham[a,b], Brachel Champion[a,b], Andrew Foster[c], Jena Hamadani[d], Warren C. Jochem[e], Gisella Kagy[f], Randall Kuhn[g]*, Jane Menken[b,h]*, Abdur Razzaque[i], Elisabeth Root[j], Patrick Turner[k]

[a] Department of Economics, University of Colorado Boulder, Boulder CO 80302
[b] Institute of Behavioral Science, University of Colorado Boulder, Boulder CO 80302
[c] Department of Economics and Department of Health Services, Policy and Practice, Brown University RI 02903
[d] Centre for Child and Adolescent Health, icddr,b, Dhaka 1212, Bangladesh
[e] School of Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, UK
[f] Department of Economics, Vassar College, Poughkeepsie NY 12604
[g] Department of Community Health Sciences, UCLA Fielding School of Public Health, Los Angeles CA 90095
[h] Department of Sociology, University of Colorado Boulder, Boulder CO 80302
[i] Health and Population Surveillance Division, icddr,b, Dhaka 1212, Bangladesh
[j] Department of Geography, The Ohio State University, Columbus OH 43210
[k] Department of Economics, University of Notre Dame, Notre Dame IN 46556

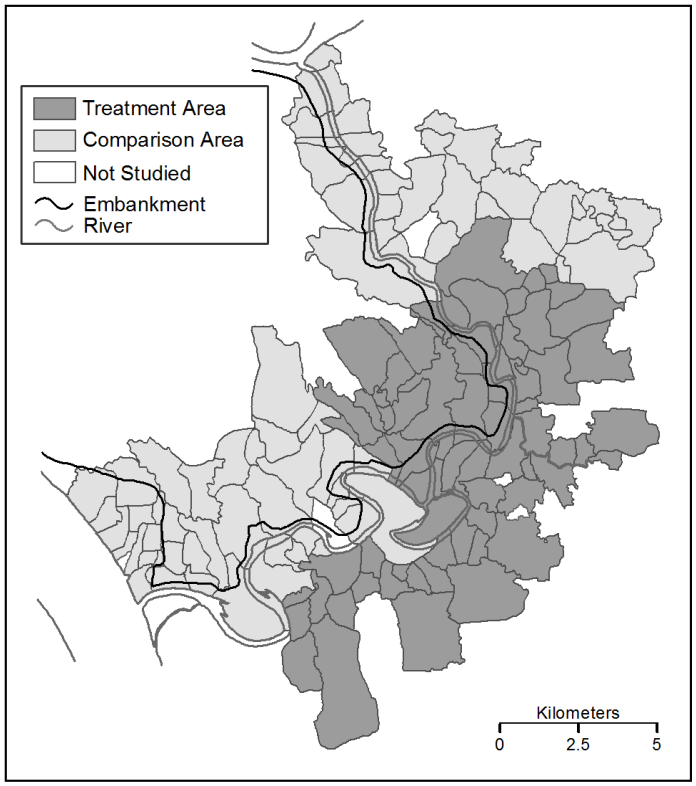
* Corresponding Authors:
Randall Kuhn
Department of Community Health Sciences, UCLA Fielding School of Public Health, 650 Charles E. Young Drive South, Los Angeles, CA, USA, (310) 825-3742,
Email: kuhn@ucla.edu
Jane Menken
Institute of Behavioral Science, University of Colorado Boulder, 483 UCB, Boulder CO 80309-0493 USA, (303) 818-9682,
Email: menken@colorado.edu

This PDF includes:

1. MCH/FP Program

The intention of the MCH/FP program established by icddr,b in the late 1970s was to gain insights that would be useful in designing a scaled-up national program. Because of limitations on women's mobility imposed by the practice of *purdah* (female seclusion), women of childbearing age received in-home delivery of contraceptives and antenatal care.(1) The program offered a wide range of methods and intensive consultation and referral on complications and method change. Government health clinics in the comparison area continued to provide contraceptives in the traditional manner, leading to a difference in take-up rates of approximately 20 percent within the first year.(2) The program incorporated additional maternal and child health interventions over time, including tetanus toxoid (for pregnant women in 1977, all women starting in 1982); iron and folic acid supplementation in the last trimester; polio, DPT and BCG immunization for children (starting 1982 in a portion of the treatment area, extended in 1985 to the entire treatment area) and infectious disease case management.(3) The Bangladesh government did not begin rolling out similar in-home interventions in the comparison area until 1989, and fertility differences persisted for many subsequent years.(4)

The program was rolled out in a block design, as shown in SI_Figure_1.



SI_Figure_1: Matlab study area

2. Baseline Balance

All analysis is based on the assumption that, in the absence of the MCH/FP program, treatment and comparison areas would have remained similar over time. This is not a testable assumption but is reasonable given the similarity between treatment and comparison areas before the program, as shown in SI_Table_1. The table reports means, standard deviations, differences in means, and normalized differences in means (difference in the means divided by the standard deviation of the mean for the comparison group) because the normalized mean is not influenced by sample size.(5) Normalized differences larger than 0.25 standard deviations are generally thought to be substantial; there are none in these data. The significantly higher levels of tubewell access in the treatment area in 1974 potentially result in differing levels of arsenic exposure, so models include arsenic levels in tubewell water in 2003, before wells were identified for contamination, to account for this potential difference.(6)

