Essays in Health and Development Economics

by

Reshmaan Nahar Hussam

S.B. Economics, Massachusetts Institute of Technology (2009)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

Signature redacted

Author .....
Department of Economics
August 27, 2015

Signature redacted

Certified by.....
Esther Duflo
Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics
Thesis Supervisor

Signature redacted

Certified by.....
Abhijit V. Banerjee
Ford Foundation International Professor of Economics
Thesis Supervisor

Signature redacted

Accepted by .....
Ricardo J. Caballero
Ford International Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays in Health and Development Economics

by

Reshmaan Nahar Hussam

Submitted to the Department of Economics
on August 27, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis is a compilation of three empirical studies exploring significant but underexamined health and development challenges of the late 20th and early 21st centuries in South Asia.

Chapter One investigates the effects of the expansion of ultrasound technology throughout India in the 1980s on the childbearing decisions of parents and the marriage market dynamics of exposed children. While ample work has documented the relationship between access to sex selection technology and heavily male-skewed child sex ratios, we know little about how such exposure translates into later life marriage market outcomes of children in highly sex-skewed regions, nor about how parental choice regarding sex selection is affected by such shifts in their children's marital prospects. I build on a theory proposed by Edlund (1999) that, in environments where hypergamy is practiced and parents derive utility from married children, a male-skewed sex ratio can generate a permanent female underclass. By examining the relationship between the child sex ratio of couples of childbearing age and that of their contemporaneous marriage market, I offer evidence that parents do indeed internalize the marriage prospects of their unborn children and adjust their use of sex selection technology accordingly. Importantly, this adjustment occurs significantly more amongst poor families than wealthy families. By exploiting spatial and temporal variation in exposure to ultrasound technology, I then examine the implications of such socioeconomically skewed ultrasound use on the marital outcomes of children in regions with high ultrasound access. I find that, relative to her unexposed counterpart, the average exposed married female has significantly poorer health and less education; there exists a wider marriage and education gap between herself and her husband; and she reports lower autonomy, less decision making power, and more abuse, among other bargaining outcomes. While existing literature suggests that scarcity of females in a marriage market should increase their bargaining power, I offer evidence to the contrary in this nationwide setting of endogenous and socioeconomically stratified sex selection. This exercise underscores the intergenerational welfare consequences of poorly regulated access to sex selection technology: not only upon the millions of 'missing women' lost to sex selection, but upon surviving females as well.

Chapter Two explores the impact of a 1999 public health campaign in Bangladesh, which sought to protect millions of individuals from exposure to arsenic-contaminated water, on infant and child mortality. The study was motivated by the dearth of literature on the effects of arsenic exposure on children (whereas its effects on adults, often manifested in the cancer arsenicosis, are well known). It quickly evolved into an examination of the unintended consequences of a highly influential but poorly planned public health campaign. Exploiting the local random nature of arsenic contamination of groundwater in Bangladesh, paired with the timing of child births and thus exposure to such contaminated water, we find that households in which children were exposed to arsenic for a shorter duration (because the household responded to the health campaign by switching away from arsenic-contaminated groundwater sources) in fact experience significantly *higher* rates of infant and child mortality relative to their counterparts. We present evidence that this unanticipated rise in mortality is due to the quality of alternatives that a switching household faced: households had to choose between arsenic-laden but easily accessible shallow tubewell water, which was protected from fecal bacteria; arsenic-free and easily accessible surface water, which was heavily exposed to fecal bacteria; or distant and inconvenient potable water, which was more likely to be exposed to bacteria at the point-of-use. As bacterial contamination is a leading cause of infant and child death in Bangladesh, we argue through a series of exercises that this is a likely driver of the rise in mortality rates amongst young children whose families switched away from arsenic-contaminated tubewells. In determining their water source, households were

essentially trading off arsenic exposure and the resulting rise in old-age mortality with bacterial exposure and the resulting rise in the mortality of their young. The study motivates caution in the execution of large-scale public health and behavioral change campaigns when alternatives to the discouraged behavior are poorly understood.

While my first two chapters investigate household health behavior, a demand-side component of the healthcare market, the next chapter explores a critical player on the supply side. Chapter Three studies the impact of a nine-month generalized training program on the knowledge and performance of private informal healthcare providers in West Bengal, India. These providers, colloquially referred to as "quacks" and described here as "informal providers" (IPs), constitute nearly 80% of the Indian healthcare provider market. However, none possess medical degrees and few have any formal certification to practice medicine. They have been the focus of considerable debate in recent years, with many pushing for their elimination while others propose their integration into the public healthcare system. To inform the debate, it is important to understand whether the quality of healthcare provided by IPs can be improved sufficiently for effective and welfare-increasing integration. The training program examined in this study was the first of its kind to be rigorously evaluated for its impact on IP knowledge and quality of care. We employ a randomized controlled trial (RCT) design, in which we randomly assigned 152 IPs to treatment and 152 IPs to control. Treatment IPs were invited to attend the program, which was taught by certified doctors and consisted of two two-hour classes per week over nine months. Endline data was collected twelve to fourteen months after the start of training. Standardized patient data, corroborated by clinical observations, demonstrate that those IPs offered the program spent significantly more time with their patients, completed a more thorough set of history questions and examinations, and provided more effective treatments. However, we see no shift in the frequency with which they practiced polypharmacy nor the dispensation of unnecessary antibiotics, two harmful practices which plague both the private and public healthcare system. We conclude that training offers a low cost, highly effective method to improve the quality of care delivered by IPs, but that deeper knowledge failures or misaligned incentives may be driving practices such as polypharmacy, for which training may not be a sufficiently powerful intervention.

Thesis Supervisor: Esther Duflo
Title: Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics

Thesis Supervisor: Abhijit V. Banerjee
Title: Ford Foundation International Professor of Economics

Contents

Acknowledgements

I am grateful to my advisors, Abhijit Banerjee and Esther Duflo, for their support throughout - from my time as an undergraduate into these years of graduate school. I am also indebted to my co-authors on the second two chapters of this dissertation: Erica Field and Rachel Glennerster for the study on infant mortality in Bangladesh; and Jishnu Das and Abhijit Chowdhury for the study on the training program for IPs in West Bengal. These experiences fueled my love for a land and a people that I have long felt removed from, and they have been sources of expansion for me in mind and heart.

I have the deepest gratitude for my earliest mentors in economics: Bart Wilson, David Porter, and Vernon Smith. It was their patient investment in me and critical and expansive approach to the field that was the door to the opportunities I have been blessed with. I cannot thank them enough.

Sara Hernandez and Natalia Rigol, the gypsy and the captain. Gracias for being exactly as you are.

And to my family. Abbu, the thinker in me comes from watching you. Ammu, here is that dissertation you put on hold for thirty years for us. Bhaiya, thank you for believing I was able. Tani, thank you for never pushing me off Ammu's lap. Shareef, you will soar - remember your heart. And to my dear Nebil, I couldn't have napped my way without you around. Love ya abya.

List of Figures

List of Tables

Chapter 1

Marry Rich, Poor Girl: The Implications of Sex Selection for Girls' Later Life Outcomes in India

Abstract

Sex ratios at birth have risen steadily over the last three decades across much of the developing world. Many attribute this rise to improved access to sex selection technologies such as ultrasound since 1980. This study seeks to understand the effect of access to ultrasound technology, and the consequently skewed sex ratios, on the later life outcomes of females in India. Existing economic theory and literature view male-skewed populations as a boon to the marital prospects of females. However, Edlund (1999) proposes an (as yet untested) theory that, in environments where hypergamy is practiced and parents derive utility from married children, a male-skewed sex ratio can generate a permanent female underclass. I extend this theory to argue that if sex ratios are skewed disproportionately amongst the rich, as the evidence suggests, then poorer matching in the marriage market can in turn lead to weaker bargaining positions for females. I test this theory and examine its implications for later life outcomes using India-wide household level data on sex ratios, ultrasound use and bargaining power. I first present evidence that sex ratios are disproportionately skewed in favor of males amongst the wealthy relative to the poor in regions with high ultrasound presence. I then demonstrate that parents are indeed considering the sex ratio of their unborn child's future marriage market when determining the sex composition of their own family. Finally, I utilize a difference-in-difference approach to identify the impact of ultrasound access on intrahousehold outcomes of affected women. I find evidence that greater parental access to sex selection technology at a son's birth is related to negative outcomes for his wife along a series of dimensions: lower education and health, greater marriage age and education gaps, and a smaller share of bargaining power as measured by autonomy and child production decision-making. My results are robust to a two-sample 2SLS specification employing distance to a major health center as my instrument. As the first cohort of females affected by ultrasound at birth have only recently entered the marriage market, this study provides timely and compelling empirical evidence of the stratification of sex ratios by wealth and the consequences of this demographic shift on the later life outcomes of females in India.

1.1 Introduction

Sex ratios, or the ratio of males to females in a population, have steadily increased over the past three decades across much of the developing world. This disproportionate and unanticipated rise in the number of males has been largely attributed to the increased availability of technologies that offer couples greater control over the sex composition of their family. Paired with son preference, technologies such as ultrasound allow parents to exercise their underlying preferences to achieve the desired sex composition of children.

The biologically anticipated sex ratio at birth is approximately 105; that is, 105 males to every 100 females. The national census of India in 2011 reported a child sex ratio of 109.4. This already skewed ratio masks considerable heterogeneity across states. While the state of Meghalaya has a sex ratio of 103, the state of Punjab has a sex ratio of 125: for every five male children under the age of five, there exist only four female children in the state of Punjab.

Son preference and the increased ability to exercise such preference through sex selection technologies like ultrasound are oft cited as the primary drivers of such remarkably skewed sex ratios. Ultrasound technology was first introduced to India in 1980, and as the technology quickly became cheaper and more mobile, it proliferated rapidly to more rural areas. Hvistendahl (2011) describes how India's aggressive push to decrease fertility in the 1980s, paired with the population's preference for sons, made ultrasound and sex selective abortion an especially appealing innovation. Demand for ultrasound as a sex selection tool was pervasively high, as it offered parents an 'easier' alternative to controlling the sex composition of their family than the emotionally taxing and socially visible practice of infanticide. Clinics that offered the technology openly advertised its efficacy as a sex selection tool: it was not uncommon in the 1980s and 1990s to find health clinic advertisements of the nature: "Pay 500 rupees today and save 50,000 rupees tomorrow;" with the latter amount referring to the dowry that parents of females must pay to marry their daughters off, and the former referring to the cost of one ultrasound to ensure that parents do not bear daughters (Jacoby 2010).

This study investigates the parental decision making dynamics of such endogenously skewed sex ratios and their implications for the marriage market of affected children. The question is a timely one, as the children of those parents to whom ultrasound was first made accessible are just passing through the marriage market in the most recently available public data in India. Though there has been anecdotal evidence of how such skewed ratios in the marriage market manifest in marriage matches (see Economist 2010 series on gendericide, Hvistendahl (2011), Trivedi and Timmons (2013) in NYTimes, among others), the present study offers a first opportunity to rigorously investigate the impacts of endogenously skewed sex ratios on later life marital outcomes on a more comprehensive scale. Importantly, it is also the first study to examine the underlying decision making process of ultrasound users within the context of marriage market prospects.

The analysis is motivated by a theoretical framework drawn from Edlund (1999), in which the author proposes an (as yet untested) theory that, in environments where hypergamy is practiced and parents derive utility from married children, a male-skewed sex ratio can generate a permanent female underclass. I extend this theory to argue that if sex ratios are skewed disproportionately amongst the rich, as the model implies and the evidence suggests, then poorer matching in the marriage market will in turn lead to weaker bargaining positions for females.

I then employ a rich set of household level data from the most recent wave of the National Fertility and Health Survey (NFHS-III) of India to test the implications of the model. I choose to use the NFHS as it is the only India-wide household level dataset with detailed information on a mother's pregnancies, ultrasound use, and intrahousehold bargaining measures. My analysis proceeds in five stages. First, I present evidence that self-reported ultrasound use, when aggregated at the village level for only first births, can serve as

an exogenous source of variation on parental access to sex selection technology. Note that the endogeneity of skewed sex ratios to the sex preferences of their reference population is a concern which plagues all the existing literature in this nascent field. A region in which sex ratios are heavily male-skewed is, by its very nature, a region where male-preference is high, which is likely to be correlated with later life outcomes in favor of males as well. Similarly, the presence of ultrasound technology may be driven by demand for sex selection technology, which should likewise be correlated with later life outcomes. I argue that a focus on ultrasound use during first births reduces the demand component of my variation; I then proceed to show that reported ultrasound use is not correlated with son preference in early childhood outcomes at the individual or village level, suggesting that even preferences which impact ultrasound use do not manifest themselves in differential later life treatment of sons versus daughters.

Second, I demonstrate that sex ratios are indeed stratified by wealth in my sample, and that my ultrasound exposure measure captures this stratification: in regions where parents have easier access to ultrasound, the sex ratios of the rich are differentially higher than those of the poor. While stratification of sex ratios by wealth has been shown in previous literature, this is the first study to demonstrate the impact of access to ultrasound on such stratification, thereby drawing attention to the consequences that increasing access to such technology may have on the long run composition of males and females in Indian society.

Third, I demonstrate that parents, when determining the sex composition of their own family, take into consideration the anticipated sex composition of their (unborn) child's marriage market. As predicted by the theoretical framework, the magnitude and direction of the response is dependent upon the wealth of the family: in regions where the marriage market is especially skewed towards males, poor parents, using the present marriage market as a proxy for their child's future prospects, respond differentially more than wealthy parents by reducing their degree of sex selection. This is true even after controlling for the intrinsic sex preference of the village, and cannot be explained purely through a story of credit constraints where the poor are simply unable to sex select while the rich sex select according to local son preference. This finding is important: it is the first empirical evidence that parental choice is impacted by marriage market prospects and, cumulatively, may be shifting the female population towards a permanent underclass in regions where sex selection is especially pervasive.

Having presented evidence that my measure of ultrasound exposure is exogenous to demand, sex ratios are stratified by wealth, and parents internalize the marriage market in their family sex composition decisions, I next utilize a difference in difference empirical framework to examine how a family's exposure to sex selection technology affects the later life marital outcomes of their children, and in particular, the intrahousehold bargaining power of their sons' wives.[0] Exposed females (wives) are considered those who were born after 1979 and married to men in regions with high village-level ultrasound exposure between 2000-2006. Recall that the channel through which skewed and socially stratified sex ratios at birth may affect intrahousehold outcomes later in life is three dimensional: first, skewed sex ratios in favor of males implies females are scarce, resulting in a marriage squeeze (Rao, 1999) where the average female will marry a man of older age; second, a marriage market affected by endogenous sex selection involves a composition effect which manifests itself in a lower 'quality' (in terms of wealth, health, education, etc.) of the average female relative to the average male; and third, greater variance in wealth and background across matched males and females may lead to a more uneven distribution of bargaining power within the marriage. I therefore examine a host of

[0]I focus on sons' wives rather than daughters because the DHS only contains data on ultrasound for a woman's village at marriage, and has no data on the woman's natal village. Since the vast majority of women in India marry outside of their natal village (while the vast majority of men remain within their natal village), the ultrasound exposure variation I employ can only be reliably linked to the parental exposure at birth for sons after 1979, not daughters.

outcome measures which describe the compositional, matching, and bargaining mechanisms at play, and find effects consistent with the model's predictions: relative to their unexposed counterparts, exposed females are less educated and in poorer health, have greater marriage age and education gaps with their spouses, and demonstrate lower bargaining power in the form of lower age at first birth, greater abuse, lower autonomy, and being further from their self-reported ideal family size. These results are robust to a placebo check using an earlier birth cohort that could not have been exposed to ultrasound as well as a two-sample two stage least squares estimation employing data on distance to the nearest health center from the NFHS-II as an instrument for access to sex selection technology.

Finally, I consider two alternative explanations to the effects I find: (1) credit constraints among the poor which limit their access to sex selection technology and (2) differential migration of poorer females into wealthy male-skewed marriage markets. While both stories are plausibly at play in the Indian sex selection and marriage market scene, I offer evidence that neither can fully explain the set of effects I have described in the preceding analyses. However, a comprehensive picture of intrahousehold dynamics in a setting of skewed sex ratios in India requires further examination of both stories and should be the topic of future work.

The rest of the paper is structured as follows: in Section 2, I present a brief literature review on the links between ultrasound, sex ratios, and bargaining power. In Section 3, I offer a stylized model of parental choice of sex selection in the social context of India. In Section 4, I describe the data employed for the empirical analysis. In Section 5, I describe a set of analyses which sketch the landscape of ultrasound use decisions in India and then test the question of interest. Section 6 presents robustness checks, and I conclude in Section 7.

1.2 Literature

1.2.1 Effect of ultrasound on sex ratios

The relationship between ultrasound access and sex ratios amongst exposed children is fairly well established in the literature. Meng (2008) investigates the effect of ultrasound diffusion on sex ratios in China using data on technological diffusion from thousands of local newspaper clippings (Local Gazeteer) from 1976-1995. demonstrating a robust relationship between the expansion of ultrasound and an increase in the ratio of males to females at birth. Bhalotra and Cochrane (2010) investigate the same relationship in India through a triple difference in which their three control groups are: (1) births that occur pre-ultrasound, (2) births of the first order, and (3) births that occur after the family has achieved its desired sex mix of children (which they identify as one girl and two boys). The authors find that availability of ultrasound in India generated an increase in the sex ratio differentially for children of later birth order in families not yet at their preferred sex mix. Finally, Rosenblum and Akbulut-Yuksel (2012) study the effect of ultrasound diffusion on sex ratios across states in India using aggregated measures of self-reported ultrasound use as their proxy for diffusion. They find a robust positive relationship between ultrasound exposure and sex ratios in states where the technology was first introduced, but no such relationship in states where it was later rapidly diffused.

1.2.2 Effect of sex selection and skewed sex ratios on bargaining power

Though there are a few oft-cited studies on the relationship between skewed sex ratios and bargaining power, most studies are based in developed countries and none investigate the relationship when son preference and sex selection are the driving mechanisms behind skewed sex ratios.

Positive effects on bargaining power

Angrist (2002) studies the effect of skewed (adult) sex ratios on second generation immigrant marriage rates and female labor force participation, employing changes in first generation immigration policies as an instrument for changing sex ratios. In accordance with a story of increasing demand for marriagable females, he finds that higher sex ratios lead to higher female bargaining power in the household and lower female labor force participation. However, it is difficult to argue that the instrument is entirely exogenous, since different immigrant populations have strong (and varying) sex preferences that are likely to be passed on from the first to the second generation, challenging the exclusion restriction. Chiappori et. al (2002) employ increases in state level sex ratios as a proxy for increases in female bargaining power, demonstrating that couples in states with higher sex ratios see less female labor force participation and more non-income transfers from husband to wife. Abramitzky et. al (2011) investigates the effect of the post-WWII scarcity of French men on marital assortative matching and other marriage market outcomes and finds a similar force at play: scarcity in marriageable men is correlated with men marrying women of relatively higher social class with a smaller age gap and fewer divorces. Lafortune (2009) examines the implications of changes in the marriage market sex ratio on pre-marital investments, and finds that a doubling of the marriage market sex ratio is associated with a 0.5 year increase in educational investment by males, suggesting that individuals anticipate their marital prospects and invest in themselves in order to be sufficiently competitive in their future marriage market. Finally, Wei and Zhang (2011) investigate skewed sex ratios in China and find that households with sons in regions with higher sex ratios have significantly higher savings rates than those with daughters and those with sons in lower sex ratio regions. The authors attribute this effect to families seeking to make their sons wealthier and thus more desirable in regions with particularly competitive marriage markets for males.

Negative effects on bargaining power

There exists no empirical literature that I am aware of which identifies negative impacts of positively skewed sex ratios on female bargaining power. This is largely because skewed sex ratios have not been examined in the context of sex selection, which, because of the endogenous nature of the behavior, is a challenging setting in which to explore the relationship in question. Crucially, populations sex-skewed at birth (rather than through war, labor markets, etc.) are also the most pervasive and thus policy relevant setting in which to investigate effects on bargaining power (Hvestendahl 2011). [1]

Beyond the increasing relevance of the effect of skewed ratios on later life outcomes in an environment with rampant sex selection, the motivation for this present study is derived from two developments in the literature. First, there exists an oft-cited empirical observation that sex ratios are more skewed in favor of males amongst the wealthy than the poor (Bhalotra and Cochrane (2010), Rosenblum and Akbulut-Yuksel (2012), among many others). Second, a provocative theoretical framework proposed by Edlund (1999) describes an equilibrium in a skewed and socially stratified marriage market with son preference in which a permanent female underclass is generated. Specifically, in environments where parents derive a higher utility from married children than unmarried children, prefer that their child marry into a family of equal or higher social/wealth status, and have an inherent son preference (all plausible conditions in the Indian setting), a skewed sex ratio in favor of males can produce a growing female underclass such that, generation after generation, the rich bear increasingly more boys and the poor bear increasingly more girls.

[1] Although Wei and Zhang's (2011) paper is set in an environment in which sex ratios are skewed due to sex selection, the authors do not investigate marital matches or post-marital outcomes.

I propose that these two observations imply growing variance in quality between the sexes in marriage markets that are skewed due to sex selection, leading to poorer assortative matching and thus poorer post-marital outcomes. In particular, because females will increasingly lie in the left tail of the quality distribution, female intrahousehold bargaining power relative to that of males may weaken.[2] A key assumption underlying this mechanism is that parents derive utility from (1) marrying their child off and (2) marrying their child to someone of an equal or higher social status; but they do *not* derive utility from the anticipated bargaining power of that child once married.

1.3 A model of sex selection

I present a stylized model of sex selection and ultrasound use adapted from Edlund (1999).

Consider a society with $N >> 0$ parents of childbearing age. N/2 parents are rich (of equally high wealth) and N/2 parents are poor (of equally low wealth). Each parent has preferences over the sex composition of her family, and must thus decide which sex selection technology, if any, to use. Two technologies exist: ultrasound (denoted US), which is employed while the child is in utero, and infanticide (denoted PS for post selection), which is employed once the parent observes the sex of the child after birth. The set of actions that each parent can take are therefore as follows:

$$A \in \{US_m, US_f, noSS, PS_m, PS_f\}$$

where US_m represents the use of ultrasound to select for a male child, US_f represents the use of ultrasound to select for a female child, likewise for PS_m and PS_f, and $noSS$ represents the use of no sex selection technology.

The cost of ultrasound, which is purely monetary, is C for the rich; the poor, however, are credit constrained and must pay CR for the same technology, with $R > 1$. The cost of infanticide, which is entirely psychological, is E for all N individuals. Because parents must choose their strategy while their child is in utero before observing the sex of the child, the expected cost of infanticide is $0.5E$. Costs are structured such that:

$$C < 0.5E < CR$$

Both technologies are perfect; in other words, once a fee is paid for the use of a technology, a parent is guaranteed to have a child of the sex she desires.[3] Parents derive utility from children through two means: inherent sex preference and the future marital prospects of the child. Thus, while a married son is preferred to a married daughter and a single son is preferred to a single daughter (i.e. inherent son preference), a married daughter is preferred to a single son. However, the society is also one in which hypergamy is practiced, making it is socially acceptable for daughters to marry up in social status but unacceptable for daughters to marry down. Preferences for the rich are thus ordered as follows:

[2]Note that Edlund's theory makes a distinction between sex preferences which are 'irrational' (due to prejudice, sexism, traditional allegiances, etc.) versus those which are 'rational' (due to economic pressures to marry). Of course this line is a blurry one: we might argue that status considerations are equally rational, and perhaps more compelling for the wealthy than the poor. The key exclusion restriction required to test my question is therefore: do the underlying motivations behind sex preference also affect the treatment of female adults in the household/in later life outcomes?

[3]Thus the cost of sex selection for either technology is the expected present discounted value of all future payments of ultrasound or infanticide for as many times as needed until the parent bears a child of the desired sex. Since the gender of a child in utero is effectively random and equally likely to be a male as a female, the expected total cost for either technology should be the same across all parents. This is a simplification a la Edlund (1999).

$$U(m_{RR}) > U(m_{RP}) > U(f_{RR}) > U(m_R) > U(f_R) > U(f_{RP})$$

and for the poor as follows:

$$U(m_{PR}) > U(m_{PP}) > U(f_{PR}) > U(f_{PP}) > U(m_P) > U(f_P)$$

where $U(m_{RR})$ denotes the utility a rich parent derives from a son who marries rich, $U(m_R)$ the utility a rich parent derives from a single son, $U(m_{PP})$ the utility a poor parent derives from a son who marries poor, $U(f_P)$ the utility a poor parent derives from a single daughter, and so forth. In accordance with the social standards of hypergamy, I set $U(f_{RP}) = -\infty$, and further set $U(m_R) = U(m_P) = 0$ for simplicity.

Marital prospects of each child are determined in equilibrium. A parent's expected utility from employing (or not employing) a sex selection technology to bear a child of a particular sex is thus a weighted average of the utilities of a married and an unmarried child of the given sex, weighted by the probability of marriage, less the cost of the technology. Denoting $U(s^0)$ as the utility gained from a child of sex s if unmarried and $U(s^1)$ as the utility gained from a child of sex s if married, with $s \in \{M, F\}$, we can write the expected utilities derived from each sex selection technology as follows:

$$U(US_s) = \pi U(s^1) + (1 - \pi)U(s^0) - CR$$

$$U(PS_s) = \pi U(s^1) + (1 - \pi)U(s^0) - E$$

$$U(noSS) = 0.5[\pi U(f^1) + (1 - \pi)U(f^0)] + 0.5[\pi U(m^1) + (1 - \pi)U(m^0)]$$

where $R = 1$ for the rich and $R > 1$ for the poor. Note that both π, or the probability of the child marrying, and whether $U(s^1)$ represents a marriage between children of the same social strata or different social strata, will be determined in equilibrium.