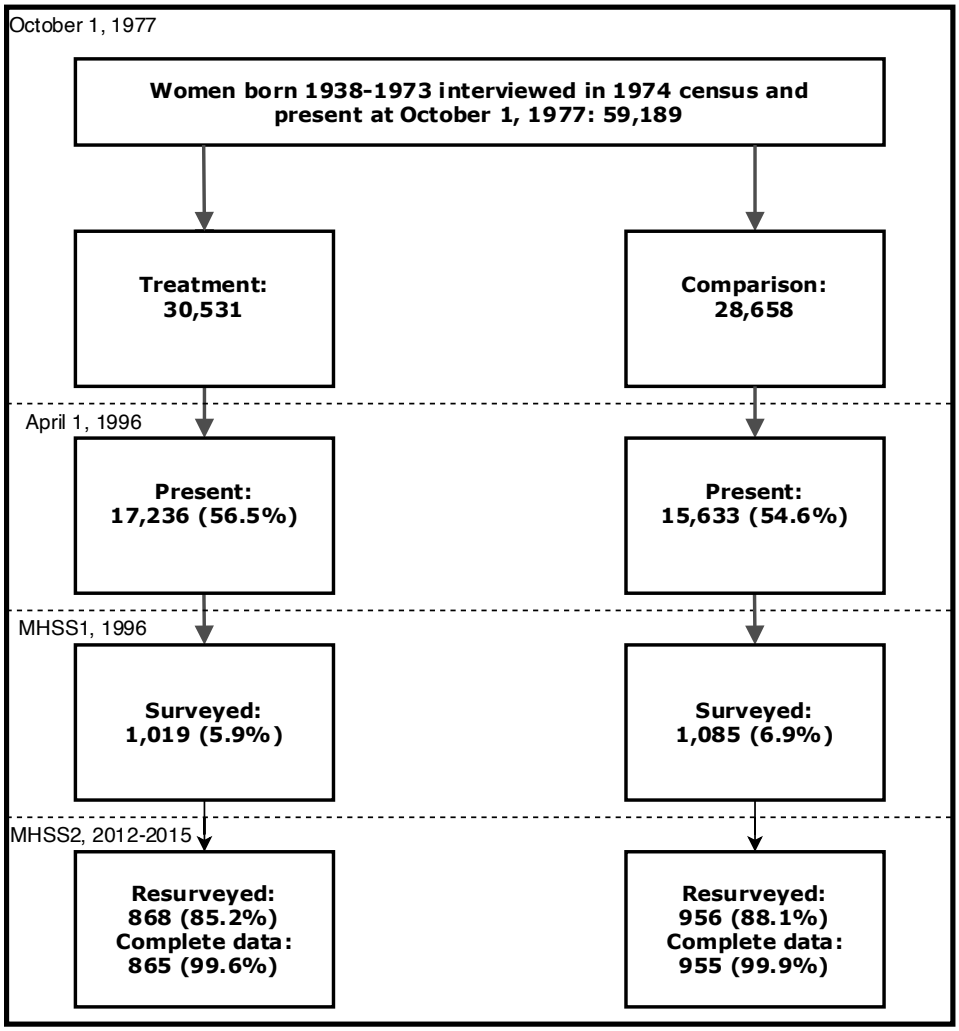
SI_Table_1: Descriptive statistics of individual demographic and household characteristics at baseline (1974) for individuals in the treatment area and comparison area, MHSS2 women born 1938-1973

	Treatment area (N=865)		Comparison area (N=955)		Difference in means		
	Mean	SD	Mean	SD	Diff	T-stat	Mean/SD
Religion Muslim (=1)	0.86	(0.03)	0.95	(0.01)	-0.09	-2.67	-0.09
Age	20.27	(0.43)	20.01	(0.37)	0.26	0.45	0.02
HH years of education	2.42	(0.12)	2.09	(0.11)	0.33	1.98	0.07
HH spouse's years of education	0.84	(0.07)	0.74	(0.06)	0.10	1.07	0.04
Bari Size	8.80	(0.39)	7.96	(0.24)	0.84	1.84	0.06
Family Size	7.00	(0.11)	6.81	(0.09)	0.19	1.33	0.04
HH works in agriculture (=1)	0.60	(0.03)	0.58	(0.02)	0.01	0.35	0.01
HH works in fishing (=1)	0.06	(0.02)	0.06	(0.01)	0.00	-0.15	-0.01
HH works in business (=1)	0.12	(0.02)	0.12	(0.01)	0.00	0.00	0.00
Wall tin or tinmix (=1)	0.32	(0.02)	0.32	(0.02)	0.00	0.12	0.00
Tin Roof (=1)	0.84	(0.02)	0.84	(0.02)	0.00	0.02	0.00
Latrine (=1)	0.84	(0.02)	0.87	(0.02)	-0.02	-0.73	-0.02
Drinking water, tubewell (=1)	0.31	(0.03)	0.16	(0.02)	0.15	3.76	0.13
Drinking water, pond (=1)	0.39	(0.04)	0.33	(0.04)	0.07	1.16	0.04
1982 Land Size	11.57	(0.84)	10.99	(0.55)	0.57	0.57	0.02
Number of rooms per capita	0.21	(0.00)	0.20	(0.00)	0.01	1.36	0.05
Number of boats	0.65	(0.04)	0.67	(0.03)	-0.02	-0.40	-0.01
Number of cows	1.52	(0.09)	1.33	(0.05)	0.19	1.81	0.06
Owns a lamp (=1)	0.65	(0.04)	0.62	(0.02)	0.03	0.65	0.02
Owns a watch (=1)	0.16	(0.02)	0.15	(0.01)	0.01	0.43	0.01
Owns a radio (=1)	0.08	(0.01)	0.08	(0.01)	0.00	-0.31	-0.01
Arsenic Level, 2003 (ppb)	194.69	(14.55)	273.62	(19.55)	-78.92	-3.24	-0.11

Note: HH= Household Head. Sources: Tabulations from 1974 Matlab Census and icddr,b 2003 tubewell arsenic data.

3. Inclusion and Mortality

SI_Figure_2 traces all women who were included in the 1974 census and were present in the HDSS on October 1, 1977, when the MCH/FP program began. MHSS1 included a sample of those who remained in Matlab on April 1, 1996; MHSS2 resurveyed over 85% of MHSS1 respondents within and outside of Matlab.



SI_Figure_2: Sample inclusion diagram

SI_Table_2 shows that mortality did not differ by area for these cohorts either for the 1977-1996 or the 1996-2012 periods. The results confirm earlier life-table analyses that demonstrated no areal differences in survival for women in these birth cohorts. Survival differences emerged only for cohorts born after 1973, those who were exposed to child health interventions.(7)

SI_Table_2: Mortality and Treatment-Comparison differences in mortality between 1974-1996 and between 1996-2012, overall and by birth cohort

	1974-1996					1996-2012				
	Proportion		Difference & Confidence Interval			Proportion		Difference & Confidence Interval		
Cohort	C	T			N	C	T			N
1938-1949	0.07	0.07	-0.01	[-0.01, 0.00]	11662	0.16	0.21	0.04	[-0.01, 0.10]	729
1950-1961	0.04	0.03	-0.00	[-0.01, 0.00]	18105	0.05	0.04	-0.01	[-0.03, 0.02]	795
1962-1973	0.03	0.03	-0.00	[-0.00, 0.00]	29422	0.01	0.02	0.01	[-0.00, 0.03]	831
Overall	0.04	0.04	-0.00	[-0.01, 0.00]	59189	0.07	0.09	0.01	[-0.01, 0.04]	2355

Note: C=Comparison group proportion; T=Treatment group proportion; N=number of observations in group; HDSS=Health & Demographic Surveillance System.
Sources: 1974-1996, Tabulations from HDSS; 1996-2012; Tabulations from linked MHSS1/MHSS2

4. Variable Construction

This section describes the creation of the intent-to-treat treatment status indicator, outcome variables and selected control variables. Economic outcome variables are described in SI_7_Economic Effects.

1. *Treatment Status and Linkage to Baseline Variables*

Access to the MCH-FP program was based on the village of residence of the individual during the program period. Because a person's residence after the program started is potentially endogenous, we use DSS and census data to create an intent-to-treat indicator based on the village of residence for the individual during the 1974 census. We take advantage of the fact that each individual has a unique ID that allows us to link the MHSS1/2 data with the HDSS and census data, and use the following sequence of linkages. First, we link a respondent to the 1974 census. If she was not present in the 1974 census, we identify her first household head in the HDSS area and link the household head to the 1974 census. Remaining unlinked individuals are assigned a treatment status using their own location in the HDSS area after the 1974 census, but before the inception of MCH-FP in 1977. Nearly all women (93.2%) were assigned treatment status from their 1974 census location, 5.4% are linked through their initial household head's location, and 1.3% are assigned based on their own location in the DSS after the census, but before the 1977 implementation. The intent-to-treat variable, hereafter referred to treatment status, takes the value of 1 if the assigned location was in the treatment area.

Baseline characteristics from the 1974 census are linked to individuals in the same manner used to construct treatment status. We fill in baseline characteristics for the few individuals who could not be linked to the 1974 census by assigning means based on treatment status, sex, and cohort.

2. Construction of Selected Outcome and Control Variables

Overall Health Index— The overall health index is constructed from MHSS2 data as the sum of three normalized indices of metabolic health, functional health and respiratory health.

These indices were created as follows: Each component of the index was standardized by using the comparison group mean and standard deviation. The sign of a component variable was switched as necessary so that an increase represents a qualitatively worse outcome. We then used the arithmetic mean of the standardized component values as the index.

Metabolic Health Index— This index was created from indicators for angina, stroke, overweight or obese BMI status, stage 1 and stage 2 hypertension.

> *Angina*— A respondent was coded as having angina if she had ever been told by a health professional that she had angina or angina pectoris and she had taken medications or other treatment for it within the last 12 months. A respondent was also considered to have angina if she said she experienced pain or discomfort in her chest when walking uphill, in a hurry, or at an ordinary pace and that she slowed down or took pain medicine to relieve the pain. The respondent had to indicate that this pain was on the left side of her chest when shown an illustration of a torso. These angina questions were modeled on the WHO Sage method.(8)
>
> *Stroke*— A respondent was coded as having had a stroke if she had been told by a health professional that she had a stroke. Respondents were also considered to have had a stroke if they reported taking medications for it in the last 12 months.
>
> *Overweight or Obese BMI status*— Respondents were considered overweight or obese if their BMI was above 23 kg/m^2, following the WHO recommendation that for Asian populations a cutoff of 23 is more appropriate than the 25 used for the US population.(9) If a respondent's height or weight was above the 99.5 percentile or below the 0.5 percentile for her age, she was dropped from the sample. Height was measured using the Shorr height board. Weight was measured using the SECA881 U digital scale (150kg maximum and 0.01kg increments).
>
> *Stage 1 and stage 2 hypertension*— Three measurements of blood pressure were taken and then averaged. A respondent was considered to have stage 1 hypertension if her systolic blood pressure was between 130-139 or her diastolic blood pressure was between 85-89. She was considered to have stage 2 hypertension or worse if her systolic blood pressure was above 140 or her diastolic blood pressure was 90. For outliers, if systolic blood pressure was less than 60 or more than 250 it was set to missing. Similarly, if diastolic blood pressure was less than 40 or more than 150 it was set to missing. Blood pressure was measured using the Lifesource 767-PV automatic blood pressure device.

Functional Health Index— This index was created from indicators for self-reported difficulty with daily activities, self-reported poor health status, poor objective physical capability score, and the adult daily living mobility score.

> *Self-reported difficulty with daily activities*— Respondents were considered to have difficulties if they self-reported having moderate or severe physical difficulties completing their work or household activities in the previous 30 days or were not able to complete any of their activities.

Self-reported poor health status— Respondents were asked to self-report their current health status as either healthy, fairly healthy or unhealthy/sick. Individuals were considered to have poor health status if they answered unhealthy/sick to this question.