Figure 1a in the Appendix presents the complete payoff matrix for all strategies of the rich and the poor. What becomes immediately apparent is that for rich parents, sex selecting for daughters is strictly dominated by no selection or sex selection for sons: if all rich parents sex select for daughters, their daughters will have no male counterparts to marry (since there are no rich sons and rich daughters cannot marry down); since unmarried sons are strictly preferred to unmarried daughters, the rich will never sex select for females. Likewise, for poor parents, sex selecting for sons is strictly dominated by no selection or sex selection for daughters: if all poor parents sex select for sons, their sons will have no female counterparts to marry (since there are no poor daughters and hypergamy prevents rich daughters from marrying poor sons); since a married daughter is strictly preferred to an unmarried son, the poor will never sex select for sons.[4] Furthermore, given the costs, the post selection technology is always strictly dominated by the ultrasound technology for the rich, while the reverse is true for the poor. Upon eliminating these strategies and imposing the normalization of $U(m_R) = U(m_P) = 0$, I am left with the following payoff matrix, where the first line within each cell represents the payoff for the poor while the second row represents the payoff for the rich:

[4]Note that this dominant strategy can be weakened (to more closely approximate reality, in which there is some evidence that the poor do sex select for sons to some extent) by introducing heterogeneity among the poor in the value they place on marriage. If we imagine the utility of having married children to be derived from the utility of passing on wealth and status to a future generation (grandchildren), then the poor may plausibly value the activity less than their wealthy counterpart. If son preference is sufficiently strong, the poor may then choose to sex select for a son even if he has no future chance of marriage. I refrain from formalizing this complication as the relevant predictions can be derived from the simpler model.

Figure 1. Simplified Payoff Matrix		
Poor/Rich	US_m	$noSS$
$noSS$	$0.5U(f_{PR})$ $0.5U(m_{RP}) - C$	$0.5[U(m_{PP}) + U(f_{PP})]$ $0.5[U(m_{RR}) + U(f_{RR})]$
PS_f	$U(f_{PR}) - 0.5E$ $U(m_{RP}) - C$	$U(f_P) - 0.5E$ $0.5[U(m_{RR}) + U(f_{RR})]$

From amongst this set of strategies, I arrive at three possible equilibria.

- Equilibrium I: The poor sex select for daughters using the post selection tool, while the rich sex select for sons using ultrasound. This will happen if $U(f_{PR}) > E$ and $U(m_{RP}) - 0.5[U(m_{RR}) + U(f_{RR})] > C$ (eg. if the costs of each technology are sufficiently small to be compensated for by the expected gains in utility from the desired sex).

- Equilibrium II: The poor sex select for daughters using the post selection tool, while the rich do not sex select at all. This will happen if $U(f_{PR}) > E$ and $U(m_{RP}) - 0.5[U(m_{RR}) + U(f_{RR})] < C$ (eg. if the cost for post selection is sufficiently small while the cost for ultrasound to the rich is sufficiently large; recall that this is constrained by the fact that $C < 0.5E < CR$).

- Equilibrium III: Neither the poor nor the rich sex select. This will happen if $U(f_{PR}) < E$ (eg. if the cost for post selection is sufficiently large, the poor will not sex select; given this, it is strictly better for the rich to not sex select regardless of the size of C).

Though highly stylized, this model captures three key points. First, parental decisions on the sex composition of their children is largely dependent on the anticipated (equilibrium) sex ratio in their child's future marriage market. Second, two out of the three possible equilibria generate a society where sex ratios are higher amongst the rich than the poor. Third, the poor have differentially lower sex ratios in areas where the rich sex ratio is especially high (a result predicted by Equilibrium I). Note that a simple story of credit constraints can potentially explain the first two observations: all individuals have inherent son preference, but the rich have access to ultrasound and can thus act on those preferences, generating higher sex ratios in their strata than those of the poor. However, such a story cannot explain the heterogeneity in behavior of the poor predicted by the third observation. Finally, note that the results are robust to the elimination of post selection as an available technology given certain values for R and each utility function. If R is sufficiently small and poor parents derive sufficiently large utility from a married daughter relative to a single son, the results will hold with ultrasound as the only available tool for sex selection.

1.4 Data

My empirical analysis primarily uses the 2005 wave of the National Fertility and Health Survey (NFHS-III) of India. I also employ the 1999 wave, or NFHS-II, briefly. The NFHS is a repeated cross-sectional dataset which surveys a representative sample of ever-married women of childbearing age (ages 15-49) across all states of India. I choose to use this dataset as it is the only India-wide household level dataset with detailed information on a mother's pregnancies (and thus ultrasound use) as well as intrahousehold bargaining measures as collected in a domestic violence module. All analyses are at the mother or child level, as specified within each section below.

1.5 Analysis

1.5.1 Ultrasound use as proxy for past access to sex selection technology

The NFHS-III includes a detailed module on antenatal care during pregnancy and after birth for women who report having been pregnant in the last five years prior to the survey (2000-2005). In this module, women who report having received antenatal check-ups during their pregnancy are asked whether they ever used an ultrasound in the checkup. This study's question of interest would be ideally answered with ultrasound access data from 1980-1989, which is the cohort of births just entering the marriage market in the 2005 NFHS-III. However, no publicly available dataset exists with detailed information on ultrasound access in India, and particularly for this time frame. The next few sections endeavor to make the case that ultrasound use in 2000-2005 is a valid proxy for ultrasound access amongst my population of interest.

I construct an "ultrasound exposure" variable by calculating the village-level averages of reported ultrasound use among first births between 2000-2005. I focus on first births as the literature has suggested that parents limit their sex selection to later births (Hu and Schlosser (2011), Bhalotra and Cochrane (2010)). Figure 2 of the Appendix demonstrates this phenomenon: parents are far more likely to bear a son after having used ultrasound in later births; at first birth, there is no statistically significant relationship between using ultrasound and bearing a son. Figure 2 elucidates two important features of the ultrasound measure available in the NFHS: (1) Self-reported use of ultrasound is an valid proxy for true ultrasound use in a village, since use is correlated with sex selection. (2) Ultrasound is being used to sex select, but not necessarily by families who have a strong preference for sons. Rather, it is being used by families who have (by chance) an 'excess' number of daughters and have not reached their ideal family sex composition. A focus on first births thus minimizes the demand effect captured in the ultrasound exposure measure.

The distribution of the village level average ultrasound measure, as shown in Figure 3 of the Appendix, is heavily skewed to the left: the median (mean) village has 16% (21%) of respondents reporting ultrasound use at first birth, and 28% of villages contain no respondent who reports ultrasound use at first birth.

Ultrasound exposure is not randomly distributed across villages, as is demonstrated in Table 1.1. Individuals who live in villages with higher levels of ultrasound exposure are less likely to be rural, are wealthier and more educated, and have spouses with more professional occupations. These results are not surprising: ultrasound technology arrived first to city centers and major health institutions and proliferated outward from there.

1.5.2 Effects of ultrasound exposure on child outcomes

Despite the non-random distribution of exposure, village-level ultrasound use in 2000-2005 may be a valid measure of access to prenatal sex selection technology in the 1980s if, conditional on observables, average self-reported ultrasound use is correlated with access to ultrasound and only impacts later-life outcomes through this access channel.

In order to address this demand-centered endogeneity concern, I examine the relationship between ultrasound exposure and early childhood outcomes. If aggregate village-level ultrasound exposure is correlated with later life marital outcomes through an underlying parental son preference, such son preference should first manifest itself in parental investments in their offsprings' early life outcomes.

Several recent papers have explored the relationship between the 'wantedness' of a child and early childhood outcomes (Almond et al. 2010, Hu and Schlosser 2011, Line et al. 2009, Shepherd 2008). The present

analysis closely follows that of Almond et. al (2010), which investigates the impact of access to ultrasound on early childhood investments in China. Building on Meng's 2009 analysis of the impact of ultrasound on sex ratios, Almond and coauthors posit that, if sex selection is driven by strong son preference, parental investments in children after birth should move in a predictable direction. Namely, assuming that ultrasound diffusion is not determined by individual preferences, postnatal investments in girls should increase in areas with high diffusion, since parents who choose to bear daughters in such regions (where access is plentiful) must be selected to have relatively lower son preference than their son-bearing counterparts. The direction of prenatal investments, however, is ambiguous: if pregnancies are carried to term, ultrasound can allow for in utero gender discrimination against girls. However, if the child is aborted, one would expect an increase in in utero investments in girls through the same preference sorting mechanism operating in postnatal outcomes.

Almond et. al find no effect of ultrasound access on differences in postnatal investments across sexes. However, they do find that early neonatal mortality increases for girls relative to boys in areas with ultrasound access, which they argue is likely a result of decreased in utero investments in girls relative to boys. The analysis suggests that, while parents' sex selection decisions are motivated by son preference, these preferences no longer manifest themselves conditional on the birth of the child.

Hu and Schlosser (2011) run a similar exercise in the Indian context employing multiple rounds of the NFHS data used in the present study. They exploit time and state-level differences in the sex ratio at birth, their proxy for parental access to sex selection technology, interacted with whether the child is female. They find that higher sex ratios are associated with lower rates of stunting and malnutrition for girls relative to boys, but find no change in girls' mortality or respondents' stated son preference.

I perform a similar exercise with a few key differences: first, I choose to employ the village-level average of reported ultrasound use rather than the sex ratio at birth as my proxy for parental access to sex selection technology. I do so because, in later analyses, I am interested in capturing the broader implications of access to the technology rather than purely those of technology use. The sex ratio at birth captures variation only in the degree to which parents use ultrasound for the purposes of abortion. The ultrasound measure arguably casts a wider net on access, since both parents who do and do not employ the technology for abortion report use. Insofar as this is an exercise in validating the ultrasound exposure measure as a proxy for parental access to sex selection technology, I focus on ultrasound use rather than sex ratios at birth for the present exercise.[5]

My second difference stems from this necessity to use ultrasound at the village level: because NFHS datasets cannot be linked at the village level over rounds, I limit my analysis to the NFHS-III round only.

Thirdly, I focus only on first births for reasons articulated earlier.

The NFHS contains detailed information on prenatal (ultrasound use) and postnatal investments in children only for those born in the last five years, so I restrict my sample accordingly.

To better understand such parent-child dynamics in play, I first run the following simple regression at the mother level for male children and female children separately:

$$ChildInvestment_{ihv} = \alpha_{ihv} + \beta Ultrasound_{hv} + \gamma_{ihv} + \theta_v + \epsilon_{ihv}$$

in which *Ultrasound* is the household-level report of whether the mother used an ultrasound in her most recent pregnancy, and the outcome variable *ChildInvestment* for child i in household h in village v spans both postnatal investments such as breastfeeding and vaccinations as well as proxies for prenatal investments

[5]Note that this is in contrast to the objectives of existing literature in this vein. While they explore the endogeneity of their measures of prenatal sex selection, I aim to identify a measure of access to prenatal sex selection technology which is not endogenous to parental preferences.

such as size of child at birth and weight and height at survey time. The standard set of controls γ_{hv} are included (rural, religion, age of mother, age of child, wealth index, education, and occupation) as well as state fixed effects θ_s.

This regression is plagued with selection issues, but that is precisely what I want to identify: do mothers who report using ultrasound and bear a son display different childcare behaviors than mothers who do not report using ultrasound and bear a son? Likewise for daughters? Any difference in child outcomes by ultrasound use may be due to a host of reasons (mother is more prone to pregnancy complications, has more maternal care knowledge, is more cautious or risk averse, or has strong underlying sex preference). If there is no clear relationship between ultrasound users and child outcomes, then selection on underlying son preferences is an unlikely channel for affecting later life child investments.

Results are presented in Table 2a. Mothers reporting ultrasound use put their child (both sons and daughters) to breast at a later age and bear healthier children of greater weight and height and lower likelihood of anemia. They also breastfeed their daughters conceived after ultrasound for less time, although this does not appear to manifest in the daughter's health outcomes. These results are broadly consistent with the fact that wealthier individuals are more likely to use ultrasound (although the wealth index of the household is controlled for, residual wealth variation may remain), but do not suggest that demand for sons or daughters and the consequent use of ultrasound plays an obvious role in how the user treats her children.

I next replicate the experimental strategy of Almond et al. (2010) with the following regression:

$$ChildInvestment_{ihv} = \alpha_{ihv} + \beta_1 Male_{ihv} + \beta_2 UltrasoundExp_v + \beta_3 Male_{ihv} * UltrasoundExp_v + \gamma_{ihv} + \theta_s + \epsilon_{ihv}$$

in which *UltrasoundExp* is a continuous variable of village-level averages of reported ultrasound use between 2000-2005, and the outcome variable *ChildInvestment* for child i in household h in village v spans both postnatal investments such as breastfeeding and vaccinations as well as proxies for prenatal investments such as size of child at birth and weight and height at survey time. The standard set of controls γ_{hv} are included (rural, religion, age of mother, age of child, wealth index, education, and occupation) as well as state fixed effects θ_s.

Results are presented in Table 2b. Male children see no differences in early childhood outcomes relative to female children in ultrasound-exposed villages, except to be of lower weight at the time of survey. For the most part, coefficients are close to zero or negative and statistically insignificant, suggesting that the measure of village level ultrasound use at first birth does not track parental selection or demand effects of ultrasound use on child outcomes. This latter result may be representative of the sorting that Hu and Schlosser (2011) find, but is not supported by the remaining outcomes.

1.5.3 Changes in the composition of the sexes

Taken cumulatively, the preceding investigation offers compelling evidence that the ultrasound measure I employ is a valid proxy for access to ultrasound technology and does not suffer from demand-driven endogeneity with later life child outcomes. Using this measure, I now explore the demographic landscape of ultrasound exposure. If we take the theory of endogenous sex selection seriously, we should see that sex ratios are skewed only among particular strata of society (the wealthy) and only in regions where sex selection is possible (high ultrasound access regions). I thus estimate the following:

$$ChildSexRatio_{hv} = \alpha_{hv} + \beta_1 UltrasoundExposure_v + \beta_2 UltrasoundExp_v * Wealth_{hv}$$

$$+\beta_3 UltrasoundExp_v * Education_{hv} + \beta_4 UltrasoundExp_v * EducationSpouse_{hv} +$$
$$\beta_5 UltrasoundExp_v * Occupation_{hv} + \beta_6 UltrasoundExp_v * OccupationSpouse_{hv} +$$
$$\gamma_{hv} + \theta_j + \epsilon_{hv}$$

where $ChildSexRatio_{hv}$ is the ratio of the total number of sons to the total number of daughters in household h in village v. $UltrasoundExposure_v$ is a village level average of reported ultrasound use, and all other characteristics are at the household level. γ_{hv} is the standard set of household level controls (education, occupation, spouse's education, spouse's occupation, age, religion), and θ_j represents village fixed effects.

Results are presented in Table 3 and follow the expected directions. A ten percentage point increase in village-level ultrasound exposure is associated with a 1.25 percentage point higher child sex ratio among wealthy families relative to their poor counterparts, implying that wealthy families are producing more sons relative to daughters than are poor families in regions where prenatal sex selection technology is available.

Interactions of village level ultrasound exposure with the education level and occupation of the mother and her spouse also move in a plausible direction: more educated parental units or those in more professional occupations have child sex compositions which are less skewed towards sons. This result may be driven by movement away from traditional customs or preferences by the more educated or greater economic opportunities for female offspring among the more educated/better employed, among many other reasons. It is interesting that education and occupation pull sex ratios in the opposite direction as wealth given that the three characteristics are often packaged together: these results suggest that the residual of wealth that is not explained by education or occupation may be more driven by status or inheritance concerns, compelling this wealthy demographic to sex select vigorously in favor of sons.

The results on education and occupation are also worth noting because previous studies have found otherwise. Namely, Akbulut-Yuksel and Rosenblum (2012) as well as Bhalotra and Cochrane (2010) both find that more educated mothers are more likely to sex select than their uneducated counterparts. Bhalotra and Cochrane find that for a given history of births, more sex selection was conducted in 1995-2005 by women with at least a secondary education and in the top quintile of the wealth distribution. Similarly, Akbulut-Yuksel and Rosenblum find that the probability of bearing a male child is considerably lower among mothers with zero years of education. However, the latter do not control for wealth in their estimation, so the education effect may simply be picking up the wealth effect discussed earlier. Bhalotra and Cochrane's result is more puzzling, since they use the same dataset and same definition of education as is used in Table 3. Note that they identify sex selection off of a triple difference of (1) births that occur post-ultrasound era, (2) births of greater than first order, and (3) births that occur before the family has achieved its desired sex mix. Thus the sex selection effect the authors are picking up is for a very specific pool of families, and may not reflect the more general demographic patterns in sex ratios across the full population of families in the Indian NFHS.

1.5.4 Do parents consider the marriage market in their fertility decisions?

Having presented suggestive evidence of endogenous sex selection - namely that skewed sex ratios in favor of males are indeed concentrated among the wealthy - I now put the theory to a more rigorous test. I examine whether parents actively internalize marriage market conditions when determining the sex composition of their family. If sex ratios are being endogenously skewed within households in anticipation of the marriage market, it is reasonable to posit that parents use the *present* marriage market as a proxy for their own child's future market and actively respond to the conditions they observe by adjusting the sex composition of their

own (still forming) family.

Loosely following the empirical strategy of Edlund and Lee (2011) in South Korea, I estimate the relationship between the within-family child sex composition of individuals born between 2000-2005 (denoted Cohort II) and the village average adult sex composition of individuals born between 1971-1986 (denoted Cohort I), all of whom should be either married or eligible for marriage by 2005. In other words, I regress the fraction of male children born in 2000-2005 at the household level on the contemporaneous village level fraction of males in the marriage market:

$$CohortIIFractionMale_{hv} = \alpha_{hv} + \beta CohortIFractionMale_{v} + \gamma_{hv} + \epsilon_{hv}$$

The coefficient of interest is β. γ_{hv} is a set of household level controls (education and occupation of respondent and spouse, respondent's age, respondent's religion, whether the household is in a rural area, ultrasound accessibility of the village, the village level sex ratio of children born between 1991-1995, and interactions of each of these variables with the fraction of males in the marriage market). I control for the village level sex ratio of children born in 1991-1995 as a proxy for village-specific sex preference.

Results are presented in Column 1 of Table 4a. The sign on the coefficient of interest is significant at the ten percent level and follows the expected negative direction, offering suggestive evidence that parents are internalizing marriage market conditions when choosing the sex composition of their own family: where the sex ratio in the marriage market is high, families adjust their child composition in favor of daughters; where the sex ratio is relatively low, families adjust their composition in favor of sons. Column 2 shows that this response is driven entirely by the poor: the poor sex select differentially less in their own family composition when the existing marriage market sex ratio is high. Note that this cannot be due only to constraints on credit, since their response is negative in levels as well as (sum of the first two rows of Column 2).

One might expect that this adjustment is primarily occurring in villages where families have the ability to actively control their family composition - in other words, villages with ultrasound accessibility. Columns 3 and 4 presents the results upon including an interaction between the fraction of males in the current marriage market (Cohort I, a village level average) and a village level ultrasound exposure variable (Column 3) as well as the additionally the interaction with wealth level (Column 4). Though noisy, estimates follow the anticipated direction: regions with greater ultrasound access see a greater negative response to skewed marriage market sex ratios, and this is differentially larger (more negative) amongst the poorer populations in these villages.[6].

Taken together, these results suggest that parents are indeed considering future marital prospects when determining the sex of their unborn child. The fact that we see this behavioral response only amongst poor families offers compelling support for Edlund's 1999 hypothesis: while wealthy families have an incentive to sex select in favor of males (thus co-moving their child sex ratio with the existing skewed marriage market sex ratio), poor families can only respond to such increases in wealthy males by adjusting their sex composition away from sons.[7]

Taking the theory a step further, I create a measure of the extent to which the village level marriage market is skewed between the rich and the poor. Specifically, I calculate the gap in the marriage market between the fraction of males in the top half of the wealth distribution (the "rich") and the fraction of males in the bottom half of the wealth distribution (the "poor"). If we take the marriage market mechanism

[6]Edlund and Lee (2011) do not empirically investigate the relationship between birth and marriage market sex ratios differentially across wealth groups nor sex selected regions; their analysis stops at the first basic regression and they do not address potential endogeneity concerns.

[7]This does not necessarily mean that the poor sex select in favor of daughters. It simply means that they do not sex select, or they sex select to a considerably lesser degree than the rich, in favor of sons.

seriously, then the larger this gap, the more poor households should react in the sex composition of their own family. I therefore run the following regression:

$$CohortIIFractionMale_{hv} = \alpha_{hv} + \beta CohortIFractionMale * Gap * Poor_{hv} + \gamma_{hv} + \epsilon_{hv}$$

All necessary interactions and controls are included. Results are presented in Column 1 of Table 4b. As expected, the coefficient on the triple interaction of interest is negative and significant: poor households in villages with not only a higher sex ratio in the total marriage market but also a larger gap in sex ratio between the rich and the poor in the marriage market (with the rich bearing significantly more boys) appear to be the key subset to react to marriage market conditions by reigning in their own relative production of boys (by, for example, not sex selecting against girls as actively). Columns 2 and 3 present the same regression results for households in villages with zero reported ultrasound use and households in villages with a positive amount of ultrasound use, respectively. As expected, the effect is coming entirely from areas where families are able to control the sex composition of their family through ultrasound.

There are four caveats to this discussion. First, we might be concerned about the endogeneity of the marriage market variable: villages of a certain type may both draw more young marriageable men to the village and have families with certain son preferences. For example, regions with labor markets that encourage the immigration of working-age men may also encourage the production of males since male labor market returns are relatively higher. Although this is plausible, this endogeneity would bias the coefficient of interest upwards, so any negative coefficient I find could be considered a lower bound on the true response. It is more difficult to think of an endogeneity story that moves in the opposite direction. Perhaps in labor markets where females complement males but the initial work is male oriented (eg. construction of factories which will eventually employ women), poorer families, upon seeing an influx of male workers, may anticipate increased labor market returns for females and thus bear more daughters in villages with a large proportion of males currently in the labor [and thus marriage] market. If ultrasound accessibility is correlated with other characteristics that make a village amenable to factory construction (which is not unlikely - closer to a city, more electricity, etc.) then I may be misattributing the observed negative coefficient to parental responses to the marriage market rather than parental responses to the labor market. I do a rough test of this story by running the following regression:

$$CohortIIFractionMale_{hv} =$$
$$\alpha_{hv} + \beta_1 CohortIFractionMale_v + \beta_2 CohortIFractionMale_v * ManualWork_{hv} +$$
$$\delta ManualWork_{hv} + \gamma_{hv} + \theta_j + \epsilon_{hv}$$

where *ManualWork* is a binary variable which equals one if the respondent's spouse is involved in skilled or unskilled manual work. Though coarse, it is the NFHS classification most closely approximating construction or factory work. Results are presented in Column 5 of Table 4a ; the coefficients on both *ManualWork* and *CohortIFractionMale*ManualWork* are noisy and close to zero.

Second, the relevant marriage market in this context may not be the village; anecdotal evidence (Economist 2010) suggests that there is frequent cross-village communication and intermarriage, especially when there is a dearth of the demographic in demand (eg. marriagable women) in ones own community. Unfortunately, due to limitations in my data, I am unable to identify neighboring villages and thus cannot expand the reach of the marriage market sex ratio to its most plausible radius. One might argue, however, that any endogenous parental response to the small village level marriage market would only be exacerbated if extended to the more relevant broader marriage market.

Third, while this analysis suggests that parents internalize current marriage market conditions, it also suggests that parents are not fully forward thinking. They appear to use the current marriage market as a proxy for their (unborn) child's future marriage market. Thus, as Edlund and Lee (2009) describe, parents may overshoot in their sex ratio adjustments, failing to recognize that all other parents may respond as they do. However, though this may have been the case in the authors' context of South Korea, which saw a steep decline in sex ratios in the course of fifteen years, the magnitudes of the negative coefficients in my analysis are quite small relative to the clear growth in the proportion of male children in ultrasound villages, as demonstrated by columns 3 and 4 in Table 4a. Overshooting does not appear to be a concern.

Finally, the above exercise on marriage market internalization explores the decision of parents who are presently (at the time of the 2005 survey) bearing children and forming a family. My ultimate question of interest, however, deals with young women who (at the time of the survey) are already married, for whom parents would have made the sex selection decision long ago. I cannot run the internalization exercise on this population of parents because, at the time of bearing the daughters in my sample, the marriage market they may have used as a proxy for their child's future market had not been exposed to ultrasound. The introduction of ultrasound had only just occurred, so any endogenous sex selection by such parents would have had to be driven by purely forward thinking, anticipatory motives. Did such endogenous selection occur amongst this earlier generation of parents? I offer some evidence that it did by investigating the differences in the population of their daughters (my sample of interest) amongst the poor and the non-poor across villages of varying ultrasound exposure.

I run the following simple specifications in which my sample is restricted to all children ages 16 to 23 years (the relevant marriage market years) as of 2005:

$$PoorFractionGirls_v = \alpha_v + \beta_1 UltrasoundExposure_v + \gamma_v + \epsilon_v$$

$$RichFractionGirls_v = \alpha_v + \beta_2 UltrasoundExposure_v + \gamma_v + \epsilon_v$$

$$DifferenceinFractions_v = \alpha_v + \beta_3 UltrasoundExposure_v + \gamma_v + \epsilon_v$$

where $PoorFractionGirls_v$ is defined as the fraction of girls who are poor out of all poor children ages 16 to 23 in the village, $RichFractionGirls_v$ is the fraction of girls who are rich out of all rich children in the village, and $DifferenceinFractions_v$ is the difference between the two. γ_v represents village level controls, namely average wealth, education, religion, and whether it is rural or urban. Results are presented in Appendix Table 1.1.

The coefficients move in the direction predicted by a story of endogenous sex selection. Girls make up a substantially larger fraction of all poor children aged 16-23 years (a one percentage point increase in village level ultrasound exposure is associated with a one percentage point decrease in the sex ratio amongst poor families) and a smaller fraction of all wealthy children aged 16-23 years in villages with higher ultrasound exposure.[8] Consistent with the first two specifications, column 3 reflects a positive relationship between ultrasound access and the gap in the fraction of girls among the poor and the rich: villages with greater

[8]It is worth stressing that the strong and significant relationship with ultrasound access arrises amongst poorer families rather than wealthier families, again suggesting that the behaviors we observe cannot be driven purely by credit constraints. A credit constraints story would suggest that in regions where ultrasound is available, wealthy families will use it readily (boosting their production of males) while the poor will be unable to afford the technology and remain unchanged in their production of males. Instead, we see that in regions where ultrasound is available, wealthy families remain unchanged, while poor families reign in their relative production of boys. This is plausible if, for example, the wealthy have alternative means of prenatal sex selection in low-ultrasound exposure villages (as measured in this study), or they are less likely to report ultrasound use when used for sex selection.

ultrasound access saw (in the 1980s) not only more sex selection in favor of males by the rich, but also a decrease in sex selection for males by the poor. Although by no means conclusive independently, the results do suggest that even as far back as 1982, soon after ultrasound was introduced, parents were differentially sex selecting the composition of their family based on their socioeconomic position, potentially due to the the anticipation of their unborn child's future marital prospects.

1.5.5 Effect of ultrasound exposure on marital and intrahousehold outcomes

Having presented evidence that my measure of ultrasound exposure is reliable and that parents internalize the marriage market in their family sex composition decisions, I now turn to the primary question of this study: how does a family's exposure to sex selection technology in the form of ultrasound affect the later life marital outcomes of their children, and in particular, the intrahousehold bargaining power of their daughters?

Recall that the channel through which skewed and socially stratified sex ratios at birth may affect intrahousehold outcomes later in life is three-fold: first, skewed sex ratios in favor of males implies a larger number of unmarried males for a given pool of females, which may lead to men reaching into younger pools of women in search of a spouse (or alternatively, men marrying at older ages); second, a marriage market affected by endogenous sex selection involves a composition effect which manifests itself in greater variance in wealth across males and females, which, third, may lead to poorer match quality in terms of socioeconomic background even within a narrow age range.

I employ the following outcome measures to examine marital and intrahousehold dynamics: BMI, weight, height, education of the respondent, the age gap between the respondent and her spouse, the respondent's marriage age, whether she reports any physical abuse or beating explicitly, whether she reports needing permission from her husband for healthcare decisions, household purchases, or family visits (coded cumulatively as 'lack of autonomy'), her age at her first child birth, the number of years between her marriage and her first birth, the number of children she has borne, the fraction of male children borne, and the distance between her stated ideal number of children and actual number of children (subtracting the former from the latter).

I employ BMI, weight, height, and education as measures of the 'quality' of the married woman, or a proxy for her wealth prior to entering the marriage. Taking the theory seriously, we should expect that the average male in a high ultrasound region is marrying a poorer woman than the average male in a low ultrasound region.[9] Note that, although I control for the wealth of the household in all figures and regressions to follow, this wealth control is a closer proxy to the wealth of the husband, and so we should expect to see effects of ultrasound on proxies for pre-marital female wealth despite (or rather, given) the inclusion of this control.

I then show, conditional on this change in the composition of marriageable women, the impact of being born in an ultrasound regime on various measures of female bargaining power within a marriage. I examine the outcomes of marriage age, marriage age gap, physical abuse, autonomy, age at first child birth, and the number of years between marriage and first birth. The age metrics are commonly used in the literature as proxies for bargaining power; I take advantage of the unique domestic violence module of the NFHS to also include direct measures of intrahousehold dynamics such as reports of physical abuse, beating, and lack of autonomy. Finally, the number of children and the fraction of male children borne by the respondent are used to demonstrate the intermediate effects of ultrasound: firstly as a pregnancy health tool (leading to lower infant mortality and thus more children) and secondly as a sex selection tool. The number of children

[9]Recall that, as demonstrated in Tables 4a and 4b, poor families in high ultrasound regions react more, or bear differentially more females than their counterparts in low ultrasound regions when compared to the wealthy families in each region.

borne may also be a measure of bargaining power through the contraception channel: women with lower bargaining power have less control over their own fertility decisions. To measure the extent of this effect, I compose an additional outcome measure of the distance between the respondent's stated ideal family size and her actual family size.

1.5.6 Pre-trends

I argue that, controlling for observables, the village average of self-reported ultrasound use between 2000-2005 has no effect on the contemporaneous marriage market and intrahousehold dynamics except through its relationship with village-level ultrasound access after 1979.

For a grasp on potential differential trends across villages, Figure 1 presents a series of graphs plotting the coefficient of interest, β_t, from the following specification:

$$LaterLifeOutcome_{ivt} = \alpha_{ivt} + \sum_{t=1971}^{1987} \beta_t UltrasoundExp_v * d_{ivt} + \gamma_{iv} + \theta_v + \delta_{st} + \epsilon_{ivt}$$

where $UltrasoundExp_v$ is a continuous variable of village-level averages of reported ultrasound use, d_{ivt} is a dummy variable indicating whether individual i in village v was born in year t, γ_{iv} represents a set of individual level controls (wealth index, religion, education, occupation, spouse's education, spouse's occupation), θ_v is village fixed effects, and δ_{st} is state by year fixed effects. For bargaining power outcomes, I also include in γ_{iv} controls for BMI, weight, and height in order to examine bargaining outcomes conditional on compositional changes.

Note that the outcomes are measured in 2005 but plotted along the birth year cohort dimension (so individuals born in 1980 are 25 years old when reporting the outcome of interest). If my identifying assumption holds, then, controlling for relevant observables, trends in the outcome should be comparable across all villages prior to 1980 (when ultrasound was not available in India and could thus not affect the later marriage market) and should diverge according to ultrasound access gradually from 1980 onwards.

Figure 1a presents the trends for outcomes on composition. Both BMI and weight follow comparable trends prior to the 1980s but exhibit sharp drops afterwards amongst exposed females. There appears to be no affect on height. Like weight, education also maintains comparable trends across regions of varying ultrasound exposure prior to 1980 but decreases differentially in more exposed regions after 1980.

Figure 1b presents the trends for outcomes on matching and bargaining. While the education gap trend is noisy, the marriage age gap clearly increases after 1979 amongst exposed females. This is concurrent with a decrease in the average marriage age of female cohorts born after 1979 in exposed regions.[10] The average age of a woman at her first child birth also decreases after exposure, and lack of autonomy, the number of children, and the distance from the respondents ideal family size show marked increases after 1979.

1.5.7 Primary difference-in-difference specification

This section presents the difference-in-difference regression analog of the previous trends analysis:

$$LaterLifeOutcome_{ivt} = \alpha_{ivt} + \beta_1 UltrasoundExp_v * BornPost1979_{ivt} + \gamma_{iv} + \theta_v + \delta_{st} + \epsilon_{ivt}$$

[10] Ex ante, it is not clear whether an excess of wealthy marriageable men/higher demand for marriageable females should generate such trends. Younger women are also often viewed as more desirable (of higher value), which would go against the typical supply and demand story. However, these trends do match much anecdotal evidence (Economist 2010, NYTimes 2013) in which young girls who would have otherwise not been eligible for the marriage market are now being married off due to their high demand, and men who would have otherwise married a woman of closer age are reaching into younger and younger cohorts in search of a partner.

where $UltrasoundExp_v$ is a continuous variable of village-level averages of reported ultrasound use, $BornPost1979_{ivt}$ is a binary variable indicating whether individual i in village v was born after 1979, γ_{iv} represents a set of individual level controls (number of antenatal visits, wealth index, religion, education, occupation, spouse's education, spouse's occupation), θ_v is village fixed effects, and δ_{st} is state by year fixed effects. In order to limit selection effects, I limit the sample to married women born between 1976 and 1987, though results are robust to extending the time frame in both directions.

Results are presented in Tables 5a, 5b, and 5c. Table 5a presents outcomes on composition, or the 'quality' of the female. Women born after 1979 in villages with higher ultrasound exposure have lower BMI and weight and are less educated than their counterparts. There are no differential changes in height amongst exposed women.

To examine whether this change in the composition of males and females is manifested in the marriage match, I run the following regression:

$$HusbandEducation_{ivt} = \alpha_{ivt} + \beta_1 UltrasoundExp_v * BornPost1979_{ivt} * WifeEducation_{ivt} + \gamma_{iv} + \theta_v + \delta_{st} + \epsilon_{ivt}$$

I can examine matching by quality only along the education dimension because education is the only pre-marriage proxy for quality that is available for both husbands and wives in the NFHS. Results are presented in Table 5b. The coefficient of interest, β_1, is negative and statistically significant: women with lower educations in exposed regimes are married to relatively more educated husbands than their counterparts in unexposed regimes.

Table 5c presents results for bargaining power and intrahousehold dynamics, controlling for compositional changes. I find that married women born after 1979 in regions with greater exposure to ultrasound marry earlier, have a larger age gap with their husbands, report higher likelihood of beating and lower levels of autonomy in their marriage, have their first child at an earlier age (a partially mechanical result of getting married earlier), and have significantly more children than their stated ideal. Coefficients on both reports of physical abuse and the interval between marriage and first child move in the direction of lower bargaining power but are noisy.

Exposed women also have significantly larger families, which is consistent both with the primary purpose of ultrasound as a maternal health tool as well as lower bargaining power (as evidenced by the distance from their stated ideal). They also have a significantly larger proportion of sons (or a significantly smaller proportion of daughters) than their counterparts, expositing the use of ultrasound as a sex selection tool.

1.6 Robustness Checks

1.6.1 1970 Placebo

I conduct a placebo check by running the same specification for the sample of married women born between 1966-1975 with treated women denoted as those born after 1970 in higher ultrasound villages (as of 2005). Given that ultrasound was not available in India during this time period, I should see no effect of the ultrasound and birth cohort interaction on my outcomes of interest. Results are presented in Tables 6a and 6b, by composition and bargaining outcomes respectively. Most coefficients are small and statistically insignificant. Marriage age, age gap, age at first child birth, and the interval between marriage and first birth are all positive and statistically significant. This may be driven by underlying difference in wealth and empowerment across villages of varying [future] ultrasound access; however the coefficients move in the

opposite direction of what is predicted by the selection on sex preference story, offering evidence against the likelihood that differential trends in my primary specification will bias my results upwards.

1.6.2 Selection on observables

In order to get an upper bound on how much of the observed effects in the primary difference in difference specification may be driven by selection, I run an exercise adapted from Altonji et. al (2002) in which the authors develop a method of assessing the amount of selectivity bias in an estimate. Bellows and Miguel (2008) employ this method in a manner I closely follow.

The process involves deriving the ratio of the selection on unobservables (omitted variables) to the selection on observables that would be necessary to attribute the entire estimated effect to selection bias. Given a rich enough set of controls, a large ratio implies that it is implausible that the estimated effect can be fully explained away by unobserved selection.

In this particular setting, I define unobservables to be the underlying sex preferences of individuals and/or alternative village-specific time trends that can impact later-life outcomes of children. For each of my later life outcomes of interest, I calculate the following ratio, where $\hat{\alpha}$ represents the coefficient on *UltrasoundExp*BornPost1979* in a regression with and without controls:

$$SelectionRatio = \frac{\hat{\alpha}_{controls}}{\hat{\alpha}_{controls} - \hat{\alpha}_{nocontrols}} \tag{1.1}$$

The ratio for marriage age gap is 7.88; for marriage age is 2.76; for beating is 2.14; for lack of autonomy is 1.4, and for age at first birth is 13.25. All others are less than one. To interpret, a ratio of 7.88 implies that underlying sex preferences must influence the marriage age gap 7.88 times more than the full set of observable controls in order to attribute the entire estimated effect to selection bias.

As a benchmark, Altonji et. al (2002) regard a ratio of 3.55 as strong evidence that unobservables are unlikely to explain away their entire effect. Given the evidence I have presented above, especially as related to parental decisions on child investment, I argue that even a ratio of 1 is evidence that unobservables are unlikely to explain away the entire effect: it is hard to conceive of a story in which underlying sex preferences would affect the outcomes of interest equal to that of strong predictors such as wealth, education, occupation, and all village-level characteristics of one's birthplace.

1.6.3 Two sample 2SLS: Access to a major health facility

To further ease concerns about the endogeneity of the ultrasound exposure variable, I run a two sample two stage least squares estimation using NFHS II data and access to a major health facility as my instrument. The two sample 2SLS is useful for two reasons: (1) Access to a major health facility is more obviously exogenous to sex preferences (controlling for observables) than self reported use of ultrasound, and (2) as the second and third rounds of the NFHS lack detailed geographical information, I cannot link the two surveys at the village level and thus cannot directly link ultrasound exposure measures from 1999 (which would be a better proxy for access to ultrasound post 1979) with marriage market outcomes from 2005. Using the two sample 2SLS design allows me to employ a proxy measure for ultrasound access in 1999 and link it to marital outcomes in 2005.

I run the following reduced form regression on the NFHS III dataset:

$$LaterLifeOutcome_{ivt} = \alpha_{ivt} + \beta_1 UltrasoundExp * BornPost1979_{ivt} + \gamma_{iv} + \theta_v + \delta_{st} + \epsilon_{ivt}$$

in which *UltrasoundExp*BornPost1979* is estimated from the NFHS II dataset in the first stage regression as follows:

$$UltrasoundExp * BornPost1979_{ivt} = \alpha_{ivt} + \beta_1 HealthFacilityAccess_{iv} * BornPost1979_{ivt} + \beta_2 HealthFacilityAccess_{iv} + \gamma_{iv} + \psi_s + \epsilon_{ivt}$$

HealthFacilityAccess is a continuous village level average of a binary variable on health facility use: households report where their most recent child was delivered and are coded as having access to a major health facility if they report doing so in a government hospital, government dispensary, UHC, PHC, non-governmental clinic, or private hospital.[11] I choose to use place of delivery as a proxy for access because delivery is the most serious medical condition that the NFHS inquires about with regards to place of care.

The reduced form equation includes village (θ) and state by year (δ) fixed effects, and the first stage includes state (ψ) fixed effects. Controls in the reduced form include the standard set described in the earlier specifications, while the first stage includes controls for the respondent's education, occupation, education and occupation of the respondent's spouse, and socioeconomic status living index (the NFHS-II version of the wealth index in NFHS-III). My sample consists of all respondents born between 1976 and 1984. I am required to decrease the upper range of birth years from that of earlier specifications because respondents in the sample used to estimate *Ultrasound*Post* can be born no later than 1984 in order to be eligible to participate in the 1999 NFHS II survey.

Results are presented in Table 7. Standard errors are corrected for the two sample 2SLS procedure following Inoue and Solon (2008). BMI, education, marriage age, and age at first birth are significantly lower for women in treated villages (those born after 1979 in villages with greater ultrasound access as instrumented for by access to a major health facility) than their counterparts. Marriage age gap, physical abuse, beating, lack of autonomy, number of children, and distance from ideal family size are significantly higher for women in treated villages than their counterparts, once again suggesting lower bargaining power for women in villages heavily exposed to ultrasound technology. All other coefficients move in the expected direction but are not statistically significant.[12]

1.7 Conclusion

The above investigation offers compelling evidence that the introduction of ultrasound has had unintended consequences on the intrahousehold bargaining dynamics of children in skewed and endogenously sex selected marriage markets. I present evidence that parents, when choosing the sex composition of their families, are informed by the conditions of the contemporaneous marriage market (which they presumably use as a proxy for the future marriage market of their unborn child). Where a sex selection technology such as ultrasound exists, this behavior generates skewed sex ratios which are stratified by wealth: in regions where the rich bear relatively more sons, the poor react by bearing relatively more daughters.

I argue that upon reaching marriageable age, this landscape of skewed and endogenously sex selected populations manifests in poorer intrahousehold bargaining outcomes for females through the mechanism of poorer assortative matching in the marriage market. I offer evidence that the 'quality' of the average female

[11] Households are denoted as not having access to a major health facility if they delivered their child in any of the following: a subcentre, a government mobile clinic, a private clinic that is not a hospital, or in someone's home.

[12] While access to a major health facility may be correlated with other village-specific time trends, all obvious correlates would predict outcomes moving in the opposite direction of what we see. For example, health facility access may be correlated with increasing wealth, urbanization, or delivery of other public services, all of which should have positive effects on exposed female cohort health and wellbeing.

in regions where ultrasound availability is high is considerably poorer than that of her counterpart in a low ultrasound regions, reinforcing the predicitons of socially stratified sex selection. I present further results that female bargaining power within a marriage, as measured by detailed data on autonomy, abuse, and childbearing dynamics, is significantly lower in regions with high ultrasound exposure relative to that in low exposure areas.

I offer two caveats to this investigation. First, the work thus far cannot differentiate between altered marriage market composition due to parental choice of locals from altered marriage market composition due to the migration/trafficking of marriageable women across villages. This is a legitimate concern for welfare measures, since the former scenario would suggest a drop in total welfare while the latter may not, depending on the migrant woman's marital conditions at her place of birth. Unfortunately, I am limited by the available dataset in the extent to which I can explore this issue: the NFHS does not ask questions on place of birth nor distance between husbands' and own village. Using differences in native language as a proxy for marrying a migrant, I find no effect of ultrasound exposure on the likelihood of marrying a migrant. Admittedly this estimate is noisy and offers less than conclusive evidence against the migration story, which should be further explored in future work.

Second, it is important to reiterate the generation gap in the various forms of evidence I have presented. I offer compelling evidence that parents of the present (2000-2005) generation are sex selecting based on anticipated marriage market conditions, but have less compelling evidence of such motives among parents of the 1980-1989 birth cohort. However, I still find considerable impacts on matching and intrahousehold bargaining power for this birth cohort upon marriage in the direction consistent with the theory. Therefore, it may be that much of the effects observed for this cohort of interest is driven by the aforementioned credit constraints story: the rich have greater sex preferences due to status considerations as well as greater access to ultrasound due to wealth, so they sex select more, producing more wealthy boys and a dearth of comparable girls. The poor remain the same given credit constraints, sex selecting only when they have access, and perhaps sex selecting less also due to preferences. Given the overall scarcity of wealthy, marriageable girls, wealthy men marry younger and poorer women (than their counterfactuals in less sex-selective areas). The lack of wealthy women and the inability to match assortatively generate the results presented. I choose to still stress the evidence on parental internalization of the marriage market because it demonstrates that the theory of endogenous selection by wealth strata will kick in with greater force during the present generation of just-born children, resulting in an inevitable exacerbation of the intrahousehold dynamics observed in this study once this generation reaches marriage age.

Figures

Figure 1-1: Composition Outcomes

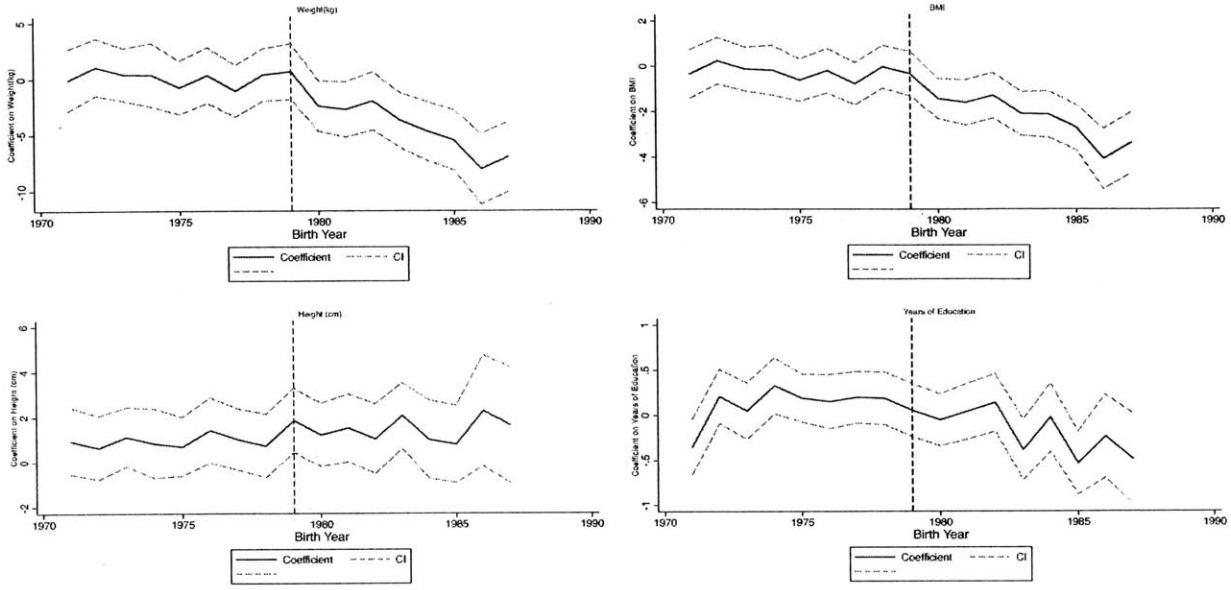


Figure 1-2: Matching Outcomes

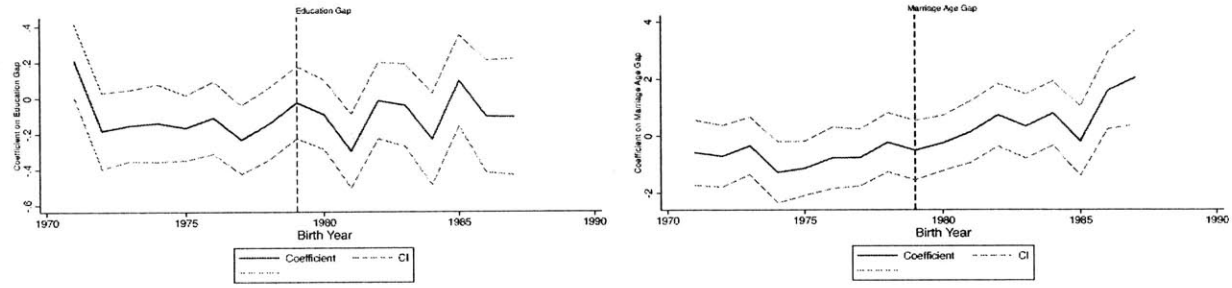
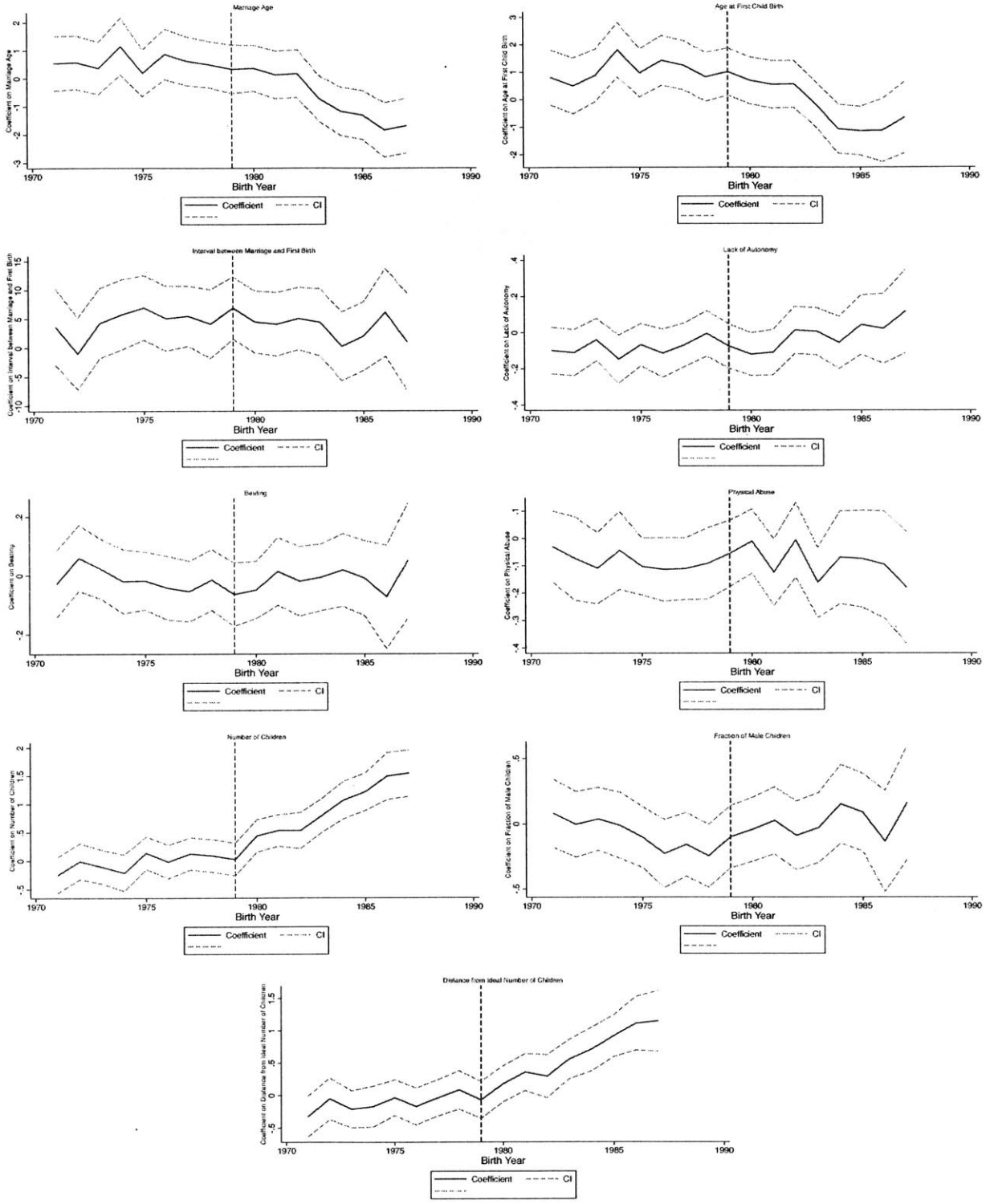


Figure 1-3: Bargaining Outcomes



Tables

Table 1.1: Ultrasound exposure and demographic characteristics

VARIABLES	Ultrasound exposure for first births (village level)
Rural	-0.0853***
	[0.00571]
Wealth index	0.0274***
	[0.00145]
Education	0.0163***
	[0.00119]
Education of spouse	0.00577***
	[0.000944]
Occupation	0.000699
	[0.000567]
Occupation of spouse	0.00365***
	[0.000611]
Religion	0.000121
	[0.000101]
Constant	0.128***
	[0.00707]
Observations	92,658

Standard errors in brackets, clustered at household level.

*** p<0.01, ** p<0.05, * p<0.1

Regression includes state fixed effects.

Table 1.2: Relationship between self reported ultrasound use and parental treatment after birth

	Males			Females		
Outcomes	Coefficient on ultrasound	Outcome mean	N	Coefficient on ultrasound	Outcome mean	N
Breastfeeding duration	-0.250	13.15	23,685	-0.670***	12.81	21,816
	[0.165]			[0.180]		
Ever received vaccination	-2.00e-05	0.87	15,453	0.00483	0.87	14,481
	[0.00933]			[0.00990]		
Age at which put to breast	6.816***	20.74	17,849	5.621***	21.21	15,355
	[1.067]			[1.076]		
Post natal check received	0.0111	0.09	7,637	0.00365	0.08	6,915
	[0.0408]			[0.0365]		
Size at birth (quintiles)	-0.0141	2.99	23,329	-0.00940	2.97	21,484
	[0.0210]			[0.0224]		
Current child weight (kg)	0.126**	9.98	21,443	0.202***	100.15	19,688
	[0.0518]			[0.0555]		
Current child height (cm)	0.301	81.35	21,374	0.588***	814.59	19,613
	[0.208]			[0.226]		
Anemia	-0.0643**	1.05	17,677	-0.0889***	2.19	15,874
	[0.0252]			[0.0287]		

Standard errors clustered at household level.