Poor objective physical capability score— Each respondent who was not pregnant completed a series of physical tasks that were then scored. Physical tasks included standing on both feet for 30 seconds without assistance, standing on one leg, chair stands, Timed-Get-Up-and-Go test for 3 and 4 meters and picking up a pencil on the ground from a standing position. Each test received a score of 0-4 depending on the interval of time required to complete each task. Scores were summed and a respondent was assigned poor objective physical capability if she scored less than 13 (max score=17).

Adult Daily Living Mobility Score— Respondents were presented with a list of daily activities and asked if they could easily do each task (=0), could accomplish the task with a lot of difficulty (=1) or were unable to do the task altogether (=2). The daily tasks included stand from sitting on the floor without help, sit on the floor with bent knees, stand from sitting on a chair without help, and bend and pick up a pen from the floor. Each individual was assigned a score that summed her responses to each task (for a maximum score of 8) so that a higher score represented worse mobility.

Maximum dominant hand grip strength— Two measurements of grip strength were taken for each individual using dynamometers. We use the maximum reading for the respondent's dominant hand.

Respiratory Health Index— This index is created from indicators for chronic lung disease and asthma.

Chronic Lung Disease— A respondent was coded as having chronic lung disease if she had ever been told by a health professional that she has a chronic lung disease, such as emphysema, bronchitis, or chronic obstructive pulmonary disease. Respondents were also considered to have a chronic lung disease if they reported taking medications for it in the last 12 months.

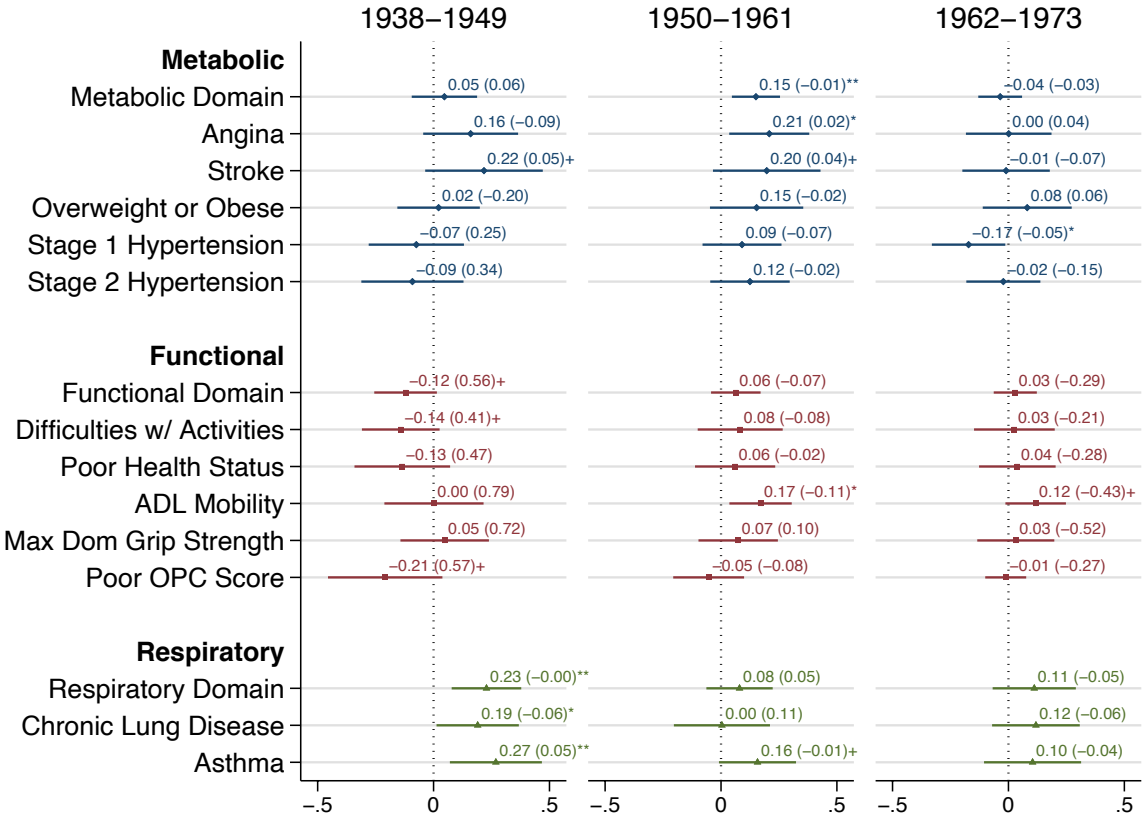
Asthma— A respondent was coded as having asthma if she had ever been told by a health professional that she has asthma and she has taken medications or other treatment for it within the last 12 months. A respondent was also considered to have asthma if she said she experienced attacks of wheezing or whistling breathing and these attacks have happened after she stopped physical activity, in the morning or without obvious cause. These asthma questions were based on the WHO Sage method.

Arsenic Exposure— The control for arsenic was created using measures of arsenic in tube well water collected by icddr,b in 2003. These data were collected prior to knowledge of arsenic in well water and so before families engaged in well switching, which could be correlated with treatment status. Wells were linked to MHSS1 households using the ID of the person who takes care of the well. For households that don't take care of a well, we use the average arsenic level in

the 3 closest wells. For households that reported not using a tubewell in MHSS1, the value of arsenic is set to zero. Arsenic is measured in parts per billion (micrograms per liter). We include binary variables of arsenic levels between 50-99, 100-149, 150-399, and >400ppb (so that 0-50ppb is the reference category). Results are similar across various methods of including the control (i.e. as a continuous variable or as a binary based on cuts at 100, 150 or 200).