*** p<0.01, ** p<0.05, * p<0.1

Controls include mother's age, child's age, household wealth index, religion, mother's education and occupation, and male interacted with the previous controls. Village FE also included.

Table 1.3: Comparing child-related outcomes for children born after 1999 across villages with varying exposure to ultrasound

	Parental Treatment						Child Outcomes			
	(1)	(2)	(3)	(4)	(5)	(6)	(11)	(12)	(13)	(14)
VARIABLES	Breastfeeding duration	Age put to breast	No treatment in case of diarrhea	Postnatal check	Ever received vaccination	Number of tetanus shots	Size at birth	Weight at survey age	Height at survey age	Has anemia
Male child	-0.309*	-0.104	-0.00434	0.0109	0.00587	0.0634**	0.0734***	0.519***	1.591***	-0.00662
	[0.174]	[0.959]	[0.0658]	[0.0124]	[0.0100]	[0.0287]	[0.0214]	[0.0477]	[0.204]	[0.0251]
Male child * Ultrasound village	0.0298	1.305	-0.0959	-0.0709	-0.0242	-0.0480	-0.0207	-0.259**	-0.495	0.0733
	[0.371]	[1.954]	[0.143]	[0.0501]	[0.0192]	[0.0489]	[0.0437]	[0.106]	[0.421]	[0.0543]
Age of mother	0.0411***	-0.0204	-0.00582**	-0.000205	0.000100	-0.0119***	0.00470***	0.00969***	0.0614***	-0.00245**
	[0.00786]	[0.0395]	[0.00296]	[0.000478]	[0.000500]	[0.00115]	[0.000976]	[0.00220]	[0.0101]	[0.00114]
Child age	2.112***	-0.166	0.00297	-0.00297	0.0318***	0.0137***	0.0205***	1.846***	8.430***	-0.178***
	[0.0296]	[0.143]	[0.0109]	[0.00219]	[0.00159]	[0.00417]	[0.00286]	[0.00752]	[0.0326]	[0.00386]
Wealth index	-0.0962	-1.263***	-0.00974	0.00679	0.0162***	0.111***	0.0409***	0.211***	0.726***	-0.0735***
	[0.0594]	[0.318]	[0.0241]	[0.00425]	[0.00351]	[0.00898]	[0.00742]	[0.0161]	[0.0692]	[0.00859]
Religion	-0.00166	-0.0183	-0.000633	5.97e-05	0.000459	-0.00105	-0.00110	0.000370	-0.00216	6.42e-06
	[0.00533]	[0.0249]	[0.00155]	[0.000277]	[0.000452]	[0.000848]	[0.000783]	[0.00154]	[0.00673]	[0.000694]
Education of mother	0.0244	-1.517***	-0.0127	0.0104**	0.0184***	0.0926***	0.0226***	0.131***	0.324***	-0.0555***
	[0.0653]	[0.366]	[0.0262]	[0.00517]	[0.00379]	[0.0101]	[0.00832]	[0.0186]	[0.0802]	[0.00982]
Occupation of mother	0.00869	-0.133	-0.00428	-0.00155	0.00290	-0.00289	0.00543	0.0243*	0.0588	-0.00193
	[0.0438]	[0.202]	[0.0150]	[0.00432]	[0.00238]	[0.00518]	[0.00501]	[0.0133]	[0.0511]	[0.00626]
Constant	8.146***	27.44***	0.616***	0.0750***	0.733***	1.704***	2.666***	5.273***	60.50***	1.802***
	[0.269]	[1.346]	[0.0971]	[0.0158]	[0.0157]	[0.0386]	[0.0327]	[0.0749]	[0.321]	[0.0382]
Observations	46,432	33,903	4,253	14,807	30,546	34,111	45,718	41,937	41,786	34,226

Standard errors clustered at household level. Village level fixed effects included; Ultrasound measure is for first births only at the village level.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.4: Composition of the sexes

VARIABLES	Child Sex Ratio (Household level)
Ultrasound exposure	-0.0159
	[0.0534]
Number of antenatal visits	-0.0147***
	[0.00268]
Ultrasound*Wealth	0.125**
	[0.0539]
Ultrasound*Education	-0.0278
	[0.0269]
Ultrasound*Education of spouse	-0.101***
	[0.0282]
Ultrasound*Occupation	-0.0290**
	[0.0128]
Ultrasound*Occupation of spouse	-0.0208
	[0.0139]
Wealth	-0.0753***
	[0.0151]
Education	-0.0875***
	[0.00769]
Education of spouse	-0.0118*
	[0.00711]
Age	0.0237***
	[0.00108]
Occupation	-0.00190
	[0.00455]
Occupation of spouse	0.000152
	[0.00423]
Rural	0.000240
	[0.000493]
Religion	0.0241**
	[0.0103]
Constant	0.319***
	[0.0231]
Observations	50,934

Standard errors in brackets, clustered at village level.

*** p<0.01, ** p<0.05, * p<0.1

State fixed effects included.

Table 1.5: Parents internalizing the marriage market

VARIABLES	(1)	(2)	(3)	(4)	(5)
	Fraction of male children born in 2000-2004 (household level)				
Fraction of males in marriage market (village level)	-0.0256*	0.0183	0.0633	0.0540	-0.0251
	[0.0155]	[0.0253]	[0.0576]	[0.0604]	[0.0200]
Marriage market males * Poor		-0.0242**	-0.0334**	-0.0240	
		[0.0111]	[0.0163]	[0.0212]	
Marriage market males * Ultrasound			-0.0984	-0.0736	
			[0.0601]	[0.0798]	
Marriage market males * Poor * Ultrasound				-0.0175	
				[0.0425]	
Poor * Ultrasound				0.00168	
				[0.0230]	
Ultrasound (village level avg.)	0.0157		0.0676*	0.0628	
	[0.0151]		[0.0347]	[0.0449]	
Poor	-0.00482	0.00736	0.0131	0.0117	-0.00560*
	[0.00304]	[0.00662]	[0.00916]	[0.0110]	[0.00295]
Marriage market males * Spouse manual labor					-0.00151
					[0.0334]
Spouse manual labor					-0.00223
					[0.0184]
Constant	0.537***	0.518***	0.489***	0.489***	0.543***
	[0.0178]	[0.0202]	[0.0341]	[0.0352]	[0.0186]
Observations	31,364	31,365	31,364	31,364	31,365

Standard errors in brackets, clustered at village level

*** p<0.01, ** p<0.05, * p<0.1

Sex ratios calculated for 2000-2005 births (sex ratio of children born in those years) and marriage market (individuals eligible in those years) only.

For marriage market, youngest person observed is 14 in 2000, oldest observed is 34 in 2005.

Full set of controls included (and not shown): village level sex ratio of children born between 1991-1995, rural, education of respondent, education of spouse, religion, and for columns (3)-(4), interactions of each of these variables with the fraction of males in the marriage market.

Table 1.6: Parents internalizing the marriage market across wealth

	(1)	(2)	(3)
	Fraction of male children born in 2000-2005 (household level)		
VARIABLES	Full sample	No ultrasound	Positive ultrasound
Marriage market males * Gap * Poor	-0.147**	-0.0889	-0.156**
	[0.0705]	[0.256]	[0.0749]
Marriage market males * Gap	0.234	-0.0933	0.225
	[0.151]	[0.695]	[0.156]
Gap * Poor	0.0654*	0.000468	0.0759*
	[0.0374]	[0.137]	[0.0397]
Marriage market males * Poor	-0.0240	0.137	-0.0250
	[0.0258]	[0.0993]	[0.0261]
Marriage market males	0.159*	-0.343	0.0651
	[0.0860]	[0.300]	[0.0772]
Gap	-0.112	0.0777	-0.111
	[0.0807]	[0.368]	[0.0833]
Poor	0.00932	-0.0593	0.00794
	[0.0141]	[0.0529]	[0.0144]
Marriage market males * education	-0.00251	0.0488	-0.0191
	[0.0316]	[0.0985]	[0.0329]
Ultrasound (village level)	0.352***	na	na
	[0.115]	na	na
Constant	0.450***	0.741***	0.506***
	[0.0496]	[0.165]	[0.0452]
Observations	16398	1532	14866

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Controls not shown include marriage market males * education, rural, education, education of partner, and re

Table 1.7: Difference in difference for composition outcomes

VARIABLES	(1) BMI	(2) Weight (kg)	(3) Height (cm)	(4) Educ. (years)
Ultrasound village * Born post 1979	-1.307***	-2.904***	0.358	-0.121***
	[0.221]	[0.548]	[0.358]	[0.0417]
Wealth index	0.452***	1.338***	0.445***	0.257***
	[0.0270]	[0.0681]	[0.0495]	[0.00643]
Religion	0.000631	-0.00179	-0.00784	-0.00167**
	[0.00319]	[0.00723]	[0.00594]	[0.000656]
Education	0.152***	0.587***	0.353***	na
	[0.0296]	[0.0766]	[0.0544]	na
Spouse's education	0.108***	0.454***	0.349***	0.306***
	[0.0279]	[0.0731]	[0.0546]	[0.00671]
Occupation	-0.0644***	-0.156***	0.0176	0.0295***
	[0.0200]	[0.0516]	[0.0339]	[0.00412]
Spouse's occupation	0.0856***	0.226***	0.0499*	0.0471***
	[0.0177]	[0.0454]	[0.0292]	[0.00375]
Constant	18.30***	40.29***	149.1***	-0.0705**
	[0.137]	[0.342]	[0.251]	[0.0320]
Observations	29,356	29,403	29,396	30,680

Standard errors clustered at village level.

*** p<0.01, ** p<0.05, * p<0.1

Village and state by birthyear fixed effects included. Outcomes observed in 2005 from NFHS III survey. Sample includes only married women born between 1976 and 1987.

Table 1.8: Matching gap in marriage market of exposed females

VARIABLES	Husband's education
Ultrasound village * Born post 1979	0.137*
	[0.0830]
Ultrasound * Post * Education of wife	-0.108***
	[0.0390]
Post * Education of wife	-0.0164
	[0.0136]
Constant	0.254***
	[0.0352]
Observations	30,680

Standard errors clustered at village level.
*** p<0.01, ** p<0.05, * p<0.1
Village level and state by time fixed effects included, as well as controls for wealthindex, religion, occupation, and occupation of partner.

Table 1.9: Difference in difference in matching and bargaining outcomes

VARIABLES	(1) Education gap	(2) Marriage age gap	(3) Marriage age	(4) Physical abuse	(5) Beating	(6) No auton.	(7) Age at first birth	(8) Interval between marriage and birth	(9) Number of children	(10) Fraction of male children	(11) Distance from ideal number of children
Ultrasound village *	0.00970	0.945***	-1.226***	-0.00464	0.0560**	0.0531*	-1.117***	-1.057	0.618***	0.0373	0.508***
Born post 1981	[0.0521]	[0.236]	[0.162]	[0.0305]	[0.0264]	[0.0295]	[0.174]	[1.211]	[0.0627]	[0.0280]	[0.0649]
Wealth index	-0.0186**	0.0377	0.198***	-0.0467***	-0.0171***	-0.0209***	0.172***	-0.257	-0.120***	0.00633*	-0.0553***
	[0.00757]	[0.0352]	[0.0232]	[0.00500]	[0.00391]	[0.00424]	[0.0246]	[0.175]	[0.00990]	[0.00361]	[0.0101]
Religion	0.00127	-0.00430	-0.00103	0.000802	0.000581	0.000820*	-0.00198	-0.0188	0.000883	-8.89e-05	-0.000792
	[0.000981]	[0.00427]	[0.00321]	[0.000626]	[0.000499]	[0.000472]	[0.00336]	[0.0126]	[0.00117]	[0.000443]	[0.00130]
Education	na	-0.346***	0.834***	-0.0508***	-0.0452***	-0.0329***	0.736***	-0.501***	-0.224***	-0.00553	-0.129***
	na	[0.0381]	[0.0260]	[0.00561]	[0.00429]	[0.00466]	[0.0274]	[0.189]	[0.0103]	[0.00389]	[0.0110]
BMI	0.178	-0.812	-0.647*	0.0266	0.0888	0.0332	0.256***	0.108	-0.0800***	0.00186	-0.0374***
	[0.120]	[0.584]	[0.393]	[0.0712]	[0.0635]	[0.0678]	[0.0264]	[0.195]	[0.0106]	[0.00388]	[0.0110]
Weight (kg)	-0.0299	0.166*	0.0752	-0.00335	-0.0146	-0.00721	-0.688*	-2.708	0.399**	0.100*	0.210
	[0.0185]	[0.0867]	[0.0594]	[0.0110]	[0.00974]	[0.0104]	[0.417]	[3.161]	[0.156]	[0.0593]	[0.162]
Height (cm)	0.00384	-0.0223	0.0193	-0.00231	0.000184	0.000625	0.0796	0.261	-0.0438*	-0.0173*	-0.0111
	[0.00567]	[0.0264]	[0.0174]	[0.00350]	[0.00288]	[0.00312]	[0.0641]	[0.482]	[0.0238]	[0.00931]	[0.0248]
Rohrers Index	-1.673	7.249	6.759*	-0.319	-0.854	-0.289	0.0184	0.120	-0.0167**	0.00219	-0.0232***
	[1.217]	[6.061]	[4.044]	[0.729]	[0.649]	[0.693]	[0.0190]	[0.125]	[0.00693]	[0.00298]	[0.00730]
Spouse's education	na	-0.134***	0.260***	-0.0285***	-0.00778*	0.00479	7.343*	33.95	-4.649***	-0.929	-2.957*
	na	[0.0388]	[0.0245]	[0.00572]	[0.00427]	[0.00451]	[4.268]	[32.10]	[1.589]	[0.604]	[1.649]
Occupation	-0.0342***	-0.0288	-0.0323*	0.0127***	-0.00157	-0.0177***	0.256***	0.108	-0.0800***	0.00186	-0.0374***
	[0.00494]	[0.0235]	[0.0176]	[0.00293]	[0.00249]	[0.00289]	[0.0264]	[0.195]	[0.0106]	[0.00388]	[0.0110]
Spouse's occupation	0.0385***	0.122***	0.0880***	-0.00604**	-0.00723***	0.00547**	-0.111***	-0.301***	-0.0377***	-0.00387	-0.0157**
	[0.00453]	[0.0209]	[0.0145]	[0.00272]	[0.00228]	[0.00254]	[0.0199]	[0.107]	[0.00604]	[0.00263]	[0.00640]
Constant	-0.294	8.286**	12.42***	0.892*	0.571	0.470	0.0791***	-0.0244	-0.0115**	0.000906	-0.00670
	[0.866]	[4.034]	[2.658]	[0.531]	[0.440]	[0.476]	[0.0156]	[0.0997]	[0.00536]	[0.00224]	[0.00563]
Observations	29,356	28,855	29,356	22,696	29,356	28,942	25,798	25,798	29,356	25,798	28,920

Standard errors clustered at village level.

*** p<0.01, ** p<0.05, * p<0.1

Village and state by birthyear fixed effects included. Outcomes observed in 2005 from NFHS III survey. Sample includes only married women born between 1976 and 1987.

Table 1.10: Placebo difference in difference in composition outcomes

VARIABLES	Coefficient on Ultrasound village * Born post 1970	Constant
BMI	-0.0981	18.55***
	[0.267]	[0.145]
Weight (kg)	0.183	41.29***
	[0.686]	[0.375]
Height (cm)	0.404	149.5***
	[0.367]	[0.219]
Educ. (years)	0.0152	-0.468***
	[0.0397]	[0.0277]

Standard errors clustered at village level.

*** p<0.01, ** p<0.05, * p<0.1

Village and state by birthyear fixed effects included. Outcomes observed in 2005 from NFHS III survey. Sample includes only married women born between 1966-1975.

Table 1.11: Placebo difference in difference in matching and bargaining outcomes

VARIABLES	(1) Education gap	(2) Marriage age gap	(3) Marriage age	(4) Physical abuse	(5) Beating	(6) No auton.	(7) Age at first birth	(8) Interval between marriage and birth	(9) Number of children	(10) Fraction of male children	(11) Distance from ideal number of children
Ultrasound village * Born post 1970	0.0292	-0.665**	1.368***	-0.000797	-0.00569	0.0221	1.681***	3.339**	-0.0362	-0.0291	-0.00255
	[0.0390]	[0.266]	[0.250]	[0.0353]	[0.0281]	[0.0311]	[0.268]	[1.515]	[0.0683]	[0.0224]	[0.0679]
Wealth index	0.226***	-0.0106	-0.0229	-0.0393***	-0.0203***	-0.00438	-0.0402	-0.103	-0.187***	0.00502*	-0.137***
	[0.00699]	[0.0443]	[0.0321]	[0.00563]	[0.00435]	[0.00449]	[0.0346]	[0.242]	[0.0149]	[0.00296]	[0.0148]
Religion	-0.00122*	-0.00385	0.000827	0.000294	-0.000565	0.000256	-0.00172	-0.0156	0.00131	0.000569*	-0.00154
	[0.000720]	[0.00490]	[0.00411]	[0.000486]	[0.000431]	[0.000378]	[0.00381]	[0.0190]	[0.00162]	[0.000298]	[0.00162]
Education	-0.673***	-0.543***	1.363***	-0.0405***	-0.0463***	-0.0336***	1.281***	-0.458*	-0.308***	0.00438	-0.179***
	[0.00763]	[0.0467]	[0.0376]	[0.00571]	[0.00456]	[0.00481]	[0.0393]	[0.245]	[0.0149]	[0.00324]	[0.0146]
BMI	0.141*	-0.617	-0.980**	0.0688	0.0382	-0.00181	-1.448***	-8.436***	0.687***	-0.00524	0.557***
	[0.0720]	[0.490]	[0.425]	[0.0617]	[0.0544]	[0.0482]	[0.426]	[2.577]	[0.143]	[0.0350]	[0.142]
Weight (kg)	-0.0168	0.120	0.104	-0.0143	-0.00310	0.00129	0.192***	1.359***	-0.0905***	0.00110	-0.0809***
	[0.0112]	[0.0771]	[0.0648]	[0.00936]	[0.00808]	[0.00758]	[0.0662]	[0.420]	[0.0230]	[0.00553]	[0.0228]
Height (cm)	0.00298	-0.0184	0.0310	0.00276	-0.00239	-0.00330	0.0110	-0.169	-0.0103	2.40e-05	-0.000812
	[0.00418]	[0.0260]	[0.0210]	[0.00326]	[0.00253]	[0.00281]	[0.0219]	[0.151]	[0.00852]	[0.00187]	[0.00827]
Rohrers Index	-1.430*	5.859	10.53**	-0.600	-0.495	-0.0484	14.45***	79.32***	-7.339***	0.0512	-5.719***
	[0.737]	[4.949]	[4.360]	[0.636]	[0.561]	[0.492]	[4.335]	[25.70]	[1.437]	[0.355]	[1.419]
Spouse's education		-0.113**	0.322***	-0.0314***	-0.0133***	0.00998**	0.298***	-0.335	-0.112***	0.00446	-0.0566***
		[0.0464]	[0.0339]	[0.00575]	[0.00448]	[0.00466]	[0.0359]	[0.243]	[0.0157]	[0.00302]	[0.0155]
Occupation	-0.00415	-0.0824***	0.109***	0.0120***	-0.0113***	-0.0187***	0.0672***	-0.123	-0.0309***	-0.00346**	-0.0218***
	[0.00309]	[0.0197]	[0.0194]	[0.00262]	[0.00203]	[0.00217]	[0.0197]	[0.113]	[0.00610]	[0.00168]	[0.00592]
Spouse's occupation	0.0965***	0.0393*	0.0870***	-0.0119***	-0.00423*	0.00494*	0.107***	0.247*	0.00238	-0.00392**	-0.00727
	[0.00377]	[0.0239]	[0.0201]	[0.00279]	[0.00226]	[0.00254]	[0.0202]	[0.126]	[0.00751]	[0.00174]	[0.00723]
Constant	-0.488	7.881**	12.25***	0.260	1.001***	1.018**	17.45***	53.14**	6.346***	0.503*	2.116*
	[0.640]	[3.985]	[3.209]	[0.498]	[0.387]	[0.429]	[3.347]	[23.05]	[1.299]	[0.285]	[1.261]
Observations	26,829	26,362	26,829	21,425	26,829	26,426	26,030	26,030	26,829	26,030	26,232

Standard errors clustered at village level.

*** p<0.01, ** p<0.05, * p<0.1

Village and state by birthyear fixed effects included. Outcomes observed in 2005 from NFHS III survey. Sample includes only married women born between 1966-1975.

Table 1.12: Two sample two stage least squares with access to major health facility

VARIABLES	(1) BMI	(2) Height (cm)	(3) Education (years)	(4) Marriage age gap	(5) Marriage age	(7) Physical abuse	(8) Beating	(9) No autonomy	(10) Age at first birth	(11) Interval between marriage and birth	(12) Number of children	(13) Fraction of male children	(14) Distance from ideal number of children
Ultrasound village * Born post 1980	-1820***	18.22	-2.701***	5.874***	-10.04***	0.402***	0.313**	0.357**	-9.947***	4.377	6.367***	0.0819	7.276***
	[450.8]	[17.88]	[0.250]	[1.259]	[0.834]	[0.155]	[0.148]	[0.154]	[0.831]	[5.684]	[0.342]	[0.132]	[0.588]
Socioeconomic standard living index	-35.55*	3.299***	0.0457***	-0.0443	-0.262***	-0.0301***	-0.00897	0.0141**	-0.241***	0.408	0.152***	-0.00110	-0.0263*
	[18.73]	[0.754]	[0.0116]	[0.0544]	[0.0374]	[0.00688]	[0.00719]	[0.00663]	[0.0388]	[0.267]	[0.0156]	[0.00448]	[0.0144]
Rural	0.435	-0.0332	0.000171	-0.00168	0.00805***	0.00110***	-0.000263	-0.000157	0.00470**	-0.0501***	1.45e-05	0.000283	0.00110
	[1.023]	[0.0335]	[0.000580]	[0.00327]	[0.00224]	[0.000423]	[0.000277]	[0.000298]	[0.00216]	[0.00985]	[0.000707]	[0.000204]	[0.00508]
Wealth index	-95.96***	1.634***	0.0409***	0.0917***	-0.0902***	-0.0130***	0.0198***	0.0121***	-0.0435***	0.408***	-0.0196***	0.00305**	-0.0541***
	[8.804]	[0.269]	[0.00455]	[0.0178]	[0.0144]	[0.00269]	[0.00327]	[0.00278]	[0.0139]	[0.0916]	[0.00518]	[0.00149]	[0.0105]
Religion	71.54***	3.723***	0.259***	0.0658*	0.138***	-0.0254***	-0.0275***	-0.0231***	0.0868***	-0.595***	-0.122***	0.00487*	-0.00194**
	[13.26]	[0.490]	[0.00715]	[0.0357]	[0.0247]	[0.00445]	[0.00454]	[0.00429]	[0.0254]	[0.171]	[0.00998]	[0.00293]	[0.000924]
Education	-8.062	2.691***	na	-0.307***	1.036***	-0.0513***	-0.0544***	-0.0328***	0.961***	-0.669***	-0.266***	-0.00203	-0.155***
	[10.53]	[0.404]	na	[0.0283]	[0.0218]	[0.00372]	[0.00351]	[0.00348]	[0.0221]	[0.140]	[0.00810]	[0.00249]	[0.00846]
Spouse's education	-13.38	3.023***	na	-0.189***	0.303***	-0.0310***	-0.00788**	0.00366	0.321***	0.216	-0.125***	0.00164	-0.0436***
	[11.96]	[0.405]	na	[0.0290]	[0.0201]	[0.00383]	[0.00350]	[0.00352]	[0.0206]	[0.142]	[0.00821]	[0.00241]	[0.00862]
Occupation	-0.846	0.455*	0.00170	-0.129***	-0.0203	0.0125***	-0.00784***	-0.0209***	-0.0561***	-0.140*	-0.00509	-0.00266*	-0.0110**
	[6.911]	[0.242]	[0.00371]	[0.0167]	[0.0145]	[0.00223]	[0.00212]	[0.00199]	[0.0146]	[0.0831]	[0.00463]	[0.00158]	[0.00506]
Spouse's occupation	39.97***	0.723**	0.182***	0.167***	0.216***	-0.00670**	-0.00754***	0.00112	0.207***	0.00401	-0.104***	-0.000978	-0.00996**
	[8.303]	[0.298]	[0.00459]	[0.0210]	[0.0162]	[0.00263]	[0.00275]	[0.00263]	[0.0158]	[0.0998]	[0.00586]	[0.00173]	[0.00447]
Constant	2585***	1517***	1.651***	2.726***	25.02***	0.312***	-0.0336	0.461***	18.87***	17.35***	0.0167	0.947***	-0.415***
	[201.0]	[6.392]	[0.101]	[0.522]	[0.447]	[0.0623]	[0.0633]	[0.0542]	[0.471]	[2.179]	[0.162]	[0.0917]	[0.0836]
Observations	44742	43022	44742	44140	44742	36667	44742	44742	41710	41710	44742	41710	44020

Village and state by birthyear fixed effects included.
Sample includes only married women born between 1976-1984, since respondents in the sample used to estimate the Ultrasound*Post variable can be born no later than 1984 to be eligible in the 1999 NFHS II.
Ultrasound village is instrumented using access to a major health facility in 1999 (from NFHS II).
Outcomes observed in 2005 from NFHS III survey.
Robust standard errors in brackets, corrected according to Inoue and Solon (2008)
*** p<0.01, ** p<0.05, * p<0.1

Chapter 2

Throwing the Baby out with the Drinking Water: Unintended Consequences of Arsenic Mitigation Efforts in Bangladesh

Abstract

The 1994 discovery of arsenic in ground water in Bangladesh prompted a massive public health effort to test all tubewells in the country and convince nearly one-quarter of the population to switch to arsenic-free drinking water sources. According to numerous sources, the campaign was effective in leading the majority of households at risk of arsenic poisoning to abandon backyard wells in favor of more remote tubewells or surface water sources, a switch widely believed to have saved numerous lives. We investigate the possibility of unintended health consequences of the wide-scale abandonment of shallow tubewells due to higher exposure to fecal-oral pathogens in water from arsenic-free sources. Significant small-scale variability of arsenic concentrations in ground water allows us to compare trends in infant and child mortality between otherwise similar households in the same village who did and did not have an incentive to abandon shallow tubewells. While child mortality rates were similar among households with arsenic-contaminated and arsenic-free wells prior to public knowledge of the arsenic problem, post-2000 households living on arsenic-contaminated land have 27% higher rates of infant and child mortality than those not encouraged to switch sources, implying that the campaign doubled mortality from diarrheal disease. These findings provide novel evidence of a strong association between drinking water contamination and child mortality, a question of current scientific debate in settings with high levels of exposure to microbial pathogens through other channels.

2.1 Introduction

Water contamination is a central cause of illness in developing countries. The primary type of contamination in most settings is fecal-oral pathogens which lead to diarrheal disease, the second most common cause of infant and child mortality worldwide. However, in Bangladesh and a handful of other countries, carcinogenic heavy metals naturally leaching into ground water is a parallel concern. Based on tests conducted by the British Geological Survey (BGS) in 1998, an estimated 20 million Bangladeshis had been drinking shallow tubewell water that contained above the government's recommended maximum arsenic concentration of 50 micrograms per liter, and many more above the level recommended by the World Health Organization of 10 micrograms per liter. Although the health effects of chronic low-level exposure to arsenic are poorly understood, many believe the Bangladeshi population to be in danger of serious health effects from long-term arsenic poisoning.[0]

The subsequent international effort to move households away from water sources contaminated with arsenic constitutes one of the most successful public health campaigns in recent history in terms of scale, speed and success rate. In 1999, with help from international donors and NGOs, the Bangladeshi government initiated a massive campaign to test over five million tubewells throughout the country and conduct awareness-building activities encouraging households to abandon contaminated sources. According to household survey data from the Bangladesh Demographic and Health Survey (BDHS), by 2004 not only was there a high level of awareness of arsenic contamination among households in endemic regions, but the majority of households had stopped drinking from wells that were known to be contaminated.[1] In a strikingly short amount of time, awareness-building efforts alone led an estimated 23% of the population to transition from backyard pumps to less convenient drinking water sources including more remote tubewells or surface water sources.

However, as we investigate in this paper, in the process of switching, millions of households may have substantially increased their exposure to water-borne disease. In a setting such as Bangladesh where surface water is heavily contaminated with fecal bacteria, which causes diarrheal disease, cholera, dysentery and other potentially fatal water-borne diseases, backyard tubewells are widely considered "the most appropriate technology in terms of microbiologically clean water" (Lokuge et al., 2004). Not only are shallow tubewells protected from the surface, and therefore have very low rates of fecal contamination compared to ponds or dugwells, but by virtue of being located close to the residence, they minimize water storage time, which is highly correlated with pathogen levels since water becomes contaminated at a rapid rate in storage (Wright et al., 2004). Since distance to water source is also likely to reduce the overall amount of drinking water consumed, morbidity and mortality from diarrheal disease are also likely to increase when households switch to less convenient sources (Pruss et al., 2002).

For both reasons, recent successful public health efforts to move households away from shallow tubewells are likely to have unintentionally increased infant and child mortality among those that discovered arsenic in their groundwater. The extent to which this occurred depends on the marginal effect of clean drinking water on mortality from water-borne disease (the most important of which is diarrheal disease), a question of scientific ambiguity and intense policy debate in settings with high levels of exposure to microbial pathogens

[0] See, for instances, media coverage in the New York Times ("Death by Arsenic: A special report; New Bangladesh Disaster: Wells That Pump Poison", November 10, 1998), the Economist ("A nation poisoned", December 20, 2001) and the British Medical Journal ("Half of Bangladesh population at risk of arsenic poisoning", March 25, 2000).

[1] For instance, in Barisal division where our study takes place, 46% of households had been drinking from contaminated tubewells in 1999 and only 1% were doing so in 2004. Furthermore, 78% of BDHS respondents had heard about the problem of arsenic in drinking water in 2004 (BDHS data, authors' tabulations).

through other sources such as bathing, dish-washing and secondary drinking water sources. That is, in settings in which diarrheal disease is endemic, drinking water source improvements may be insufficient to interrupt transmission of waterborne pathogens, which also occurs via ingestion of contaminated food and other beverages, person-to-person contact, and by direct or indirect contact with infected feces.

To investigate this question, we quantify the impact on infant and child mortality of switching water sources in response to the arsenic testing and awareness campaign in one subdistrict of Bangladesh (Barisal). To do so, we make use of the high degree of natural variation within villages in the rate of arsenic in shallow groundwater, which is uncorrelated with observable measures of land quality within small distances. This small-scale variability enables us to employ a difference-in-difference estimation strategy using data from a random sample of 3100 households spread across 155 villages that compares households living relatively close to one another who tested positive versus negative for arsenic contamination, and track the change in health outcomes of children born before versus after well-testing took place.

To identify households that were encouraged to switch to more distant water sources, we collected water samples from the drinking water kept in each household's kitchen and from the closest shallow tubewell. Consistent with census data collected by the government in 1999, our tests of the latter indicate that over 65% of households were drawing water from arsenic-contaminated shallow tubewells prior to 2000. However, only 1% of households in our sample tested positive for arsenic in their stored drinking water, implying that over two-thirds switched from shallow tubewells to alternative arsenic-free water sources between 2000 and 2009. Since there is no piped water in these rural villages, arsenic-free water sources include either deep tubewells, uncontaminated shallow tubewells in neighboring houses, or surface water sources.

We then estimate the trends in infant and child mortality with a village fixed effect specification that absorbs differences in mean characteristics between relatively exposed and relatively unexposed villages arising from potential correlations between the spatial clustering of arsenic contamination at a macro level and characteristics such as income that may influence health. In doing so, our identification strategy relies on the assumption that the spatial distribution of arsenic contamination is quasi-random within distances as small as villages, which these and other data support.[2]

Our estimates indicate that, while infant and child mortality rates were almost identical in contaminated versus uncontaminated households before 2000, these outcomes diverged sharply immediately after. Post-2000, households with arsenic-contaminated wells - those likely to have switched sources - exhibit a 27% increase in infant and child mortality relative to those in the same village with arsenic-free wells. This figure implies that the abandonment of shallow tubewells approximately doubled rates of diarrheal disease in the population of switchers.

We also undertake a similar exercise at the national level using data from the 2004 BDHS. Just as we did in 2009, the 2004 BDHS collected drinking water samples from each household and tested them for arsenic contamination. While the majority of households had already switched to arsenic-free drinking water by 2004, 8% of households spread across 29% of villages were still drinking from arsenic-contaminated tubewells despite public health efforts to change behavior. We make use of this variation to test whether households that we know did not switch away from backyard tubewells exhibit relatively lower rates of infant mortality after 2000 compared with households that can be presumed to have switched to more distant sources. While this approach raises concerns about the endogeneity of switchers, we view it as a consistency check on the more tightly identified estimates in Barisal that also allows us to look in more detail at potential mechanisms

[2]For instance, in a large multi-village dataset from the Araihazar subdistrict, the spatial distribution of arsenic is orthogonal to observable household characteristics within but not across villages (Madajewicz et al., 2007).

through which switching may adversely affect health.

Indeed, consistent with our estimates from Barisal, switchers have significantly higher rates of infant and child mortality after but not before the well-testing campaign relative to non-switchers. Furthermore, the negative effect of abandoning shallow tubewells appears to be equally large when deep tubewells are available as an alternative source. This provides evidence that clean but remote water sources are poor substitutes for backyard tubewells in terms of mortality risk, either due to high rates of recontamination in storage or more frequent use of water from secondary surface water sources.

Although it is difficult to argue that non-switchers identified in the BDHS data are a representative subsample of households, observable characteristics and pre-campaign mortality levels are similar across the two comparison groups, so the pattern does not reflect simple convergence in child mortality between low- and high-SES households that happens to coincide with the well-testing campaign. In addition, when we use BDHS verbal autopsy data to classify causes of death, we find that abandoning shallow tubewells is associated with an increase in deaths due to diarrheal disease but find no such divergence in mortality from pneumonia or fever, further strengthening evidence of a causal relationship between shallow tubewells and child mortality as opposed to general trends in mortality between switchers and non-switchers that are unrelated to water source.

Together, these two sets of results provide novel evidence of a strong link between improved drinking water sources and mortality from diarrheal disease in settings with high risk of exposure to fecal contamination through other channels. Given the potentially small benefit offered by drinking microbiologically safe water in settings where there is constant exposure to fecal matter through bathing, food preparation, and dish washing, the marginal health benefit of protected water sources in countries such as Bangladesh is to date an unresolved question in the public health literature. The debate has become particularly heated in light of previous results showing little difference in rates of diarrhea by water source in such settings (Esrey and Habicht, 1986; Lindskog et al., 1987; Caldwell et al., 2003; Kremer et al., 2010). Hence, in addition to contributing to the current debate over arsenic mitigation efforts in Bangladesh, our results have important policy implications for more general efforts to reduce infant and child mortality in the most afflicted settings.

Our results also highlight the need to proceed cautiously when issuing public health recommendations when there is insufficient information concerning competing risks. In the case of rural Bangladesh, should the use of shallow tubewells contaminated with arsenic continue to be discouraged in regions with an absence of equally clean and convenient alternative water sources? Our results suggest that continued efforts to do so could have dire consequences for the health of infants and children, which need to be weighed carefully against the less understood health consequences of chronic low-level arsenic exposure.

2.2 Background

2.2.1 Public health efforts surrounding shallow tubewells

Largely because of its geographic vulnerability to flooding combined with its high population density, Bangladesh has historically had one of the highest incidence of water-borne viral and parasitic infections and corresponding infant and child mortality in the world. To reduce chronic cholera and diarrheal disease outbreaks, an estimated 8.6 million shallow tubewells were constructed throughout the country from the 1970s to the 1990s.[3] These efforts succeeded in moving an estimated 95% of rural Bangladeshis from

[3]Tubewell construction was funded by the Bangladeshi government, UNICEF, World Bank, and numerous other public and private organizations, and also financed privately by households.

parasite-infected surface water to protected ground water (Caldwell et al., 2003).

Unfortunately, these improvements in sanitation were short-lived due to the discovery of arsenic in the major shallow aquifers.[4] Geologists first discovered traces of arsenic in Bangladesh groundwater in 1987, and physical manifestations of arsenicosis, the disease caused by substantial ingestion of arsenic, were first documented in 1994. Three years later, the World Health Organization (WHO) publicly declared groundwater arsenic contamination to be a "major public health issue," and issued a grant to address the emergency.

In 1998, BGS conducted a nationwide study measuring levels of contamination in a sample of shallow tubewells across Bangladesh. Results indicated that 21 million people (15% of the population) were in grave danger, drinking water with more than 50 ppb (parts per billion, or micrograms) As, and 42 million in lesser danger, drinking water with more than 10 ppb As.[5] In the late 1990s and early 2000, the Bangladeshi government, along with UNICEF and a host of other aid organizations, conducted a blanket screening of all shallow tubewells in contaminated regions of the country. Wells that tested contaminated (1.4 million) were painted red and those that tested safe (3.3 million) were painted green (Johnston, 2006).

Households were and continue to be strongly encouraged to stop drinking from red tubewells and switch to alternative sources (Jakariya, 2007). Potential alternatives include deep tubewells, piped water, dug wells, treatment of surface water, rainwater harvesting, sharing of safe shallow tubewells, and treatment of arsenic contaminated water. Among these, deep tubewells are one of the most commonly promoted alternatives. Although they are prohibitively expensive for most households to build, between 1998 and 2006, the Arsenic Mitigation Water Supply Project built over 9,000 deep tubewells across 1800 villages in Bangladesh where sufficiently deep aquifers could be found.[6]

Unfortunately, analyses of post-construction deep tubewell water found that arsenic can leach into the wells over time (Feroze Ahmed, 2002; WorldBank, 2007). A fear that further use of deep tubewells would lead to arsenic contamination of Bangladesh's deep aquifers led the 2004 National Policy for Arsenic Mitigation report to stress a "preference of surface water over groundwater as a source for water supply." According to a World Bank evaluation, this report had a notable effect on patterns of water usage "effectively foreclos[ing] use of the less costly option of tubewells as a safe source for small communities, leaving the less popular dug wells, rainwater harvesting, and pond sand filters as options for other areas. Many dug wells were abandoned, and some communities installed new shallow wells (with uncertain arsenic levels) or reverted to surface water from ponds (where water quality is suspect)" (WorldBank, 2007).

Though less emphasized among policymakers, the sharing of safe tubewells has been a relatively popular option in some parts of the country, including the heavily studied district of Araihazar. VanGeen et al. (2002) report that 43% of exposed individuals in Araihazar preferred switching to a nearby safe shallow tubewell over other alternatives such as deepening their well (31%) or using surface water (20%).[7] Within two years of well testing in the district, Schoenfeld (2005) reports that approximately 30% of individuals exposed to greater than 50 ppb As and 15% of individuals using unknown (unpainted) wells switched to nearby green-painted wells.

Finally, nationwide public education campaigns about the presence and dangers of arsenic have been

[4]Arsenic-bearing sediments buried in the aquifers come from rocks that eroded from the Himalayas and were deposited in the low-lying areas which now make up West Bengal and Bangladesh. Arsenic sediment is released into ground water by a natural process called "oxyhydroxide reduction".

[5]This estimate has more recently been increased by the Government of Bangladesh to 30 million and 70 million, respectively (WHO, 2008).

[6]The cost of constructing deep tubewells in most locations is estimated to be well over $500, while the cost of constructing shallow tubewells is estimated to be $38 (Caldwell et al., 2003).

[7]The authors also note, however, that Araihazar District has more shallow tubewells than the rest of Bangladesh.

widespread since 1999.[8] The impact of these educational campaigns are reportedly considerable: 80% of the population is aware that arsenic may be a danger in groundwater (relative to less than ten percent in the late 1990s), and 70% of households report changing their behavior to avoid arsenic (UNICEF, 2008).

2.2.2 Health benefits of switching away from shallow tubewells

Arsenic is a known carcinogen that has been shown in laboratory studies to cause or catalyze several forms of cancer, particularly of the lung and bladder (Kozul et al., 2009; Suzuki et al., 2008; Rossman et al., 2002).[9] Hence, it is generally accepted that exposure to high levels of arsenic (> 100 micrograms) will lead to a major increase in cancer-related deaths and morbidity in the older adult population. There is a notable lack of hard evidence on the health effects at the lower end of the exposure doses. However, in large part due to the long latency of most arsenic-related health problems, the National Research Council concludes that "arsenic-related disease due to chronic exposure through drinking water has a relatively low incidence" in settings with low average life expectancy such as Bangladesh (Research Council, 2001).

One exception to this perspective are recent results from an epidemiological study following over 10,000 adults in the Araihazar District in Bangladesh, which reported very high mortality associated with arsenic exposure (Argos et al., 2010). The authors estimated that approximately 20% of all deaths documented over nine years were attributable to arsenic, with mortality rates nearly 70% higher for those exposed to arsenic concentrations of over 150 ppb relative to those exposed to less than 10 ppb. However, an important caveat to this study not addressed by the authors is that arsenic concentrations in groundwater are not orthogonal to socioeconomic status in this setting. As shown in Madajewicz et al. (2007), due to the spatial clustering of arsenic across the 54 villages in this study area, prior to testing households with uncontaminated wells happen to have significantly higher average income and assets (with 42% more assets and 16% more expenditures) compared to households living on contaminated land.[10] Although the differences disappear when accounting for village fixed effects, the Argos et al. (2010) study fails to do so and, as a result, mortality differentials found in their study are almost certainly biased upwards.[11]

On the other end of the spectrum, the calculations by Lokuge et al. (2004) of the disease burden from arsenic exposure that take into account only "strong causal evidence" from existing studies estimate that arsenic-related disease leads to the loss of 174,174 disability-adjusted life years (DALYs) per year among the population exposed to arsenic concentrations of more than 50 ppb, which amounts to 0.3% of the disease burden, compared with diarrheal disease which accounts for between 7.2% and 12.1% of the total disease

[8]Programs focus on raising awareness of the impact of arsenic ingestion, alternative safe water sources, remedial measures against poisoning, and the understanding that arsenicosis is not contagious (BMOH, 2004). During the testing campaign of 1999-2000, UNICEF had its tubewell testers spend their waiting time sharing basic information about arsenic, dispelling common myths, and then directly showing the villagers the result of the well test. In more recent years, UNICEF has established an educational curriculum integrating hygiene and sanitation with arsenic awareness and also involved the community in choosing alternative water sources best suited to their needs.

[9]Field studies have also found a strong dose-response relationship between skin cancer and arsenic exposure through drinking water (Chen et al., 2006; Mazumdar et al., 1998; Tucker et al., 2001).

[10]VanGeen et al. (2003) and Ahsan et al. (2006) describe these spatial patterns in detail, though not as they relate to SES. VanGeen et al. (2003) notes that "Most of the wells with the lowest As concentrations are located in the northwestern portion of the study area", which appears to contain higher SES villages. According to Madajewicz et al. (2007), "Arsenic is released when the accumulation of plant matter during the formation of river delta deposits drives groundwater to anoxia. The process may generate a correlation between soil types and arsenic levels and therefore possibly between arsenic levels and incomes. However, this correlation would not be likely to appear within villages. Wells are located within small, densely inhabited villages. The surrounding fields are fairly uniform geologically, while the dispersion of incomes and wealth within villages is large."

[11]Furthermore, age is significantly higher and the number of relatives in the study is significantly lower among high-concentration households (Madajewicz et al., 2007). Age is less of a concern in terms of bias since the Argos et al. (2010) estimates control for age. However, households with fewer social network connections are likely to have higher mortality due to a deficit of informal insurance and health care networks.

burden.

Researchers almost universally agree that the relationship between arsenic exposure and morbidity and mortality in younger populations is minimal. One highly publicized study of children in Araihazar found that arsenic exposure inhibits the mental development of children (Wasserman et al., 2004), but the estimates face the same bias that the Argos et al. (2010) study faces so should be interpreted with caution. Similarly, a handful of studies have reported reproductive health consequences of arsenic exposure, although the evidence is mixed (Vachter, 2008; Tofail et al., 2009; Milton et al., 2005; Liaw et al., 2008).

In general, since arsenic exposure also tends to be correlated at a macro level with socioeconomic conditions influencing child development measures, causality cannot be easily inferred from studies that show a correlation between arsenic exposure and various health outcomes (Tofail et al., 2009).

2.2.3 Health costs of switching away from shallow tubewells

Although abandoning shallow tubewells contaminated with arsenic is likely to have a measurable latent effect on reducing mortality in older populations, given the relatively high burden of diarrheal disease, it could come at a significant cost to the health of younger populations. Because shallow tubewells are supplied to individual households (generally built in the backyard close to the residence), they are an extremely convenient water source, which increases the frequency with which water is collected and therefore reduces water storage time and increases water consumption. Storage time is an important determinant of contamination with fecal matter, as water that is not stored properly is continuously exposed to dirty hands and cups or utensils, and previous studies find strong correlations between distance from water source and diarrheal disease (Esrey, 1996). Inconvenience also implies a potential decrease in the amount of water consumed (Hoque et al., 1989), which can have important health consequences for children facing dehydration from diarrheal disease. In fact, according to one previous study, the quantity of water used is a better predictor of child health than the quality of water used (Esrey, 1996).[12]

The only water sources equally convenient to shallow tubewells are surface water sources such as ponds that are also likely to be close to the residence. However, while they are free of arsenic, these sources are significantly more likely to be contaminated with fecal matter. While water filtering and cleaning methods can address point of use contamination, survey data indicate that these have largely been abandoned in rural Bangladesh since the construction of shallow tubewells (Caldwell et al., 2003).

Taking into account all of these changes in risk exposure, Lokuge et al. (2004) estimate that abandonment of shallow tubewells would increase a household's risk of diarrheal disease by 20%. Until now, there has been no empirical estimation of this possibility and health messages promoted by governmental and non-governmental agencies continue to stress the importance of moving away from shallow tubewells that are contaminated with arsenic.[13]

[12] Using experimental methods, Kremer et al. (2010) estimate in rural Kenya that on average, water quality deteriorates by one third between point-of-source and point-of-use.

[13] For instance, the Bangladeshi Ministry of Health recently had the following message posted on its web site: "The public health of the country is now facing a severe threat as a section of existing tube-wells are contaminated with arsenic. Now time has come to return to our old habit. Because we can keep ourselves safe from arsenic pollution by drinking surface water." (March, 2010)

2.3 Estimation Strategy

2.3.1 Data and setting

To study trends in child mortality, we capitalize on extensive household survey data, including reproductive and child health outcomes for all children in the household, that were collected by the authors in 2007 as part of an impact evaluation of an adolescent empowerment program currently being implemented in one district in the southeast of the country. The data set covers 155 villages and 3093 households in Barisal District of Barisal Division, one of the areas most heavily contaminated with arsenic in the country. According to village-level well testing data collected by the government in 1999, over 70% of tubewells in the area were contaminated.[14]

Barisal was also a relatively "successful" region in terms of the public health campaign that followed. Data from the BDHS reveal a uniquely high rate of switching away from contaminated water sources in Barisal, attributed largely to the geology of the region, which made it possible to construct deep tubewells in almost all villages. According to estimates from the Bangaldeshi Government's National Arsenic Mitigation Information Center (NAMIC), there is currently one deep tubewell per approximately 100 households in rural Barisal.

The full household survey, of which we use a part, collected data from 9048 households in three districts and five subdistricts of Barisal. Households included in the study were randomly drawn from within the five participating sub-districts in a two-stage sampling process in which villages were first sampled from the universe of villages containing more than 50 and fewer than 500 households, and then 20 households per village were selected at random from village-level census data.[15] Only one of the three districts, Barisal District, is contaminated with arsenic. Hence, our present analysis is restricted to the 3158 households in Barisal district. For the purpose of this analysis, these households were revisited in 2009 for water testing and a brief survey of water use and arsenic awareness, and 3093 households were successfully surveyed at follow-up. In order to link data on child health histories with water source, we also tested each household's closest shallow tubewell for arsenic level and collected survey data on household water sources before and after the 1999-2000 well testing campaign, in addition to respondents' knowledge about arsenic contamination.

Our analysis sample includes all children born in the present home between 1980 and 2007 to heads of households with complete arsenic survey information.[16] The final sample encompasses 2817 households and 11,766 children, 3685 of whom reside in low concentration households and 8081 in high concentration households.[17]

[14]Correspondingly, in a 2000 nationally representative household survey on arsenic contamination, 28% of respondents reporting arsenic symptoms were in Barisal, which contains 9% of the country's population (Caldwell et al., 2006).

[15]Households were eligible for random selection only if they included at least one adolescent girl. Villages of medium size were included in the sample frame because this was seen as an appropriate size for the adolescent girls program the survey was designed to evaluate.

[16]The 2697 children (18%) born after 1980 but before the household moved into the current residence are dropped from the analysis, although the results are robust to including them. As predicted, the point estimate falls but remains statistically significant. We also exclude from the sample 167 individuals whose mother's age at birth is less than twelve years, greater than 45 years, or missing, and one household is dropped because identifying data do not match well between the baseline and arsenic surveys.

[17]Numbers are specific to defining high concentration as those households with wells with greater than 60 ppb As according to our closest-well test results.

2.3.2 Identification strategy

Our identification strategy makes use of the fact that there is significant small-scale variability of arsenic concentrations in ground water uncorrelated with observable land characteristics (Yu et al., 2003).[18] In particular, an estimated 88% of contaminated wells are located within 100 meters of an uncontaminated well (VanGeen et al., 2003), giving rise to substantial within-village variation in contamination: in 47% of villages in our sample between 20 and 80% of wells are contaminated, and in 65% between 10 and 90% of wells are contaminated. Within a village, local pockets of contamination are impossible to predict as they have not been found to be correlated with any observable features of the land.[19] Hence, while certain villages contain a much higher percentage of contaminated groundwater than others, within a village it is impossible for households to know whether a given property is situated on contaminated groundwater prior to digging the well and testing it. This variation in well contamination makes it possible to compare otherwise identical households residing close to one another who are and are not encouraged to abandon shallow tubewells in 1999 based on revealed arsenic exposure in a difference-in-difference (DID) estimation strategy.

We define a binary level of arsenic exposure using two methods. The first, denoted "measured contamination," categorizes wells (and implicitly households) as contaminated if the concentration of arsenic in the shallow tubewell closest to the surveyed household, as tested by a standard arsenic testing kit, is greater than 60 ppb when measured by our field team in 2009.[20] The second, denoted "reported contamination," categorizes households as contaminated if any of the shallow tubewells ever used by the household are reported in survey data collected in 2009 to have tested positive for arsenic, been painted red, been deemed unsafe for drinking, been abandoned, or been built less than three years before the survey.[21] The 2% of households that lack information on shallow tubewell use because of non-response are categorized using the "measured contamination" method for both measures. The two measures of contamination correspond for 87% of households.

Since there is some concern that households underreport use of contaminated wells, we favor the "measured contamination" variable over the "reported contamination" variable.[22] In the case that underreporting of contaminated wells is unrelated to household characteristics associated with child health outcomes, this measurement error will bias our estimate downward. However, if there is a more complicated reporting bias - for instance, if households that are more aware of health risks are more prone to hide contaminated wells - there is risk that our estimates are biased upwards.

A key assumption in our "measured contamination" method is that distance is a valid proxy for utilization (i.e. that the closest shallow tubewell for a given household was in fact the main source of drinking water prior to the arsenic awareness campaign), which is probably a fairly accurate assumption since most households

[18] Small-scale variation in arsenic levels is due to heterogeneity of near-surface geology and the resulting biogeochemical environments, both of which are uncorrelated with agricultural land quality.

[19] Because of this difficulty, encouraging households to build new shallow tubewells on uncontaminated land is not a viable policy alternative, although it has been explored extensively. There are many times village-level observable features of land such as the permeability of nearby soils that predict probabilities of arsenic contamination for all households in the village, but such characteristics do not provide information about within-village location of pockets of arsenic in underground aquifers (Madajewicz et al., 2007).

[20] To test arsenic levels in water samples, we used the Wagtech Digital Arsenator testing kit in a laboratory setting. We chose 60 ppb as the cutoff to reflect the 50 ppb WHO cutoff, taking into account an estimated 2% per year increase in arsenic levels, so that contaminated wells in our sample are those believed to have tested above 50 ppb in 1999. Relatively constant groundwater As concentrations have been reported in a number of time series studies in the area (VanGeen et al., 2003).

[21] The latter condition is included under the assumption that wells installed recently were built to replace contaminated wells.