5. Unpacking Health Effects

SI_Figure_3 unpacks effects in each domain by plotting treatment effects for each normalized item that contributed to the index. For the 1950-1961 cohort, the significantly poorer treatment area metabolic health reflected a combination of significant or borderline significant effects, with only angina risk significant at the 5% level (+0.21 SD, 95% CI 0.040 to 0.385, $p<0.05$), but all were signed in the same direction and ranged from +0.09 SD in Stage 1 hypertension risk (95% CI -0.080 to 0.258, $p=0.302$) to +0.20 SD in stroke risk (95% CI -0.029 to 0.434, $p=0.086$). In the functional health domain, women in the treatment area had worse ADL mobility by +0.17 SD (95% CI -0.037 to 0.306, $p<0.05$), Within the respiratory domain, we found borderline significantly higher risk of asthma (+0.16 SD, 95% CI -0.009 to 0.323, $p<0.10$). Turning to the oldest cohort (1938-1949), women in the MCH/FP area experienced significantly higher risk of asthma (+0.27 SD, 95% CI 0.070 to 0.467, $p<0.01$) and a modestly higher risk of chronic lung disease (+0.19 SD, 95% CI 0.013 to 0.368, $p<0.05$) and stroke (+0.22 SD, 95% CI -0.037 to 0.470, $p<0.10$). Within the youngest cohort (1962-1973), we observed significantly lower risk of Stage 1 hypertension (-0.17 SD, 95% CI -0.330 to -0.013, $p<0.05$) among women in the MCH/FP area.



SI_Figure_3: Program effects on health domain components: Adjusted difference between treatment and comparison area means by birth cohort

Note: Points represent the difference between means; lines represent the 95% confidence interval. The comparison area mean is reported in parentheses. Positive differences indicate poorer health in the treatment area. See main text Table 1 note for adjustment model. ADL=Activities of Daily Living, OPC=Objective Physical Capacity. +p<0.10; *p<0.05; **p<0.01.

6. Mediating Effects

We conducted a simple mediation analysis to better understand negative health outcomes within the 1950-1961 cohort by introducing controls for key mediating pathways between MCH/FP and health outcomes. Column 1 of SI_Table_3 replicates results shown in main text Figure 1 while columns 2-6 controlled for each potential confounder listed. We did not observe any mediating effects of children ever born (indicators for having 0-2 children, having exactly 3, 4, 5 or 6 children, and having 7 or more children the reference category) (column 2) or birth spacing (age at first birth, last birth and average birth interval) (column 3). Controlling for BMI in 1996 explained about 1/3rd of the metabolic effect (from +0.15 SD in base model to +0.09 SD) and the overall health effect. Neither functional nor respiratory health were meaningfully affected by the inclusion of additional confounders.

SI_Table_3: Program effects on health domains: Adjusted differences between treatment and comparison area means controlling for potential confounders for 1950-1961 cohort

	(1) Baseline	(2) Children Born	(3) Birth Spacing	(4) MHSS1 BMI	(5) MHSS2 BMI	(6) Metabolic Index
Overall Health	0.11*	0.11*	0.11*	0.08	0.08+	NA
	[0.01, 0.20]	[0.01, 0.20]	[0.01, 0.22]	[-0.02, 0.17]	[-0.00, 0.17]	NA
Metabolic	0.15**	0.14**	0.14**	0.09+	0.12**	NA
	[0.05, 0.25]	[0.04, 0.25]	[0.04, 0.25]	[-0.01, 0.19]	[0.03, 0.21]	NA
Functional	0.06	0.08	0.08	0.08	0.05	0.04
	[-0.04, 0.17]	[-0.03, 0.19]	[-0.03, 0.20]	[-0.04, 0.19]	[-0.06, 0.15]	[-0.06, 0.15]
Respiratory	0.08	0.07	0.09	0.07	0.05	0.04
	[-0.06, 0.22]	[-0.07, 0.22]	[-0.07, 0.24]	[-0.09, 0.24]	[-0.09, 0.19]	[-0.10, 0.17]
N	698	697	676	618	679	697

Note: Point estimate is the model-adjusted difference in treatment and comparison mean of row health index (measured in standard deviations) controlling for column confounder. See main text Table 1 note for adjustment model. Column 1 is the original model for comparison, columns 2-6 include the baseline controls in column 1. +p<0.10; *p<0.05; **p<0.01. Source: Tabulations from linked MHSS1/MHSS2.

7. Economic Outcomes

1. Construction of Outcome Variables

Consumption— Both MHSS1 and MHSS2 collected data on food and nonfood items measured over different recall periods, as is typical in the World Bank Living Standard and Measurement surveys. The household head was asked to report consumption of various items over 7-day, 30-day, and 12-month recall periods. 7-day recall included 118 food, drink or tobacco related items that were purchased, produced, and transferred to the household. The 30-day recall recorded expenditure of basic household items (such as items for basic hygiene), services, and utility expenses, and the 12-month recall included personal and household items such as clothing, kitchen items, appliances and furnishings, and vehicle repair. For food items, when available we used the value and quantity of purchased food to assign a value to the quantity of food produced or transferred. For households without purchased food, we used average prices determined from households in nearby areas. Additionally, we removed outlier values by item, defining the outlier cutoff as the smallest value that fell more than two standard deviations above the nearest value.

We constructed annual aggregate consumption measures at both household and per capita levels for two main reasons. First, treated households are on average larger than non-treated because individuals in the treated area were less likely to migrate. Second, the composition of households

over a year can change, making per capita measures potentially noisy, especially in high-migration areas. All consumption dependent variables are given in 2012 US dollars.