[22] Survey data also show evidence of underreporting: Our survey data on history of shallow tubewell use, when compared to our measured contamination, indicates a tendency to underreport use of highly contaminated wells prior to their being tested. The "reported contamination" method also lacks complete data, as many individuals responded "Don't Know" to relevant survey questions.

have only one shallow tubewell in the interior courtyard close to their dwelling and convenience has been shown to be an important predictor of amount of water consumed. Although households could also lie about which well is the closest, our "measured contamination" method is less prone to such biases since enumerators were instructed to visually inspect the area surrounding each dwelling to identify the closest well, and typically backyard tubewells are quite close to the dwelling and highly visible.

To test the validity of our key identifying assumption regarding the quasi-random nature of variation in arsenic exposure, Table 2.1, Panel 1 presents a host of time-invariant sample characteristics, with mean values shown separately for those in low concentration households and those in high concentration households based on both measured and reported contamination. All averages are regression-controlled means that account for village fixed effects, as do reported t-statistics of the differences in means across samples. Characteristic of the rural population in this area, households in our sample are relatively poor and uneducated: Mothers completed an average of three and a half years of schooling and fathers completed nearly four. The mean monthly income of a household was approximately $11.40, with 40% of households working primarily in agriculture and 15% of households working primarily in business. Households owned on average less than one acre of land and lived in a home with fewer than three rooms. However, the majority is not destitute: approximately 90% of respondents reported having sufficient food for the family in a given week, and more than half (54%) of households had some type of outstanding loan.

In terms of similarity of our comparison groups on observable characteristics, baseline differences across low and high contamination households are small and statistically insignificant, supporting our identification strategy. Only one variable out of 23 - whether Muslim - is significantly different across the two subsamples at the 10% level. This is true under the measured contamination measure but not the reported contamination measure, and the point estimate of the difference is extremely small. Other measures of socio-economic status indicate that the samples are balanced on income and wealth, and an F-test of joint significance indicates that the samples are balanced on observables within villages (p=0.54). Nonetheless, we present all estimates with and without controlling for a number of family background variables to reduce the scope for imbalance to bias our estimates. Interestingly, as shown in Appendix Table 2.1, the same exercise conducted without accounting for cross-village variation shows a high degree of imbalance, as is also observed in other study areas such as Araihazar. In our setting, however, spatial clustering across villages produces a pattern in which arsenic contamination is disproportionately concentrated in relatively well-off villages.

In terms of endogenous variables, differences in infant and child mortality across sub-samples are evident from sample means alone: High concentration households have higher rates of infant and child mortality over the entire period, although we see no difference in fertility, sex ratios, or the timing of births, all of which could potentially be influenced by differences in child mortality and complicate the interpretation of trend differences. Interestingly, individuals in low contamination households reported statistically significantly higher home values than those in high contamination households, which is presumably a causal effect of having a contaminated well. Most households list at least two sources of drinking and cooking water, and about 70% report that the closest shallow tubewell was tested and painted during the campaign of 1998-2000, consistent with estimates from national data.

2.3.3 Estimating equation

We test for changes in infant and child mortality that correspond to the timing of the testing campaign by estimating the following difference-in-difference equation for individual i in household j and village v, which includes village fixed effects (θ):

$$Y_{ijv} = \alpha_{ijv} + \gamma * HighConc_{jv} + \delta * EarlyLifeExp_{ijv} + \beta * (HighConc_{jv} + EarlyLifeExp_{ijv}) + \theta_v + _{ijv} \quad (2.1)$$

HighConc is a dummy variable taking the value of one if the individual is in a household exposed to arsenic contamination. *EarlyLifeExp* denotes the fraction of a child's life below the age-of-death cutoff being considered in the outcome variable that he or she was exposed to microbiologically unsafe drinking water from surface sources or deep tubewells as a result of the testing campaign in 1999-2000. Hence, for under 1 mortality, *EarlyLifeExp* is simply a dummy variable that takes a value of 1 if the child was born after 2000 and 0 if born before 2001, but for under 2 mortality, *EarlyLifeExp* takes a value of 1 if the child was born after 2000, 0.5 if born in 2000, and 0 if born before 2000.[23] Since it is difficult to verify in exactly which of these two years the majority of households was tested, we also run analogous estimates using 1999 as a cutoff point in place of 2000. Although the estimates are robust to either cutoff, we choose 2000 as our preferred specification since we presume that behavioral change towards alternative drinking water sources had at least a slightly lagged response. Standard errors are clustered at the household level.

We are interested in the coefficient estimate of β, describing the change in mortality due to abandoning shallow tubewells. Proper identification relies on the assumption that other natural processes or human interventions occurring over the observed time period did not differentially affect infant and child mortality rates for households exposed to high concentrations versus low concentrations of arsenic. The high degree of variation in arsenic exposure across very small distances and the similarity across comparison groups in relevant baseline characteristics and mortality levels prior to revelation of arsenic contamination lend credibility to this assumption. To test this assumption, we also run a placebo check described in Section 2.4.1 in which we test whether an alternative cutoff well above 60 ppb produces similar patterns within a subsample of households that were all encouraged to abandon shallow tubewells (those with arsenic concentrations higher than 60 ppb). Since this specification compares switchers with switchers, we should observe a significant effect of the cutoff only if unobservable determinants of mortality are correlated with arsenic in groundwater.

Along with the parsimonious specification, we also estimate versions of Equation 2.1 with controls for the individual's sex, parity, birth year, and birth year squared, and a wider set of control variables that includes age of mother at birth, mother's education, father's education, years since birth of last child, solvency, land size, number of rooms in house, electricity, whether Muslim, and monthly income per capita.

2.4 Results

Figures 2-1 to 2-3 present the trends in one, two, and five year mortality between 1978 and 2007 based on the raw data using the measured contamination method to divide the sample into switchers and non-switchers. For smoothness, mortality rates are averaged across two-year periods. For the most part mortality trends in high concentration households closely follow those in low concentrations households until 1998-1999, at which point they begin to diverge. Both child and infant mortality rates are substantially higher among individuals in high concentration households relative to those in low concentration households immediately

[23]The maximum number of years of exposure is the mortality interval (of one, two, or five years) over which infant and child deaths are being measured in each outcome variable.

after the arsenic testing campaign (2000-2001), and these differences are sustained to the time of the survey in 2007 (though there is some indication of convergence in the last two-year interval). This suggests that most switching (and the resulting mortality effects of exposure to microbiologically unsafe water) occurred immediately after the campaign.

Table 2.2 presents the corresponding regression results from equation (2.1) for infant, under two, and under five mortality using our measured contamination measure. As reflected in the mortality graphs, the coefficient estimates indicate a substantial and statistically significant increase in mortality after 2000 among individuals living in households with high levels of arsenic in their shallow tubewells. These results are robust to the inclusion of both the basic and full set of controls (detailed in Section 2.3.3). Referring to the full control specification (columns 3, 6 and 9) of Table 2.2, an additional year of exposure to the post-campaign environment for an individual with a contaminated shallow tubewell is associated with a 2 percentage point (27%) increase in the likelihood of death within 12 months, a 3.2 percentage point (33%) increase in the likelihood of death within two years, and a 3.9 percentage point (28%) increase in likelihood of death within 5 years. This implies that mortality from diarrheal disease, which was estimated to account for approximately one-quarter of deaths under age five in 2000 (Morris et al., 2003), approximately doubled after the well-testing campaign for households that abandoned backyard tubewells.

These estimates are large in comparison to the increased burden of diarrheal disease that is predicted in response to the abandonment of shallow tubewells in the projections of Lokuge et al. (2004) (20%), although there are several possible sources of discrepancy. First, the Lokuge et al. (2004) estimate was taken directly from a study by Esrey (1996) that was based on DHS data from eight countries, all of which have diarrhea prevalence below that of rural Bangladesh. Projections were based on the simple correlation between access to improved water supply and reported incidence of diarrhea in children under 5, which could produce downward biased estimates of the causal effect of changes in water supply on diarrheal disease if improved water services are, conditional on income, targeted to areas with highest rates of mortality from diarrheal disease. This is particularly problematic since the Esrey (1996) study was based on extremely small samples within each country.

A potentially more appropriate benchmark is the reduction in rates of diarrheal disease that are associated with the widescale construction of tubewells in rural Bangladesh, which has roughly fallen in half since the 1970s. Unfortunately, as noted by Caldwell et al. (2003), it is unclear how much of a role can be attributed to the use of tubewells given the concomitant adoption of public health measures such as immunization, antibiotics and oral rehydration therapy (ORT). Still, it is worth noting that, according to autopsy data from the demographic surveillance site of Matlab, diarrheal disease accounted for an estimated 47% of deaths to children ages 1-4 in 1966-1977, then fell to 34% of deaths in 1978-1987, and by 1999 accounted for only 20% of deaths (Baqui et al., 1994), suggesting that the adoption of shallow tubewells could have reduced mortality from diarrheal disease by as much as 57% (or, correspondingly, reverting to surface water sources would increase diarrheal disease by 135%). Hence, we take 20-135% (which encompasses our estimates of 100%) as an appropriate range of possible mortality increases due to the abandonment of shallow tubewells.[24]

The results reported in Table 2.3, using the reported rather than measured contamination of the household water source prior to the testing campaign, show very similar patterns. The estimates are consistently larger in magnitude under the reported contamination method, which could be driven by either higher precision or reporting bias, as described in Section 2.3.2.[25]

[24]Data from 1966-1987 reported by the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B) Demographic Surveillance System - Matlab: Registration of Demographic Events.

[25]Furthermore, mortality patterns are similar across gender (unreported), although the sample sizes are too small to draw

2.4.1 Robustness checks

Our estimates are robust to a number of alternative specifications and placebo checks, the results of which are presented in Appendix Tables 2.2 to 2.7. Appendix Table 2.2 presents the same regressions as Table 2.2, replacing the individual with the household as the unit of observation, such that the outcome variable is fraction of deaths in the household under a certain age, and early life exposure is now measured as the fraction of children born after 2000. Coefficient estimates are noisier but virtually unchanged. The regressions in Appendix Table 2.3 include a binary in place of a continuous measure of early life exposure, in which exposure is equal to 1 if the child was born after 2000 and 0 otherwise. Again, the results are noisier but very similar, as we would expect.

Since approximately one-third of households report both well testing and switching drinking water sources after 2000, Appendix Table 2.4 makes use of survey data on the year in which a household's well was reportedly tested to try to gain precision on the anticipated date of switching within a given household. Here, we replace the binary indicator of a child being born after 2000 with an indicator of a child being born after the household's closest well was tested, according to survey reports. In this specification, the DID estimate is comparable in magnitude and gains precision, as one would expect if we take the survey reports at face value. However, because there is no way to confirm reports of testing dates, there is a possibility of non-random measurement error biasing these results.

In the regressions reported in Appendix Table 2.5, we run a falsification test in which we exclude households with arsenic contamination levels below 60 ppb (non-switchers) and construct a false cutoff point of 100 ppb. We then estimate a DID regression analogous to Equation 2.1 in which we compare households above and below 100 ppb. Since all of those households were equally encouraged to switch sources after 1999, we should see no difference in trends if our previous estimates truly reflect the causal effect of switching water sources. In contrast, if level of arsenic contamination in groundwater is correlated with unobservable characteristics of the household that are giving rise to differential time trends in child and infant mortality, we should expect to see positive and significant point estimates on the interaction terms in both regressions. As the estimates reveal, we see no significant effect on mortality of having arsenic levels above 100 ppb relative to having arsenic levels between 50 and 100 ppb, which reduces the likelihood that our estimates reflect differential time trends in mortality that are correlated with a household's level of arsenic exposure through mechanisms other than switching drinking water sources. Since we only observe a significant DID estimate when the true cutoff for well-switching is used, we can deduce that the estimate reflects the causal effect of changing water sources rather than time trends in unobservables correlated with arsenic exposure.

Appendix Table 2.6 shows Equation 2.1 estimated only for households whose nearest well was built more than eight years ago. Exclusion of recently installed wells ensures that all individuals in the sample had access to the existing shallow tubewell prior to the testing campaign, and subsequent decisions on water source and usage would have been made with consideration of the campaign. The DID estimate is significant and larger in magnitude than those of the original specification: Obtaining drinking water from surface sources or deep tubewells since birth is associated with a 2.6 percentage point increase in likelihood of death within one year, a 3.8 percentage point increase for two years, and a 4.1 percentage point increase for five years. While the estimates are in theory more accurate, since year of well construction is likely subject to recall bias and potential misreporting, it is possible that estimates that take account of these reports are biased, so our preferred estimates are those in Table 2.2. Finally, Appendix Table 2.7 uses 1999 instead of 2000 as the cutoff date of the campaign, with very similar results.

precise comparisons.

2.4.2 Nationwide trends in DHS data

We next look for nationwide evidence of changes in risk of diarrheal disease attributable to the arsenic mitigation campaign using an analogous estimation strategy with national data on infant and child mortality and water sources available in the 2004 BDHS. The 2004 BDHS tested household drinking water for arsenic contamination and found that 8% of households distributed across 29% of BDHS villages had not switched to arsenic-free drinking water sources in spite of the massive campaign efforts. That is, the presence of arsenic in their drinking water confirms that these households were still using contaminated shallow tubewells in 2004 even though deep tubewells existed in at least 18% of these affected villages.

We make use of this within-village variation in household response to the arsenic mitigation campaign to test whether child mortality trends before and after 2000 look worse for households that switched to arsenic-free sources relative to those that continued to drink from shallow tubewells. If switching away from shallow tubewells is associated with greater exposure to microbiologically contaminated water, we would expect child mortality to increase with early life exposure for those households in arsenic-contaminated villages who switched to arsenic-free water post-campaign.

Because we are restricting ourselves to within-village comparisons, our analysis sample is implicitly restricted to the 29% of villages in which at least one household is still drinking from an arsenic-contaminated source.[26] Hence, although we cannot observe in the BDHS data whether well water in a specific village is contaminated with arsenic, our sample is necessarily restricted to a subset of villages in which arsenic contamination is present by virtue of the fact that we observe it at least once in the data. Based on the spatial concentration of arsenic deposits in our data from Barisal, for villages in which arsenic is present, the median rate of contamination is 77%, and in only 25% of villages are less than half of shallow tubewells contaminated. Hence, although in the BDHS we are unable to distinguish whether households with clean drinking water in 2004 have switched away from contaminated shallow tubewells or continue to drink from tubewells that were never contaminated, in our subsample of "exposed" villages, we can assume that the majority of households with arsenic-free drinking water in 2004 are households that were encouraged to abandon shallow tubewells.[27] This ambiguity also implies that we are underestimating mortality effects of switching since not all households with arsenic-free water switched sources around the time of the campaign.

Our sample contains all births reported in the birth history module of the DHS that occurred between 1990 and 2004.[28] Our final dataset consists of 19,919 children born in 361 villages of Bangladesh, but our effect is estimated off of 6003 births in 104 villages in which we observe at least one household drinking arsenic-laden water. Our outcome of interest being child and infant mortality, we observe all deaths occurring under 12 months, under 24 months and under 60 months.

We estimate the following difference-in-difference equation for individual i in household j and village v, which includes village fixed effects (θ):

[26] In no villages are all sampled households drinking contaminated water.

[27] Though the BDHS does ask about current drinking water sources, we cannot make use of these data since the category "tubewell" does not clearly distinguish between deep and shallow tubewells, nor does it distinguish between backyard tubewells and tubewells in neighboring houses. Finally, given public health efforts, there is some concern that people misreport drinking water sources they have been told to avoid.

[28] We limit our sample to individuals born in 1990 or later to minimize noise by restricting the comparison to mothers of the same age range and also to minimize measurement error in reported death age. However, the estimates are robust to expanding the period of observation by at least 5 years.

$$Y_{ijv} = \alpha_{ijv} + \gamma * AsFree_{jv} + \delta * EarlyLifeExp_{ijv} + \beta * (AsFree_{jv} + EarlyLifeExp_{ijv}) + \theta_v + _{ijv} \quad (2.2)$$

In this regression, *AsFree* is an indicator that household drinking water is free of arsenic when tested in 2004, our proxy for whether a household switched water sources after the well-testing campaign (in this sense, it is the opposite of the *HighConc* variable of Equation 1). As in the previous set of regressions, we are interested in the coefficient estimate on the interaction between being born after the well-testing campaign and being a "switcher" household (arsenic-free). If our identifying assumption holds, this coefficient captures the change in mortality from switching to a less convenient water source.

As shown in Appendix Table 2.8, regression-controlled means (that account for village fixed effects) of a wide range of household and respondent characteristics are very similar across arsenic-exposed and arsenic-free households. However, to account for potential differences between switchers and non-switchers, our regressions control for the following household and child characteristics: sex, parity, birth year, birth year squared, age of mother at birth, education of mother, education of father, mean birth interval, household wealth (solvency), amount of land owned by household, number of rooms in house, whether household has electricity, and whether Muslim.[29] Standard errors are clustered at the household level.

Table 2.4 presents the coefficient estimates for Equation 2.2 with full controls. We present results for both exposure after 1999 and exposure after 2000, as it is ambiguous how quickly the campaign led to switching. As expected, the coefficient estimate on the interaction term is positive: Among households that are arsenic-free in 2004, one additional year of exposure to the post-campaign environment is associated with at least a 1.8 percentage point increase in likelihood of death within one year, a 2.5 percentage point increase in likelihood of death within two years, and 3.2 percentage point increase in likelihood of death within five years. The two-year mortality estimate is significant at the 10% level and the 5-year mortality estimate is significant at the 5% level, and the magnitudes of the estimates are similar to our estimates from Barisal (Table 2.2). The results are qualitatively similar using the 1999 versus the 2000 cutoff, indicating that switching behavior was spread across both years.

As described earlier, one shortcoming of our *ArsenicFree* measure is that we cannot distinguish switcher households from households that were never exposed to arsenic in groundwater, and so are underestimating the effect of abandoning tubewells potentially by a great deal. Hence, to gain more precision in identifying switchers, in the next set of regressions we make use of information provided by village leaders on the primary source of water for households in each village. In villages in which the primary water source is identified to be anything other than shallow tubewells, households with arsenic-free drinking water are more likely to be switchers than households in villages in which the primary water source is shallow tubewells. Furthermore, we can look separately at switcher households that most likely moved to surface water sources compared to those who most likely moved to deep tubewells in order to estimate the relative impact of switching to alternative sources.

Table 2.5 presents separate regression estimates for these three categories of villages: those in which the main source of drinking water is piped water into or outside of the house (column 1), those in which the main source of drinking water is deep tubewells (column 2), and those in which the main source of drinking water is some type of surface water source (ponds, lakes, streams, etc.) (column 3). As expected, the difference-in-difference estimate is small in magnitude and insignificant in villages in which arsenic-free

[29] Regression estimates without controls produce very similar and in most cases statistically robust results.

households are most likely using piped water, which is relatively safe in terms of exposure to fecal matter, and is large and significant in villages in which arsenic-free households are most likely to be drinking from deep tubewells or surface water sources. This suggests that the patterns we are observing in the DHS data are not driven by convergence in mortality rates over time between switcher and non-switcher households, which is a concern due to the endogeneity of switching behavior. Interestingly, there is little difference in the negative effect of switching away from shallow tubewells when the alternative source is surface versus deep tubewells, possibly indicating that the higher rate of fecal contamination in surface water relative to deep tubewell water at the source reduces to similar levels when measured at the point of use, consistent with previous studies in other settings.

Our final exercise with the BDHS makes use of detailed verbal autopsy data collected for the majority of child and infant deaths reported between 1998 and 2004 in order to verify that the patterns on child mortality we observed in the previous regression estimates are driven by deaths due to an increase in water-borne illnesses, as our interpretation implies.[30] Using these reports, we classify infant and child deaths into proximate causes of death due to water-borne pathogens, pneumonia, and fever, and run regressions analogous to Equation 2.2 in which the dependent variable is now a specific cause of mortality. These estimates are presented in Table 2.5 in columns 4-6. As expected, we observe a significant DID estimate of switching to arsenic-free drinking water on deaths attributable to water-born illnesses, but no concurrent pattern with respect to deaths attributed to fever or pneumonia. Not only does this provide an important consistency check on our interpretation of the child mortality patterns, but it minimizes the likelihood that our estimates reflect simple convergence in infant and child mortality between relatively high and relatively low SES households.

2.5 Conclusion

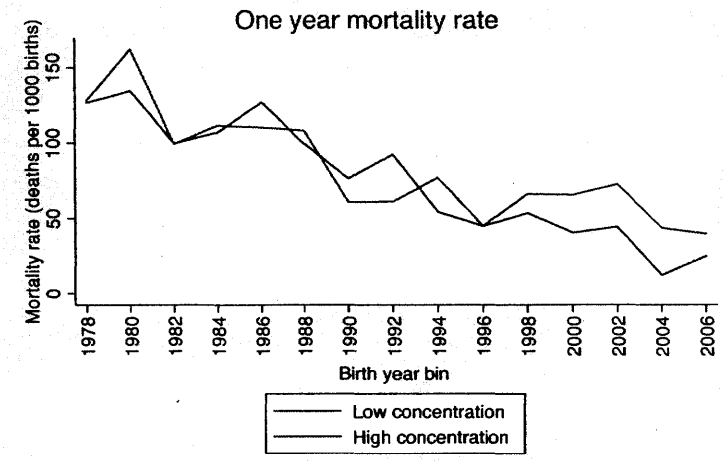
While the arsenic mitigation campaign in Bangladesh has been heralded by the international medical community as a life-saving effort, our estimates indicate substantial negative health consequences of public health efforts to move Bangladeshi households away from shallow tubewells as sources of drinking water. Using data from a district in Bangladesh in which shallow tubewells were readily abandoned for less convenient but arsenic-free deep tubewells, we find that households with an incentive to switch sources experienced a significant increase in the rate of infant and child mortality after arsenic levels were revealed. Hence, evaluation of future public health interventions need to reconsider efforts to convince households to abandon shallow tubewells when alternatives that are equally safe in terms of water-borne pathogens are not readily available.

Perhaps most importantly, our findings provide rigorous evidence of substantial benefits in terms of reductions in infant and child mortality to point-of-source improvements in water quality in a setting of endemic diarrheal disease.

[30] Due to nonresponse, the BDHS verbal autopsy data are only available for 572 of the 606 infant and child deaths that we observe in the data between 1998 and 2004.

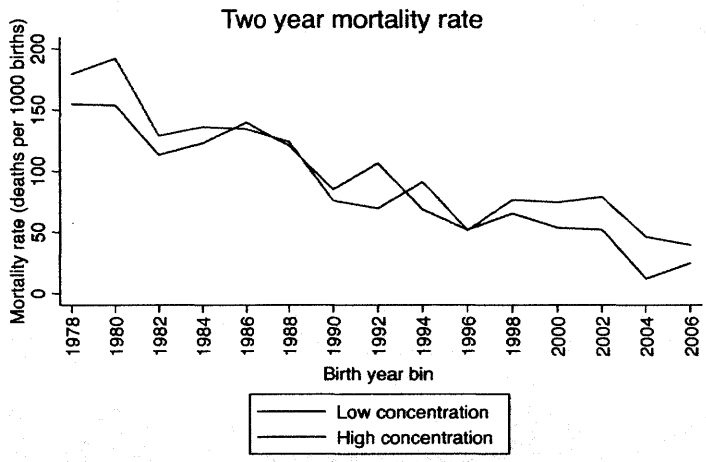
Figures

Figure 2-1: Infant mortality rate (0-12 months)



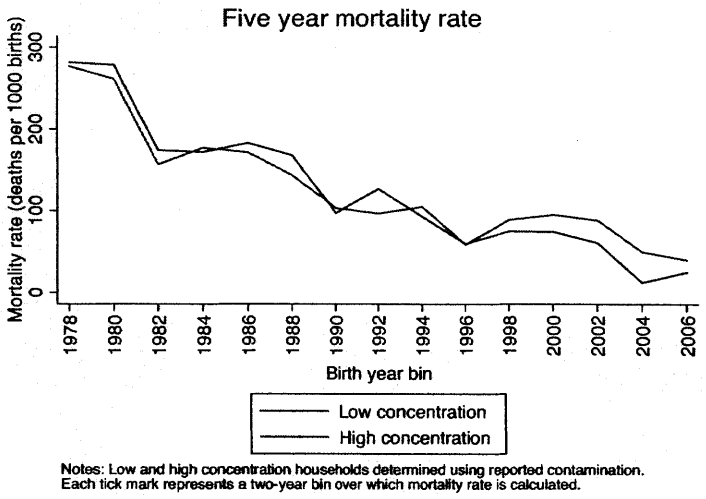
Notes: Data from the 2007 Kishoree Kontha Baseline Survey and 2009 Household Arsenic Survey in Barisal subdistrict, Bangladesh. Sample includes 11,766 births across 2817 households in 162 villages. "Infant mortality rate" is deaths between 0 and 12 months of age per 1000 births observed in each two-year period. Average mortality rates are calculated controlling for village means. "Low conc" are households in which arsenic concentration of closest shallow tube well <60ppb, as measured in the 2009 survey, and "High conc" are households with >60ppb arsenic concentrations.

Figure 2-2: Under two year mortality rate (0-24 months)



Notes: Data from the 2007 Kishoree Kontha Baseline Survey and 2009 Household Arsenic Survey in Barisal subdistrict, Bangladesh. Sample includes 11,766 births across 2817 households in 162 villages. "Under two mortality rate" is deaths between 0 and 24 months of age per 1000 births observed in each two-year period. Average mortality rates are calculated controlling for village means. "Low conc" are households in which arsenic concentration of closest shallow tube well <60ppb, as measured in the 2009 survey, and "High conc" are households with >60ppb arsenic concentrations.

Figure 2-3: Under five year mortality rate (0-60 months)



Notes: Data from the 2007 Kishoree Kontha Baseline Survey and 2009 Household Arsenic Survey in Barisal subdistrict, Bangladesh. Sample includes 11,766 births across 2817 households in 162 villages. "Under five mortality rate" is deaths between 0 and 60 months of age per 1000 births observed in each two-year period. Average mortality rates are calculated controlling for village means. "Low conc" are households in which arsenic concentration of closest shallow tube well <60ppb, as measured in the 2009 survey, and "High conc" are households with >60ppb arsenic concentrations.

Tables

Table 2.1: Sample Means

	Measured Contamination				Reported Contamination			
	Panel I: Exogenous variables							
VARIABLES	(1) High Conc.	(2) N	(3) Low Conc.	(4) T-statistic	(5) High Conc.	(6) N	(7) Low Conc.	(8) T-statistic
Age of mother	38.58	1928/889	38.01	1.5	38.44	2332/485	38.17	0.65
Age of mother at earliest birth	21.03	1928/889	21.14	-0.4	21.03	2332/485	21.22	-0.61
Education of mother	3.43	1928/889	3.47	-0.31	3.46	2332/485	3.37	0.53
Education of father	3.76	1928/889	3.85	-0.44	3.78	2332/485	3.81	-0.15
Mean birth interval	2.28	1928/889	2.29	-0.17	2.27	2332/485	2.32	-0.87
Solvency	0.66	1928/889	0.68	-1.2	0.66	2332/485	0.69	-1.42
Land size (acres)	0.85	1928/889	0.85	0.04	0.85	2332/485	0.85	-0.02
Number of rooms in house	2.75	1928/889	2.75	0.02	2.76	2332/485	2.69	1.28
Electricity	0.39	1928/889	0.39	0.01	0.39	2332/485	0.40	-0.40
Muslim	0.97	1928/889	0.96	1.84	0.97	2332/485	0.96	0.58
Fraction of children living in HH	0.80	1654/775	0.82	-1.6	0.81	2011/418	0.81	-0.18
Respondent's age	42.63	1928/889	42.31	0.64	42.54	2332/485	42.48	0.11
Male respondent	0.16	1928/889	0.16	0.08	0.16	2332/485	0.18	-1.14
Sufficiency of food per week	0.92	1928/889	0.93	-1.37	0.92	2332/485	0.93	-0.49
Outstanding loan	0.54	1928/889	0.54	0.13	0.54	2332/485	0.53	0.58
Years living in house	25.89	1928/889	24.89	1.55	25.73	2332/485	25.21	0.74
Years living in village	30.49	1547/729	30.26	0.3	30.28	1868/408	31.07	-0.94
Mean monthly income of HH	11.37	1928/889	11.35	0.04	11.38	2332/485	11.28	0.21
Head of HH works in agriculture	0.43	1928/889	0.41	0.55	0.41	2332/485	0.46	-1.57
Head of HH works in business	0.15	1928/889	0.16	-0.62	0.16	2332/485	0.14	1.20
	Panel II: Endogenous variables							
Arsenic concentration (ppb)	94.71	1928/889	31.96	54.71	80.61	2332/485	47.50	19.29
Fraction of deaths under 12 mo.	0.07	1928/889	0.05	2.24	64.50	2332/485	44.77	2.84
Fraction of deaths under 24 mo.	0.08	1928/889	0.06	2.73	54.39	2332/485	54.36	2.77
Fraction of deaths under 60 mo.	0.09	1928/889	0.08	2.3	91.49	2332/485	74.09	2.13
M:F sex ratio	0.42	1928/889	0.47	-3.51	0.44	2332/485	0.45	-1.05
Number of offspring in family	4.21	1928/889	4.05	1.67	4.19	2332/485	4.03	1.57
Number of drinking sources used	2.05	1928/889	2.05	0.45	2.05	2332/485	2.04	1.21
Number of cooking sources used	2.24	1928/889	2.25	-0.3	2.25	2332/485	2.22	1.28
Whether closest well tested	0.69	1700/836	0.70	-0.44	0.72	2091/445	0.55	6.88
Whether closest well painted	0.68	1542/772	0.65	1.39	0.71	1922/392	0.49	8.24
Value of house ($)	2050.62	1835/864	2308.14	-2.82	2125.35	2230/469	2169.73	-0.44

Notes:

(1) All averages are calculated controlling for village means.

(2) Sufficiency of food defined as family members taking at least two meals a day last week; solvency defined as last week's expenses being within the budget.

(3) In columns 1-4 ("measured contamination"), high concentration versus low concentration defined according to field test of shallow tubewell closest to residence. High concentration households are those with tubewells that contain arsenic concentrations greater than 60ppb. In columns 5-8 ("reported concentration"), high concentration households are those who report that their well tested positive for arsenic concetration, or (if household has no recollection of well being tested or test result) if closest shallow tubewell currently contains arsenic concentration greater than 60ppb.

Table 2.2: Measured arsenic contamination and early life exposure to post-campaign environment

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00785	0.00606	0.00461	0.00998	0.00763	0.00563	0.00533	0.00132	-0.00105
	[0.00723]	[0.00722]	[0.00712]	[0.00815]	[0.00812]	[0.00797]	[0.00962]	[0.00958]	[0.00940]
Early life exposure	-0.0403***	-0.00116	0.000706	-0.0597***	-0.00495	-0.00763	-0.0941***	0.00453	-0.00675
	[0.00879]	[0.0127]	[0.0127]	[0.00930]	[0.0146]	[0.0146]	[0.0112]	[0.0189]	[0.0188]
High concentration * Early life exposure	0.0142	0.0163	0.0203*	0.0249**	0.0279**	0.0317***	0.0300**	0.0346**	0.0389***
	[0.0117]	[0.0117]	[0.0117]	[0.0121]	[0.0121]	[0.0120]	[0.0139]	[0.0138]	[0.0138]
Sex		0.0292***	0.0279***		0.0284***	0.0267***		0.0261***	0.0239***
		[0.00499]	[0.00499]		[0.00529]	[0.00528]		[0.00596]	[0.00594]
Parity		0.00308*	0.00329		0.00442***	0.00431*		0.00622***	0.00344
		[0.00157]	[0.00216]		[0.00167]	[0.00221]		[0.00176]	[0.00236]
Birth year		-0.00636***	-0.00693***		-0.00781***	-0.00889***		-0.0130***	-0.0150***
		[0.00204]	[0.00207]		[0.00208]	[0.00213]		[0.00207]	[0.00211]
Birth year squared		0.00006	0.00009		0.00007	0.000117		0.000135*	0.000239***
		[0.00007]	[0.00008]		[0.00008]	[0.00008]		[0.00008]	[0.00008]
Age of mother at birth			-0.000243			-0.000197			0.00008
			[0.000540]			[0.000576]			[0.000622]
Mother's education			0.00117			0.00178			0.00257
			[0.00148]			[0.00155]			[0.00165]
Father's education			-0.00146			-0.00163			-0.00208*
			[0.00107]			[0.00114]			[0.00126]
Years since birth of last child			-0.0107***			-0.0134***			-0.0172***
			[0.00178]			[0.00188]			[0.00204]
Solvency			0.000313			-0.00162			0.00178
			[0.00556]			[0.00599]			[0.00633]
Land size (acres)			0.00474***			0.00419**			0.00383**
			[0.00164]			[0.00170]			[0.00192]
No. of rooms in house			-0.0133***			-0.0143***			-0.0152***
			[0.00268]			[0.00298]			[0.00317]
Electricity			-0.00936			-0.0104			-0.0123*
			[0.00615]			[0.00667]			[0.00732]
Muslim			0.0115			0.0208			0.0168
			[0.0145]			[0.0148]			[0.0157]
Monthly income per capita			0.000214			0.000231			0.000324
			[0.000313]			[0.000328]			[0.000356]
Mean among offspring with zero exposure in households with low arsenic concentration	0.07649	0.07649	0.07649	0.09618	0.09618	0.09618	0.14108	0.14108	0.14108
	[0.26582]	[0.26582]	[0.26582]	[0.29489]	[0.29489]	[0.29489]	[0.34819]	[0.34819]	[0.34819]
Observations	11766	11766	11766	11766	11766	11766	11766	11766	11766

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

Table 2.3: Reported arsenic contamination and early life exposure to post-campaign environment

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00754	0.00702	0.00749	0.00703	0.00625	0.00640	-0.00747	-0.00924	-0.00926
	[0.00820]	[0.00815]	[0.00800]	[0.00919]	[0.00918]	[0.00898]	[0.0113]	[0.0113]	[0.0110]
Early life exposure	-0.0520***	-0.0123	-0.00779	-0.0679***	-0.0121	-0.0126	-0.115***	-0.0146	-0.0237
	[0.0112]	[0.0147]	[0.0147]	[0.0122]	[0.0168]	[0.0168]	[0.0142]	[0.0207]	[0.0208]
High concentration * Early life exposure	0.0256*	0.0266**	0.0266**	0.0297**	0.0306**	0.0311**	0.0488***	0.0505***	0.0515***
	[0.0131]	[0.0131]	[0.0132]	[0.0139]	[0.0140]	[0.0141]	[0.0159]	[0.0161]	[0.0161]
Sex		0.0291***	0.0278***		0.0282***	0.0265***		0.0259***	0.0236***
		[0.00499]	[0.00498]		[0.00530]	[0.00528]		[0.00596]	[0.00594]
Parity		0.00312**	0.00334		0.00446***	0.00436**		0.00625***	0.00352
		[0.00157]	[0.00216]		[0.00167]	[0.00221]		[0.00176]	[0.00237]
Birth year		-0.00641***	-0.00696***		-0.00790***	-0.00895***		-0.0131***	-0.0150***
		[0.00204]	[0.00207]		[0.00208]	[0.00213]		[0.00207]	[0.00211]
Birth year squared		0.00006	0.00009		0.00007	0.000119		0.000138*	0.000240***
		[0.00007]	[0.00008]		[0.00008]	[0.00008]		[0.0008]	[0.00008]
Age of mother at birth			-0.000244			-0.000198			0.00006
			[0.000540]			[0.000576]			[0.000622]
Mother's education			0.00116			0.00178			0.00263
			[0.00149]			[0.00155]			[0.00165]
Father's education			-0.00147			-0.00165			-0.00215*
			[0.00106]			[0.00114]			[0.00126]
Years since birth of last child			-0.0106***			-0.0133***			-0.0170***
			[0.00177]			[0.00187]			[0.00203]
Solvency			0.000415			-0.00162			0.00152
			[0.00557]			[0.00600]			[0.00633]
Land size (acres)			0.00476***			0.00421**			0.00387**
			[0.00163]			[0.00169]			[0.00192]
No. of rooms in house			-0.0134***			-0.0144***			-0.0152***
			[0.00268]			[0.00298]			[0.00317]
Electricity			-0.00925			-0.0103			-0.0122*
			[0.00616]			[0.00668]			[0.00733]
Muslim			0.0118			0.0217			0.0178
			[0.0145]			[0.0149]			[0.0157]
Monthly income per capita			0.000216			0.000235			0.000330
			[0.000315]			[0.000330]			[0.000358]
Mean among offspring with zero exposure in households with low arsenic concentration	0.07826	0.07826	0.07826	0.09715	0.09715	0.09715	0.14929	0.14929	0.14929
	[0.26867]	[0.26867]	[0.26867]	[0.29627]	[0.29627]	[0.29627]	[0.35653]	[0.35653]	[0.35653]
Observations	11766	11766	11766	11766	11766	11766	11766	11766	11766

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if household reported in survey that their drinking water source tested positive for arsenic contamination or, if respondent could not recall whether well was tested or test results, if shallow tubewell closest to the residence revealed in field test to have arsenic concentration above 60 ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

Table 2.4: Arsenic in drinking water in 2004 and exposure to post-campaign environment

	1999 campaign cutoff			2000 campaign cutoff		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Death under 12 mo.**	**Death under 24 mo.**	**Death under 60 mo.**	**Death under 12 mo.**	**Death under 24 mo.**	**Death under 60 mo.**
Arsenic free	-0.00949	-0.0178	-0.0190	-0.00833	-0.0171	-0.0176
	[0.0130]	[0.0138]	[0.0146]	[0.0126]	[0.0134]	[0.0141]
Born after campaign	-0.00475	-0.00539	-0.0175	-0.00838	-0.0148	-0.0232
	[0.0153]	[0.0164]	[0.0168]	[0.0157]	[0.0169]	[0.0172]
Arsenic free * Born after campaign	0.0176	0.0245*	0.0317**	0.0177	0.0287*	0.0347**
	[0.0139]	[0.0149]	[0.0151]	[0.0141]	[0.0152]	[0.0154]
Sex	0.0166***	0.0149***	0.0115***	0.0166***	0.0149***	0.0116***
	[0.00384]	[0.00400]	[0.00424]	[0.00384]	[0.00400]	[0.00424]
Parity	0.00880***	0.00990***	0.00997***	0.00877***	0.00987***	0.00994***
	[0.00209]	[0.00222]	[0.00229]	[0.00209]	[0.00222]	[0.00229]
Birth year	-0.00003	0.00325	0.00207	-0.000704	0.00211	0.00166
	[0.00556]	[0.00577]	[0.00603]	[0.00633]	[0.00660]	[0.00695]
Birth year squared	-0.00009	-0.000215	-0.000210	-0.00006	-0.000164	-0.000187
	[0.000168]	[0.000174]	[0.000181]	[0.000191]	[0.000200]	[0.000209]
Age of mother at birth	-0.00189***	-0.00215***	-0.00200***	-0.00188***	-0.00213***	-0.00199***
	[0.000586]	[0.000617]	[0.000646]	[0.000587]	[0.000618]	[0.000646]
Mother's education	-0.00260***	-0.00295***	-0.00272***	-0.00261***	-0.00297***	-0.00273***
	[0.000802]	[0.000842]	[0.000874]	[0.000803]	[0.000843]	[0.000875]
Father's education	0.000004	0.00007	-0.000570	-0.000002	0.00006	-0.000578
	[0.000629]	[0.000655]	[0.000677]	[0.000629]	[0.000654]	[0.000677]
Years since birth of last child	-0.00147*	-0.00162*	-0.00221***	-0.00147*	-0.00162*	-0.00221***
	[0.000799]	[0.000826]	[0.000851]	[0.000799]	[0.000827]	[0.000851]
Solvency	0.000256	0.000338	0.00128	0.000284	0.000384	0.00131
	[0.00456]	[0.00476]	[0.00498]	[0.00457]	[0.00477]	[0.00499]
Land size (acres)	0.00121	0.000993	0.000639	0.00122	0.00100	0.000644
	[0.000869]	[0.000872]	[0.000911]	[0.000870]	[0.000873]	[0.000912]
No. of rooms in house	-0.00448***	-0.00530***	-0.00660***	-0.00449***	-0.00531***	-0.00661***
	[0.00137]	[0.00141]	[0.00147]	[0.00137]	[0.00141]	[0.00147]
Electricity	0.000926	0.000491	-0.00230	0.000866	0.000409	-0.00237
	[0.00556]	[0.00576]	[0.00604]	[0.00556]	[0.00577]	[0.00604]
Muslim	-0.00334	-0.000825	0.00286	-0.00326	-0.000673	0.00297
	[0.00943]	[0.00962]	[0.0101]	[0.00943]	[0.00961]	[0.0101]
Mean among offspring born before the campaign in households with high arsenic concentration	0.0801173 [0.27149]	0.0897654 [0.28586]	0.1040445 [0.30533]	0.0789935 [0.26974]	0.0884559 [0.28397]	0.1018434 [0.30245]
Observations	19919	19919	19919	19919	19919	19919

Notes:

(1) OLS regressions, linear probability models. Data from the 2004 Bangladesh Demogrpahic and Health Survey. Observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) "Arsenic-free" defined from survey field test of household drinking water, and equal to 1 if arsenic concentration less than 50ppb.

(3) In columns 1-3, "Born after campaign" equal to 1 if child born after 1999; in columns 4-6, "Born after campaign" equal to 1 if child born after 2000.

Table 2.5: Heterogeneity according to village-level water source and cause of death

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Death under 60 months			Water related death	Death by pneumonia	Death by fever
	Subsample: All individuals in villages where primary water source is piped water	*Subsample: All individuals in villages where primary water source is deep tubewell*	*Subsample: All individuals in villages where primary water source is surface water*	*Subsample: All individuals aged 8 yrs and below in villages where primary water source is deep tubewell or surface water*		
Arsenic free	-0.0444	-0.0553	-0.0388	-0.0960	-0.0183	-0.0160
	[0.0423]	[0.0528]	[0.0475]	[0.0659]	[0.0397]	[0.0531]
Born after 2000	0.0117	-0.101*	-0.0370	-0.0879	-0.0258	-0.0383
	[0.0490]	[0.0561]	[0.0661]	[0.0673]	[0.0426]	[0.0630]
Arsenic free * Born after 2000	0.0327	0.100**	0.139**	0.110*	0.0508	0.0475
	[0.0465]	[0.0485]	[0.0547]	[0.0638]	[0.0403]	[0.0582]
Sex	0.0229	0.00350	-0.00949	-0.00184	-0.00683	-0.0155
	[0.0147]	[0.0126]	[0.0221]	[0.00966]	[0.00876]	[0.0105]
Parity	0.00999	0.00138	-0.0134	0.00215	-0.00260	0.00435
	[0.00828]	[0.00582]	[0.00984]	[0.00406]	[0.00505]	[0.00513]
Birth year	0.0232	0.0317*	0.0516	0.134**	0.0563	-0.0240
	[0.0230]	[0.0190]	[0.0356]	[0.0568]	[0.0486]	[0.0703]
Birth year squared	-0.000775	-0.000925	-0.00174	-0.00322**	-0.00139	0.000376
	[0.000700]	[0.000581]	[0.00109]	[0.00131]	[0.00114]	[0.00157]
Age of mother at birth	0.000183	-0.000462	0.00122	0.000689	0.000901	-0.000719
	[0.00211]	[0.00165]	[0.00315]	[0.00104]	[0.00124]	[0.00113]
Mother's education	-0.00566**	-0.00392	-0.00707	-0.00216	-0.00152	-0.00418*
	[0.00241]	[0.00280]	[0.00493]	[0.00200]	[0.00174]	[0.00237]
Father's education	0.00403*	-0.00353*	-0.00484	-0.00229	-0.000778	0.000982
	[0.00214]	[0.00183]	[0.00311]	[0.00142]	[0.000994]	[0.00156]
Years since birth of last child	-0.00391	-0.000123	0.00344	0.000861	0.00134	0.00176
	[0.00307]	[0.00232]	[0.00643]	[0.00171]	[0.00144]	[0.00156]
Solvency	0.000568	0.00529	0.00903	0.00721	-0.00963	-0.00881
	[0.0165]	[0.0152]	[0.0259]	[0.0118]	[0.00845]	[0.0114]
Land size (acres)	-0.000870	-0.00133	-0.00361**	-0.000997	-0.00125	-0.000303
	[0.00163]	[0.00168]	[0.00155]	[0.000831]	[0.00101]	[0.000857]
No. of rooms in house	-0.00901**	0.00148	-0.000155	0.00257	0.00415	0.00561
	[0.00440]	[0.00419]	[0.00627]	[0.00362]	[0.00306]	[0.00391]
Electricity	-0.0372	0.00667	-0.0512	0.0247	-0.00625	-0.0104
	[0.0276]	[0.0187]	[0.0337]	[0.0167]	[0.0153]	[0.0175]
Muslim	0.00414	-0.0317	0.0399	0.0519**	0.0487	0.0152
	[0.0285]	[0.0359]	[0.0465]	[0.0260]	[0.0331]	[0.0319]
Mean among offspring born before the campaign in households with high arsenic concentration	0.0833333	0.15	0.1506849	0.1111111	0.037037	0.0740741
	[0.28233]	[0.36008]	[0.36022]	[0.32036]	[0.19245]	[0.26688]
Observations	1444	1787	928	1122	1122	1122

Notes:

(1) OLS regressions, linear probability models. Data from the 2004 Bangladesh Demogrpahic and Health Survey. Observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) "Arsenic-free" defined from survey field test of household drinking water, and equal to 1 if arsenic concentration less than 50ppb.

(3) In columns 1-3, "Born after campaign" equal to 1 if child born after 1999; in columns 4-6, "Born after campaign" equal to 1 if child born after 2000.

(4) Subsample for (1) is households with arsenic-contaminated drinking water or households whose primary water source is reported to be piped water. Subsample of (2) is households with arsenic-contaminated drinking water or households whose primary water source is reported to be a deep tubewell. Subsample of (3) is households with arsenic-contaminated drinking water or households whose primary water source is reported to be surface water (surface wells, ponds, lakes, streams, etc.)

(5) Subample for (4), (5), and (6) are individuals born after 1997 and households with arsenic-contaminated drinking water or whose primary water source is reported to be a deep tubewell or surface water.

Chapter 3

The Impact of Training Informal Providers on Clinical Practice in West Bengal, India: A Randomized Controlled Trial

Abstract

Background: Informal providers in the private sector with little or no formal medical training comprise the majority of India's rural primary healthcare workforce. Programs that improve their quality of care can complement more effective regulation.

Methods: We conducted a single-blind, randomized controlled trial of a training program that consisted of 72 sessions over nine months designed to improve quality of care for multiple conditions. The trial allocated 304 informal providers (IPs) from 203 villages in West Bengal, India, equally to treatment or control groups with providers in the treatment group offered participation in the training program. Clinical practice was assessed at endline using unannounced standardized patients as well as clinical observations. Intention-to-treat and instrumental variables analysis were used to assess the impact of the training program. The study was registered on ClinicalTrials.gov at NCT02291575.

Findings: Attendance in the training program, which lasted between January and October 2013, was 56% (95% CI: 51%-62%) for IPs in the treatment group with no contamination from the control group. Primary outcome data were available for 267 of 304 participants (87.8%) at endline. Providers allocated to the treatment group were 4.1 percentage points (95% CI: 0.017-0.065) more likely to adhere to condition-specific checklists and to correctly manage the conditions presented by SPs (OR: 1.40, 95% CI: 1.017-1.931), but there were no differences in the unnecessary use of antibiotics, injections, or polypharmacy.

Interpretation: The program improved IPs adherence to checklists and correct case management, but there was no reduction in unnecessary or potentially harmful practices. Given the limited presence of public sector clinics in such regions, IP training can be a valuable complement to effective regulation.

Funding: National Rural Health Mission, West Bengal, Knowledge for Change Program, World Bank and Bristol Myers' Squibb Foundation's "Delivering Hope and Center for Excellence" program.

Research in context

Evidence before this study

The limited availability of formal providers in rural India implies that informal providers (IPs) are regularly asked to provide care for a range of conditions. As detailed in the reviews of Shah et al. (2010), Bloom et al. (2011), and Suhinaraset et al. (2013), existing evaluations show that interventions can improve quality of care for specific conditions, but the impact of non-disease specific training on clinical practice is currently unknown. Furthermore, while most studies focus on the impact of training on pharmacists and medicine sellers, there is a dearth of rigorous research on the impact on IPs. This is a key gap, as non-specific training for IPs is increasingly discussed as a valuable complement to better regulation and public provision of health care.

Added value of this study

Our randomized controlled trial assesses the impact of a multi-topic training course for IPs consisting of 72 training sessions over 9 months. The study is novel in that the quality of clinical practice was assessed using standardized (simulated) patients and the designers and implementers of the training course did not know the conditions that the IPs would be assessed on. This eliminated emphasis on the evaluated conditions during the training. The results suggest that training improved adherence to checklists and correct case management across a range of conditions but had no impact on unnecessary care such as the use of antibiotics.

Implications of all the available evidence

Our results demonstrate that necessary and unnecessary care respond in different ways to training and that multi-topic training can lead to improvements in the former. Such information is critical to designing programs that educate IPs, who are responsible for the bulk of primary care in India. The methodology adopted here with the simultaneous use of SPs and the blinding of implementers from the conditions to be evaluated has wide applicability for evaluating a range of potential interventions.

3.1 Introduction

India, like many low-income countries, has a pluralistic health care delivery system with a significant deficit in adequately trained health care human resources to match the population requirements (Chen et al. 2004, Bloom et al. 2011). More than 75% of ambulatory outpatient care is managed by private providers with varying degrees of training, competence and quality of care (Government of India 2005, Das et al. 2015). Amongst private providers, those who are fully trained tend to concentrate in urban areas; consequently, curative primary care in rural India is largely provided by private informal providers (IP), who account for the majority of health care providers in the average Indian village (De Costa and Diwan 2007, Gautham et al. 2014, Banerjee et al. 2004, MAQARI 2011). These providers are self-employed with little or no formal medical training and function on a fee-for-service basis with scarce de facto regulation (Sudhinaraset et al. 2013). While their suboptimal quality of care is widely discussed, their strong community ties and significant presence where few alternatives exist necessitate their consideration in healthcare policy discussions (George and Iyer 2013).

Better training of IPs, which would allow them to integrate into existing health care systems and improve their quality of care, is increasingly discussed as a viable first step to complement more effective regulation (Telegraph 2014). To be successful, such programs have to account for the varied clinical tasks that IPs may be required to provide for patients in the rural setting. These can range from (a) triage into higher quality care to (b) immediate treatment to (c) diagnosis and maintenance care for chronic patients. Training programs should therefore demonstrably improve quality across a range of conditions that IPs are asked to manage.

Existing evaluations have shown that training can improve practice among pharmacists, medicine sellers and IPs, but have not been designed to address these multiple objectives (Shah et al. 2010, Sudhinaraset et al. 2013). First, most training programs evaluated to date are short in duration and designed to improve care for specific conditions rather than on multiple topics (Stenson et al. 2001, Abuya et al. 2009, Adu-Sarkodie et al. 2000, Peltzer et al. 2006). Second, evaluations that assess practice immediately or shortly after the completion of training leave open valid questions of persistence over the medium to long term (Adu-Sarkodie et al. 2000, Chakraborty et al. 2000). Third, assessments of training are confounded if providers self-select into training programs. Fourth, evaluations based on clinical observations with real patients may be biased if there are intervention-specific Hawthorne effects whereby trained providers utilize their training only when an observer is present (Rowe et al. 2012, Leonard and Masatu 2010).

We present estimates of the impact of a semi structured multi-topic training program on the quality of healthcare delivered by IPs in rural West Bengal 3-6 months after the completion of training and 9-10 months after the completion of the illness-specific portion of the training. Using unannounced standardized patients in conjunction with a randomized study design, we provide robust evidence that helps address several of the limitations in existing studies.

3.2 Methods

3.2.1 Ethics Statement

The study protocol was approved by the Massachusetts Institute of Technology Institutional Review Board, Cambridge, Massachusetts, USA (study number 1209005234) and the Institute for Financial Management and Research Institutional Review Board, Chennai, India (IRB00007107). All surveyed IPs were informed

by enumerators of the study objectives, how the data would be used, confidentiality, risks, benefits, and respondent rights, and provided written consent prior to surveying. The study was registered as a randomized controlled trial on ClinicalTrials.gov at NCT02291575.

3.2.2 Trial Design

The training program was designed and executed by The Liver Foundation, a public health organization based in West Bengal. Between May and June 2012, the organization invited 360 providers in 203 villages across Birbhum district (one of 20 districts in the state), who had been practicing for at least 5 years, to participate in the training. Of these, 304 agreed to participate, and in September 2012 they were randomized into treatment and control groups with equal allocation to both groups. Providers randomized into treatment were offered admission to the Liver Foundation Rural Healthcare Practitioner Training Program, which commenced in January 2013.