Annual Earnings— In MHSS1, annual wages were constructed from a survey module that recorded annual income for occupations that garnered wages. Earnings from self-employment or any family business or farm activities were not recorded. In MHSS2, annual earnings were constructed from a similar survey module that captured both paid and unpaid work from a set of eight general employment activities that was designed to cover all possible types of work (e.g. salaried work, piece-rate work, self-employment, etc.). Unlike MHSS1, questions in MHSS2 were asked by employment category to reduce the chance that respondents would forget to report income if they worked multiple jobs. In both cases, earnings were deflated to 2012 values using World Bank national accounts data and then converted from Bangladeshi taka to US dollars using an exchange rate of 78Tk/US. There are some large outcomes in the dataset, but the results are similar when the data are trimmed by setting to missing the earnings values that are above the 95th percentile, separately by birth cohort.

Employment and Savings— Women self-reported if they worked for wages in the previous 12 months for each of the eight activities and were labeled employed if they had done any work for pay in the previous 12 months, and if they had any cash savings.

Productive Assets— Women self-reported if they owned any assets separate from jointly-owned household assets that could be used to earn money, such as cattle, sewing machine or loom.

2. Program Effects on Economic Outcomes

SI_Table_4 and SI_Table_5 presents associations between programmatic exposure and economic outcomes at MHSS2 and MHSS1. SI_Table_4 examines effects on per capita household consumption and log of per capita household consumption at both MHSS1 and MHSS2. There were no significant effects on consumption in either MHSS1 or MHSS2, except for a marginally significant increase in MHSS1 consumption in the youngest group when we did not adjust for baseline covariates, which included household assets. In SI_Table_5, we looked at MHSS1 wage income as well as MHSS2 total earnings, a dummy variable for any employment in the past year, a dummy variable indicating any individual savings, and a dummy for any individual productive assets. Based on the models that include controls, treatment area women in the oldest cohort were 6 percentage points more likely to have savings (95% CI 0.01 to 0.11, $p<0.05$), almost double the rate with savings in the comparison group. There was no impact on annual earnings for the oldest group. In the middle cohort, we observed significantly lower earnings (-$57.13, 95% CI -$108.85 to -$5.42, $p<0.05$). There was no significant impact on employment for the youngest cohort. Women in the oldest cohort were more likely to have their own savings (0.06, 95% CI 0.01 to 0.11, $p<0.05$). There was no significant impact on individual asset ownership. No effects were significant when all groups were combined, with the exception of a marginal significant negative impact on MHSS2 earnings.

SI_Table_4: Program effects on consumption: Differences between treatment and comparison area means by birth cohort

	(1)	(2)	(3)	(4)
	Per Capita Consumption (2012 USD)		Log Consumption	
	MHSS1	MHSS2	MHSS1	MHSS2
Panel A: Adjusted				
1962-1973	17.65	3.87	-0.04	0.00
	[-28.20, 63.50]	[-44.08, 51.82]	[-0.20, 0.11]	[-0.09, 0.10]
1950-1961	12.36	-1.34	-0.03	-0.05
	[-37.96, 62.69]	[-67.40, 64.73]	[-0.18, 0.13]	[-0.14, 0.03]
1938-1949	-0.53	34.51	-0.03	0.03
	[-67.42, 66.36]	[-20.37, 89.39]	[-0.22, 0.15]	[-0.07, 0.14]
1938-1973 (combined)	14.01	6.15	-0.03	-0.01
	[-19.83, 47.85]	[-28.24, 40.54]	[-0.15, 0.10]	[-0.07, 0.05]
Panel B: Unadjusted				
1962-1973	41.45+	19.53	0.02	0.03
	[-7.51, 90.42]	[-23.34, 62.39]	[-0.13, 0.17]	[-0.05, 0.12]
1950-1961	21.04	7.02	0.01	-0.03
	[-32.17, 74.26]	[-59.77, 73.80]	[-0.15, 0.17]	[-0.12, 0.05]
1938-1949	11.01	28.39	-0.02	0.02
	[-55.29, 77.32]	[-23.00, 79.77]	[-0.20, 0.16]	[-0.08, 0.12]
1938-1973 (combined)	27.85	17.22	0.01	0.01
	[-9.66, 65.35]	[-18.52, 52.96]	[-0.12, 0.13]	[-0.05, 0.07]
Mean Control 1973	331.01	449.86	5.61	5.96
Mean Control 1961	342.99	490.31	5.61	6.06
Mean Control 1949	402.97	432.38	5.78	5.94
N	1820	1816	1820	1816

Note: Point estimates are differences in treatment and comparison mean of column outcome by row cohort. Headers denote survey round. Columns 1 and 2 are measured in 2012 USD. For Panel A adjustment model, see main text Table 1 note. Panel B is unadjusted.
Source: Tabulations from linked MHSS1/MHSS2.

SI_Table_5: Program effects on economic production: Differences between treatment and comparison area means by birth cohort

	(1)	(2)	(3)	(4)	(5)
		MHSS2			
	MHSS1 Annual Wages (2012 USD)	Annual Earnings (2012 USD)	Employed (=1)	Has Her Own Savings (=1)	Has Her Own Productive Assets (=1)
Panel A: Adjusted					
1962-1973	-23.57	-23.77	0.06	-0.02	0.02
	[-98.93, 51.79]	[-78.72, 31.19]	[-0.03, 0.15]	[-0.08, 0.04]	[-0.04, 0.09]
1950-1961	-25.21	-57.13*	-0.06	0.02	0.04
	[-63.43, 13.01]	[-108.85, -5.42]	[-0.13, 0.02]	[-0.04, 0.08]	[-0.02, 0.10]
1938-1949	27.14	10.96	0.04	0.06*	0.02
	[-17.82, 72.09]	[-13.85, 35.76]	[-0.01, 0.09]	[0.01, 0.11]	[-0.02, 0.06]
1938-1973 (combined)	-10.47	-30.33+	0.01	0.01	0.03
	[-47.94, 27.00]	[-63.44, 2.79]	[-0.03, 0.06]	[-0.03, 0.04]	[-0.01, 0.06]
Panel B: Unadjusted					
1962-1973	2.54	-5.05	0.05	0.01	0.01
	[-74.01, 79.09]	[-56.34, 46.24]	[-0.03, 0.14]	[-0.05, 0.08]	[-0.05, 0.08]
1950-1961	-13.76	-34.84+	-0.07+	0.02	0.02
	[-56.94, 29.42]	[-74.76, 5.09]	[-0.15, 0.01]	[-0.04, 0.08]	[-0.03, 0.08]
1938-1949	30.05	9.74	0.05+	0.06**	0.03
	[-18.20, 78.30]	[-13.12, 32.60]	[-0.00, 0.09]	[0.02, 0.11]	[-0.01, 0.08]
1938-1973 (combined)	3.00	-12.14	0.01	0.03	0.02
	[-34.41, 40.42]	[-39.03, 14.75]	[-0.04, 0.06]	[-0.01, 0.06]	[-0.01, 0.05]
Mean Control 1973	53.21	65.40	0.46	0.17	0.19
Mean Control 1961	58.91	56.20	0.47	0.10	0.14
Mean Control 1949	19.66	6.17	0.11	0.05	0.04
N	1783	1776	1816	1765	1773

Note and Source: See SI_Table_4. +p<0.10; *p<0.05; **p<0.01.

8. Robustness Checks

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

Erosion and Flood Risk— One potential confounder is the Meghna Dhonnogoda Irrigation Project. In 1987 the government of Bangladesh completed this project, which involved constructing a river embankment along the northern bank of the major Meghna River where it meets the west bank of the smaller Dhonnogoda River, which runs through Matlab (see Figure 1). The villages near this project were all located in the comparison area, and the embankment had two important consequences for these villages. First, seven villages in this area lining the river were partially or fully inundated between 1984 and 1986 as part of the project. All

households in these villages were displaced, with most initially relocating to adjoining villages within the comparison area. Second, owing to the size and strength of the Meghna River, the embankment was relocated mid-project to a more stable position farther from the river, so there are a number of villages in the Meghna area between the river and the embankment that are more likely to suffer from flooding. Indeed, there were major floods on this river in 1987 and 1988. Migration rates were slightly higher in general in these two areas before the embankment project because of more frequent flooding. To control for potential differences in the Meghna area in general, we include two variables indicating whether a person's treatment village was submerged as a result of the project (Eroded Village=1) or was not submerged but was between the Meghna River and the embankment (Adjacent Meghna River Village=1).

1993 BRAC Microcredit Experiment— In 1993, BRAC introduced an experiment that provided landless women with access to microcredit. The rollout was designed to include villages in both treatment and comparison areas, but the presence of the program could still bias our results. We include an indicator for whether the village ever participated in BRAC (BRAC=1) during its experimental period.

Pourashava Area Village— There was gradual urbanization of villages near Matlab town and the surrounding pourashava area (the local government municipality). Households in these villages may engage in different work (and thus have different health risks), have different access to the local health supply because they are closer to the hospital, and potentially have different migration patterns. Furthermore, 75% of these villages were located in the treatment area, potentially biasing the effects of living in a treatment village. We control for whether an individual lived in one of these villages (Pourashava Village=1).

Migration Behavior— We control for whether the individual was presently living in an urban district or in a district outside of Chandpur district in which Matlab is located. It is possible that the labor opportunities, healthcare access, living standards, etc. affected health outcomes differently than those still living in Chandpur or some other rural district.

SI_Table_6 shows means of these additional potential confounders by treatment and comparison area, and their differences. None of the confounders had normalized treatment/comparison difference greater than the recommended benchmark of 0.25.(5)

SI_Table_6: Descriptive statistics for additional potential confounders, MHSS2 women born 1938-1973

	Treatment area (N=865)		Comparison area (N=955)		Difference in means		
	Mean	SD	Mean	SD	Mean	T-stat	Mean/SD
Eroded Village (=1)	0.00	(0.00)	0.09	(0.04)	-0.09	-2.05	-0.07
Adjacent Meghna River Village (=1)	0.00	(0.00)	0.28	(0.07)	-0.28	-4.13	-0.13
BRAC (=1)	0.65	(0.08)	0.52	(0.08)	0.13	1.14	0.04
Pourashava Village (=1)	0.28	(0.08)	0.09	(0.05)	0.19	2.09	0.07
In Urban District (=1)	0.05	(0.01)	0.06	(0.01)	-0.01	-1.14	-0.04
Outside of Chandpur (=1)	0.06	(0.01)	0.07	(0.01)	-0.01	-1.10	-0.04

SI_Table_7 presents the results for the main health outcomes of interest (analogous to SI_Table_3) with the inclusion of additional controls. Column 1 replicates the results in main text Figure 1 for comparison. Column 2 includes dummies for whether the respondent previously lived in a village that eroded due to the irrigation project or lived in a village between the embankment and the Meghna river. Column 3 includes a dummy indicating if the respondent lived in a pourashava village near the Matlab town area. Column 4 includes a dummy indicating if the respondent lived in a village with access to BRAC. Column 5 includes a dummy indicating if the respondent lived in an urban location or out of the Chandpur district at MHSS2. Lastly, Column 6 removes all controls. Any effects of these confounders are minor.

SI_Table_7: Program effects on health domains: Adjusted difference between treatment and comparison areas means controlling for additional potential confounders, 1950-1961 cohort

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Eroded & Adjacent Meghna River	Pourashava Village	BRAC	Own Migration	No Controls
Overall Health	0.11*	0.08	0.08+	0.10*	0.11*	0.12**
	[0.01, 0.20]	[-0.02, 0.18]	[-0.01, 0.17]	[0.00, 0.19]	[0.01, 0.20]	[0.03, 0.20]
Metabolic	0.15**	0.14*	0.12*	0.15**	0.15**	0.16**
	[0.05, 0.25]	[0.03, 0.25]	[0.01, 0.22]	[0.04, 0.25]	[0.05, 0.25]	[0.07, 0.26]
Functional	0.06	0.04	0.05	0.05	0.07	0.06
	[-0.04, 0.17]	[-0.07, 0.15]	[-0.05, 0.15]	[-0.05, 0.15]	[-0.04, 0.17]	[-0.03, 0.15]
Respiratory	0.08	0.05	0.07	0.09	0.08	0.09
	[-0.06, 0.22]	[-0.11, 0.21]	[-0.08, 0.22]	[-0.06, 0.24]	[-0.06, 0.22]	[-0.05, 0.23]
N	698	698	698	698	698	698

Note and Source: See SI Table 3. +p<0.10; *p<0.05; **p<0.01.

9. Weights

The main results are weighted for attrition between 1974 and MHSS2 using inverse propensity weights. The analysis sample includes respondents from MHSS1/2 who were born between 1938 and 1973. We also include a sample of individuals from MHSS1 households who had migrated out of the HDSS area prior to the survey conducted in 1996. To reduce potential bias due to selective in-migration into treatment versus comparison area, we excluded individuals who entered the HDSS area after October 1977. The main reasons for non-response are death, primarily during infancy, and migration in early adulthood.

We construct weights taking into account the likelihood of attrition from (1) the MHSS1 sampling frame and (2) the subsequent follow-up of MHSS1 respondents in the MHSS2 survey. We estimate these two probabilities separately. First, we estimate the conditional probability that an individual born into one of our cohorts and living in the DSS area as of 1977 was present to be surveyed in MHSS1. To do this, we assign our treatment variable to the universe of women living in the DSS between 1974 and 1977 and born between 1938 and 1973. We estimate a probit model on the probability an individual is present in the DSS on January 1, 1996 using the set of baseline household and household head characteristics, their interactions with the treatment variable, year of birth fixed effects, and indicators for whether an individual was from a village that experienced erosion or was exposed to the Meghna Dhonnogoda Irrigation Project,

separately for each of our three cohort groups. The second probability is constructed in a similar manner. We estimate the probability of non-attrition between the two survey waves for each cohort group separately using a probit model and the same set of covariates. The resulting attrition weight is the inverse of the product of the two probabilities.

Column 1 of SI_Table_8 repeats the results shown in Figure 3 of the main paper based on these attrition rates. Column 2 is based on the sample without weights. Results are near-identical.

SI_Table_8: Program effects on health domains: Model-adjusted difference between treatment and comparison area means, with and without attrition weights

	(1) Attrition Weights	(2) Unweighted
Panel A: 1962-1973		
Overall Health	0.05	0.05
	[-0.04, 0.14]	[-0.04, 0.13]
Metabolic	-0.04	-0.03
	[-0.13, 0.06]	[-0.13, 0.06]
Functional	0.03	0.03
	[-0.06, 0.12]	[-0.07, 0.12]
Respiratory	0.11	0.11
	[-0.07, 0.29]	[-0.07, 0.29]
N	570	570
Panel B: 1950-1961		
Overall Health	0.11*	0.11*
	[0.01, 0.20]	[0.02, 0.20]
Metabolic	0.15**	0.16**
	[0.05, 0.25]	[0.06, 0.25]
Functional	0.06	0.07
	[-0.04, 0.17]	[-0.04, 0.17]
Respiratory	0.08	0.08
	[-0.06, 0.22]	[-0.06, 0.23]
N	698	698
Panel C: 1938-1949		
Overall Health	0.08	0.07
	[-0.02, 0.18]	[-0.03, 0.17]
Metabolic	0.05	0.05
	[-0.09, 0.19]	[-0.09, 0.19]
Functional	-0.12+	-0.13+
	[-0.26, 0.01]	[-0.26, 0.01]
Respiratory	0.23**	0.22**
	[0.08, 0.38]	[0.07, 0.37]
N	552	552

Note and Source: See SI Table 3. +p<0.10; *p<0.05; **p<0.01.

References

1. J. F. Phillips, R. Simmons, M. A. Koenig, J. Chakraborty, Determinants of Reproductive Change in a Traditional Society: Evidence from Matlab, Bangladesh. *Stud. Fam. Plann.* **19**, 313–334 (1988).

2. W. S. Stinson, J. F. Phillips, M. Rahman, J. Chakraborty, The Demographic Impact of the Contraceptive Distribution Project in Matlab, Bangladesh. *Stud. Fam. Plann.* **13**, 141–148 (1982).

3. T. Barham, R. Kuhn, J. Menken, A. Razzaque, Thirty-five years later: Evaluating the impacts of a child health and family planning programme in Bangladesh (Impact Evaluation Report 39). *New Delhi India Int. Initiat. Impact Eval.* (2016).

4. M. Z. Haq, *et al.*, Health and Demographic Surveillance System–Matlab, v. 51. Registration of health and demographic events 2016 (2018).

5. G. W. Imbens, J. M. Wooldridge, Recent Developments in the Econometrics of Program Evaluation. *J. Econ. Lit.* **47**, 5–86 (2009).

6. D. Chakraborti, *et al.*, Arsenic groundwater contamination in Middle Ganga Plain, Bihar, India: a future danger? *Environ. Health Perspect.* **111**, 1194–1201 (2003).

7. J. Menken, *et al.*, Long-term effects of the Matlab Maternal and Child Health and Family Planning Program on Cohort Survival and Health in *2017 International Population Conference*, (IUSSP, 2017).

8. G. A. Rose, The diagnosis of ischaemic heart pain and intermittent claudication in field surveys. *Bull. World Health Organ.* **27**, 645–658 (1962).

9. E. C. Who, Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet Lond. Engl.* **363**, 157 (2004).