Baseline data were collected prior to the randomization and included information on provider background and practice characteristics as well as medical knowledge evaluated through vignettes for multiple conditions. Endline data collection commenced three months after the completion of training. Unannounced standardized patients (SPs), increasingly used to assess condition-specific quality of care in low-income countries, were sent to the entire study population (Das et al. 2015, Mohanan et al. 2015, Sylvia et al. 2014, Das et al. forthcoming). To benchmark the performance of trained IPs, the SPs were also sent to all 11 public Primary Health Centers (PHCs) in the 203 villages. After SP data collection, one-day long clinical observations were conducted with each provider, and baseline measures were recollected. As the use of SPs and clinical observation could have primed providers to expect such cases in the future, both were restricted to the endline survey data collection only. Further details on measurement tools are presented in the Supplemental Appendix.

3.2.3 Randomization and masking

Participants were randomized using a random number generator. RH conducted the randomization and the Liver Foundation notified IPs allocated to treatment that they were eligible for training. As the identity of all participants was known prior to the randomization, we could verify that there was no contamination in the post-randomization allocation by The Liver Foundation.

Program implementers and impact researchers were not blinded to group assignment, as invitation to and enrollment in the training program was public knowledge. Program implementers were, however, blinded from the final SP conditions that the IPs would be evaluated on, and therefore could not tailor their training to the conditions that would eventually be evaluated. SPs remained blinded to group assignment for the duration of their visits. Field surveyors were also blinded to group assignment at endline, although providers may have revealed their group status during clinical observations.

3.2.4 Study population

In the study population (Table 3.1, Column 1), 290 of 304 providers (95%) were male with a mean age of 40 years (95% CI: 39.16-41.22) and 13.1 years of experience (95% CI: 12.2-14.08); 187 of 304 (62%) had completed high school and 228 of 304 (75%) reported no formal training. The remainder were in possession of certificates of limited legal validity in terms of medical training, but perceived to be of help in the rural medical market.

3.2.5 Intervention program

The Liver Foundation training consisted of a nine-month module with 72 sessions and 150 teaching hours conducted between January and October 2013. The supplement details the full structure and curriculum for the program. Classes were conducted using a semi structured curriculum, two days a week for two hours each in the district capital (9km to 45km from the providers' clinics), and were taught by certified medical doctors. IPs were not financially compensated for training, but they were provided a transportation allowance and meal at each session. Throughout the training, IPs continued to operate their clinics.

The training included a wide variety of topics ranging from anatomy and physiology to first aid in trauma and public health programs. Theoretical classes were supplemented with periodic patient simulations and clinical demonstrations of problems encountered in primary care. Emphasis was placed on basic medical conditions, triage, and avoidance of harmful practices. Training on specific health conditions was conducted between the third and sixth months of the course, preceded by an introduction to medicine and followed by a focus on community medicine and humanity in medicine. From the beginning, the Liver Foundation stressed that IPs would not receive any certificates upon completion of the training and should not think of themselves as qualified doctors.

3.2.6 Outcome definition and measurement

Quality of clinical care was assessed between January and May 2014 using unannounced SPs and clinical observation. Unannounced SPs depicted one of three conditions: chest pain highly suggestive of coronary artery disease, respiratory distress suggestive of asthma, or diarrhea suggestive of dysentery in a child asleep at home. These conditions had been previously validated in the Indian context and were designed to mimic the fact that IPs have to deal with a wide range of conditions given the paucity of care from other sources (Das et al. 2015). These range from immediate referral for chest pain to identification of a lung condition needing primary care and referral to the provision of treatment (ORS in diarrhea). SPs were recruited from the local community and trained for 120 hours to present symptoms consistent with these conditions. After the completion of each SP interaction, SPs were debriefed using a structured questionnaire that recorded history taking, examinations completed, and treatments provided.

Our primary outcomes are condition-specific metrics obtained from the SP interactions. We assessed potential improvements in necessary care through condition-specific checklists of recommended care and rates of correct case management (Jindal et al. 2005, Integrated Management Booklet, Ischaemic Heart Disease Booklet). We also assessed potential declines in unnecessary or harmful care through the use of antibiotics, injections and polypharmacy. Antibiotic use was assessed for all three conditions as well as for asthma and angina alone, as in the latter two instances antibiotics are almost never required. The Supplemental Appendix describes the recruitment of SPs, the conditions assessed and the condition-specific recommendations for the checklist and case management.

Whether the case was managed correctly was assessed in two ways. First, we defined correct case management for each interaction (1 if correctly managed, 0 otherwise) on the basis of critical case-specific actions, even if paired with additional unnecessary treatments. To account for the widespread use of unnecessary medicines and to allow for finer assessments, three independent medical professionals (blinded to provider identity) at a leading teaching hospital in the state capital, Kolkata, also rated the quality of case management using a five-item Likert scale ranging from "Lowest quality case management" (1) to "Case management of choice" (5). Using these ratings, we construct two additional outcome variables: whether

case management was "Average quality or higher" and whether case management was the "Lowest quality."

These condition-specific quality metrics were complemented with secondary outcomes that were not condition-specific but that have been shown to be related to higher quality care (Das et al. 2015). The general measures of care assessed included consultation length, history taking, and examinations performed, elicited from both the SPs and clinical observations. We recognize that clinical observations may be subject to Hawthorne effects as well as potential confounds: e.g. the patient population may be different for trained and untrained providers. It is also the case that the true underlying patient condition remains unknown to the observer. Nevertheless, clinical observations provide information on a broader sample of conditions that affect real patients and allow assessments of the impact of training on the actual use of injections and IVs, which were intentionally avoided by SPs.

3.2.7 Statistical methods

The study was designed to have 80% power to detect a 0.3 standard deviation effect size on a standardized baseline competence score with a two-sided alpha of 5%. A 0.3 standard deviation effect size was identified as the relevant minimum detectable effect size based on past research.

To identify the impact of the program on clinical practice we focus on results from the standardized patients and clinical observations. We present age-adjusted intention-to-treat (ITT) estimates with standard errors adjusted for clustering at the village (SP data) or provider (clinical observation data) level. ITT estimates assess the impact of assignment to treatment on provider practice and help answer the question: "What is the impact of the training program, bearing in mind that not all providers will take up the offer and not every provider will complete the program?" We also present instrumental variable (IV) estimates, which are widely used in the program evaluation literature and increasingly used in assessments of health interventions (King et al. 2009, McClellan et al. 1994). IV estimates the effect of the program on those who take up the training offer, or the treatment on the treated. However, unlike the "as-treated" estimator (the difference in means among the treated and untreated), the IV estimator adjusts for potentially non-random compliance by instrumenting attendance rates with assignment to treatment, which, being (a) randomly determined and (b) correlated with attendance, satisfies the criteria for a valid instrument (Dunn et al. 2005).

For dichotomous outcome variables, we present relative effect sizes using odds-ratios and absolute effect sizes using marginal effects from the logistic regression model as well as the linear regression model. The instrumental variables estimates presented are the marginal effects from the linear instrumental variables model for both continuous and dichotomous variables. Model specifications and justifications are detailed in the Supplemental Appendix. Data were analyzed using Stata 13 (Stata Corp, Texas) and study questionnaires and protocols are available upon request.

3.2.8 Role of funding source

The study was funded by the National Rural Health Mission, West Bengal and The World Bank's research department. Bristol Myers' Squibb Foundation funded the Liver Foundation in the development of content and curriculum for the training. The funding source did not have any role in the collection of data, review of findings or any say in the production of this article.

3.3 Results

3.3.1 Attrition, compliance with randomization, and baseline balance

Participants were recruited between May 10 to June 15, 2012. Figure 3-1 summarizes the participant flow and loss to follow-up in the original cohorts. Of the study population of 304, SPs were completed for 267 providers (87.8%). Amongst those lost to follow-up, 22 were no longer practicing and 7 had either permanently relocated or were not seeing patients at the time of the endline. In the remaining eight cases, the SP was seen by another IP in the same clinic who was also in the study sample (typically, a husband-wife team). There were no statistically significant differences in provider characteristics by treatment status between providers assessed at endline and those who were not (Appendix Table A3.1). All baseline covariates were well balanced (Table 3.1).

3.3.2 Participation in training

Mean attendance over the duration of the training course was 56% (95% CI: 51%-62%). The mean attendance reflects differences across providers and across days. Appendix Table A3.2 shows that 33 of 152 providers (21.7%) attended less than 10% of the sessions, but 65 of 152 (42.8%) attended more than 80%, with the remainder in between. Attendance in each session varied between 55% and 70% and was negatively correlated with distance to the training site and rainfall on the day of the training (Appendix Figure A3-1 and Table A3.3).

3.3.3 Impact of training using standardized patients

Table 3.2 reports the results. The mean SP interaction in the control group lasted 3.2 minutes (95% CI: 3.08-3.43) with 27.3% (95% CI: 25.9-28.7) adherence to the checklist of recommended questions and examinations (Table 3.2). Correct case management was offered in the control group in 52% of interactions (95% CI: 47.1-57) with independent raters grading 11.4% (95% CI: 8.6-14.9) of all treatments in the control group as "average quality or higher." Polypharmacy and antibiotic use, unnecessary in all cases, was high.

At endline (Table 3.2, Panel B), intention-to-treat estimates show that correct case management was higher among providers assigned to treatment (OR: 1.40, 95% CI: 1.012-1.93) as was the likelihood of providing "average quality or higher" treatment (OR: 1.68, 95% CI: 1.15-2.46). Higher correct case management rates were accompanied by greater adherence to condition-specific checklists (Table 3.2, Panel A) among trained providers (4.1 percentage points, 95% CI: 0.017-0.065). The marginal effects from the linear instrumental variable specifications (Table 3.2, Panel B, Column 4) show that full attendance increases correct case management rates by 13.3 percentage points (95% CI: 0.009-0.258). Both ITT and IV estimates are unable to detect any statistically significant differences in the use of antibiotics, injections, polypharmacy, or the likelihood of very low quality case management as assessed by the independent raters (Table 3.2, Panel C).

Each component of correct case management was higher for providers assigned to training (Figure A3-2). Across the three conditions presented by the SPs, providers assigned to training were more likely to prescribe inhalers and refer for respiratory distress; recommend Oral Rehydration Salts for the child with diarrhea; and refer and administer aspirin for chest pain. However, these results are imprecisely estimated and therefore generally not statistically significant, except in the use of inhalers for respiratory distress and correct case management for chest pain.

3.3.4 Patient observation

Data from 2311 observations of clinical practice with real patients show similar improvements (Table 3.3) with increases in the number of examinations conducted and the number of history questions asked. Finally, as in the case of SPs, there is no evidence of any difference in the use of polypharmacy, injections or antibiotics.

3.3.5 Benchmarking with public care in the geographical area

Relative to public sector providers in the same villages, both trained and untrained IPs were more likely to adhere to condition-specific checklists and spend more time with patients (Table 3.4). Polypharmacy, offers of injections, and antibiotic prescriptions were also lower among IPs relative to the public providers. However, public sector doctors were more likely to correctly manage a case than untrained IPs, and this difference is halved for IPs in the treatment group. Neither difference is statistically significant. The likelihood of average or higher quality case management was equally high among public providers and IPs in the treatment group, although the likelihood of very low quality case management was lower for public doctors.

3.3.6 Program costs

The total costs of the program included rent for the training center, staff salaries, and materials, travel stipend, and food allowance for the providers. Assuming at least 50% attendance (compared to the 56% actually attained), the cost per-student of the program was Rs.10528 ($175 at $1=Rs.60).

3.4 Discussion

This study demonstrates that multi-topic training over a nine-month period increased adherence to checklists and correct treatments for three tracer conditions that neither the providers in the study population nor the implementing group for the program were aware would be the subject of the test. The results provide the first evidence on the impact of non-disease specific training for IPs, a program that is being considered by several Indian states as a complement to regulation (Varghese 2010).

SPs do have limitations, as they are restricted to presenting conditions where (a) the lack of physical findings does not confound diagnosis of the underlying condition; (b) invasive examinations are not required; and (c) a child need not be present at the clinic. Therefore, we cannot extend these findings to all cases that IPs may be required to manage.

The results from SPs also pertain to the specific characteristics of the SPs presented; for instance, in our case, the recruited SPs were more educated than the average patient. SPs, though local, were also clearly from outside the village, which may have biased care away from strategies that require follow-up visits. Although this does not affect internal validity since all providers faced the same SPs, we cannot extrapolate the findings to the entire patient population that IPs treat in their clinics. Previous studies find little variation by SP characteristics on provider behavior, and the parallel evidence of greater history taking and physical examinations for real patients among trained providers suggests that the SP findings may hold for a wider set of illnesses and patients (Das et al. forthcoming). However, direct evidence for such an assertion requires using a larger number of SPs with greater diversity in the SP population or measuring population health outcomes. Finally, the impact of training beyond the nine-month period is uncertain and depends on whether the effect of training decays over time and the rate at which such decay occurs.

Despite these limitations, the use of SPs to detect changes in actual provider practice across multiple conditions is novel, allowing us to assess the impact of the program across multiple conditions. When SP cases are blinded from the implementers and SPs are blinded from group assignment, as in the case of the present study, there is less scope for "teaching to the test" or bias arising from knowledge of treatment assignment. Although select previous studies have used SPs to assess program impact, training in these studies was typically on the condition that was eventually assessed. In our case, the implementers did not know what conditions would be assessed and the method therefore allows us to widen the scope of such evaluations to a broader set of interventions that alter clinical practice for a range of conditions that IPs are asked to manage. This is particularly important as SPs allow us to detect differences for the same patient across treatment and control groups free of observation bias, an important alternative when patient data (for instance, patient charts) are absent and data based on clinical observations may be biased due to Hawthorne effects or differences in patient mix arising from the intervention. Further, because researchers know the conditions presented by SPs, each case can be evaluated for condition-specific metrics including adherence to checklists, correct treatment rates, and the use of unnecessary medications.

The results show that the clinical practice of IPs, despite an average of thirteen years of experience, is highly malleable across multiple topics and that training IPs can improve the quality of care for the rural population of India. In most Indian villages, as in our sample, there is little access to care from fully qualified providers in the public or private sector, with the bulk of care delivered by IPs (Gautham et al. 2014, Banerjee et al. 2004). In our study setting, the IPs came from 203 villages, but there were only 11 public PHCs that patients could access. Parallel to our findings, another study shows that IPs use fewer unnecessary medications and antibiotics than fully trained doctors, but their correct treatment rates represent a 15-20% deficit relative to trained doctors (King et al. 2009). Training reduced this deficit by half. Further, there is no evidence that trained IPs were less likely to refer patients for chest pain and respiratory distress, as referrals among trained IPs are higher, though the difference is not statistically significant. This suggests that training led to both better case management in the clinic and more timely referrals. Finally, reducing the distance to the training center would increase attendance and thus the beneficial impacts of training (Figure A3-3). For instance, if the training center was 5 km from each clinic, expected attendance rates would be above 75%.

At the very least, the program sets one benchmark for multi-topic quality improvement efforts in resource-poor settings, suggesting that a $175 investment per-provider ($2.3 per session) can increase correct treatment rates by 7.8 percentage points for a range of conditions. We could not find cost estimates for 9 of 11 evaluations of training programs for pharmacists or IPs, but these costs are comparable to those that we found; in one study the cost was $50 for two days of training14 while in the other it was $200 for a longer duration program designed to improve specific practices among pharmacists (cost calculations for the latter may not have included the time of government employees) (Chalker et al. 2002). More optimistically, multi-topic medical training may offer a short-run strategy to improved healthcare provision to complement critical investments in medical education infrastructure and the quality of public care.

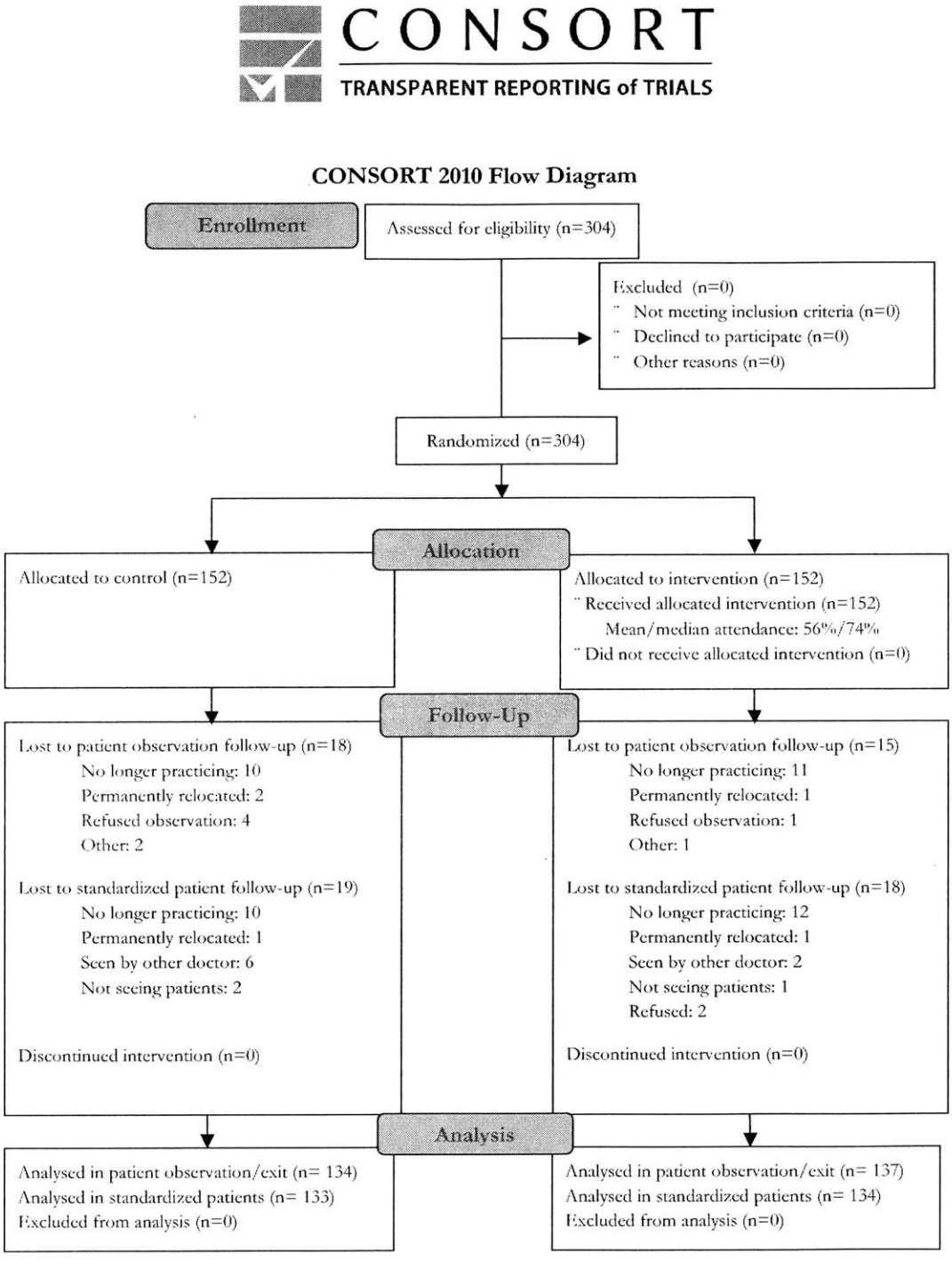
At the same time, the lack of any effect of training on the use of unnecessary medicines, injections, or antibiotics is worrying. Given that unnecessary or harmful practices are higher in the public sector, it seems possible that such polypharmacy is driven by perceived (or active) patient demand for such drugs. Alternately, beliefs about the efficacy of antibiotics may have been too rigid to change through this training. This remains an area for active future research.

In conclusion, both the evidence of a potential beneficial effect of IP training as well as that of persistence

in certain provider practices should be important inputs into the formulation of strategies regarding IP regulation and integration in developing countries.

Figures

Figure 3-1: Consort Flow Diagram



Tables

Table 3.1: Provider characteristics in baseline, by allocation to treatment and control

	(1) Full Sample N = 304	(2) Control N = 152	(3) Treatment N = 152
Mean Age	40.19	39.29	41.10
	(39.16, 41.22)	(37.87, 40.71)	(39.6, 42.6)
Mean Years of Experience	13.14	12.97	13.32
	(12.21, 14.08)	(11.63, 14.31)	(12, 14.63)
Mean Consultation Fee (INR)	10.55	10.78	10.33
	(9.41, 11.7)	(9.07, 12.49)	(8.78, 11.88)
Mean Consultation Fee (USD)	0.17	0.17	0.17
	(0.15, 0.19)	(0.15, 0.2)	(0.14, 0.19)
Mean Village Literacy Rate	0.53	0.53	0.52
	(0.52, 0.53)	(0.52, 0.54)	(0.51, 0.53)
Proportion Male	0.95	0.95	0.96
	(0.92, 0.97)	(0.9, 0.97)	(0.92, 0.98)
Proportion Completed High School	0.62	0.61	0.63
	(0.56, 0.67)	(0.53, 0.68)	(0.55, 0.7)
Proportion with No Formal Training	0.75	0.74	0.76
	(0.7, 0.8)	(0.67, 0.81)	(0.68, 0.82)
Proportion with Minimal Formal Training	0.25	0.26	0.24
	(0.2, 0.3)	(0.19, 0.33)	(0.18, 0.32)

Notes: Fees in 2014 Indian rupees (INR) and 2014 U.S. dollars at ($1 = INR62.5). Purchasing-power-parity adjustment would result in an exchange rate of INR16.76 per dollar. "No Formal Training" includes providers with no formal medical training, although they may have degrees in community medicine. Literacy rates are from the 2011 Indian population census. 95% confidence intervals are presented below the means, and Wilson intervals without continuity correction are used for dichotomous variables.

Table 3.2: Impact of training on main standardized patient outcomes

	Control Group		Intervention Group		(1)	(2)	(3)	(4)
	N	Mean	N	Mean	ITT Logit OR	ITT Logit ME	ITT Linear ME	IV Linear ME
PANEL A: CONTINUOUS OUTCOME VARIABLES FOR DESIRABLE AND NON-DESIRABLE OUTCOMES								
Mean History and Exam Checklist %	396	0.27	388	0.31			0.041***	0.069***
		(0.259, 0.287)		(0.299, 0.327)			(0.017, 0.065)	(0.031, 0.107)
Mean Consultation Length (min)	396	3.25	388	3.49			0.247	0.416
		(3.078, 3.426)		(3.312, 3.677)			(-0.046, 0.54)	(-0.064, 0.896)
Mean Number of Medicines	396	2.16	388	2.21			0.065	0.109
		(2.032, 2.291)		(2.07, 2.346)			(-0.162, 0.291)	(-0.269, 0.487)
PANEL B: DICHOTOMOUS OUTCOME VARIABLES FOR DESIRABLE OUTCOMES								
Correct Case Management	396	0.52	388	0.59	1.402**	0.083**	0.079**	0.133**
		(0.471, 0.569)		(0.545, 0.641)	(1.017, 1.931)	(0.004, 0.162)	(0.004, 0.155)	(0.009, 0.258)
Average Quality Case Management or Better	394	0.11	384	0.17	1.679***	0.060***	0.062***	0.104***
		(0.086, 0.149)		(0.14, 0.215)	(1.145, 2.462)	(0.017, 0.104)	(0.016, 0.107)	(0.03, 0.178)
Referred Case (Asthma/Angina)	263	0.29	258	0.34	1.306	0.057	0.055	0.093
		(0.234, 0.343)		(0.289, 0.403)	(0.881, 1.937)	(-0.026, 0.14)	(-0.026, 0.137)	(-0.042, 0.228)
PANEL C: DICHOTOMOUS OUTCOME VARIABLES FOR NON-DESIRABLE OUTCOMES								
Lowest Quality Case Management	394	0.14	384	0.10	0.664	-0.042	-0.043	-0.072
		(0.111, 0.18)		(0.076, 0.137)	(0.411, 1.073)	(-0.091, 0.007)	(-0.094, 0.008)	(-0.156, 0.012)
Antibiotics (All)	396	0.48	388	0.48	1.059	0.014	0.011	0.018
		(0.429, 0.526)		(0.431, 0.529)	(0.752, 1.49)	(-0.071, 0.099)	(-0.055, 0.077)	(-0.092, 0.129)
Antibiotics (Asthma/Angina)	263	0.33	258	0.33	1.047	0.010	0.009	0.016
		(0.277, 0.39)		(0.278, 0.391)	(0.718, 1.526)	(-0.071, 0.091)	(-0.067, 0.086)	(-0.112, 0.143)
Offered Injection (Asthma/Angina)	263	0.01	258	0.02	1.612	0.004	0.007	0.012
		(0.004, 0.033)		(0.008, 0.044)	(0.337, 7.717)	(-0.012, 0.021)	(-0.017, 0.031)	(-0.028, 0.051)
Any Unnecessary or Harmful Medicine	396	0.71	388	0.70	0.978	-0.004	-0.004	-0.007
		(0.66, 0.75)		(0.648, 0.739)	(0.673, 1.421)	(-0.077, 0.068)	(-0.07, 0.062)	(-0.116, 0.102)

Notes: Mean history and exam checklist completion measures the number of items that were completed from the structured questionnaire. Correct case management is defined as: an inhaler, corticosteroid, or referral for asthma; asking to see the child or recommending ORS for child diarrhea; and referral or aspirin for angina. Unnecessary or harmful medicines are any other class of medication. All regressions control for case presentation fixed effects and the age of the attending provider, and standard errors are clustered at the level of the village. 95% confidence intervals are presented below the means and estimated coefficients. Estimates that are significant at the 95% level of confidence are marked (**) and at the 99% level of confidence (***). Wilson intervals without continuity correction are used for dichotomous variables.

Table 3.3: Impact of training on main clinical observation outcomes

	Control Group		Intervention Group		(1)	(2)	(3)	(4)
	N	Mean	N	Mean	ITT Logit OR	ITT Logit ME	Linear ME	IV Linear ME
PANEL A: CONTINUOUS OUTCOME VARIABLES FOR DESIRABLE AND NON-DESIRABLE OUTCOMES								
Mean Consultation Length (min)	1096	5.38	1215	5.49			0.060	0.095
		(5.192, 5.567)		(5.332, 5.657)			(-0.443, 0.563)	(-0.692, 0.882)
Mean Number of Questions Asked	1096	4.70	1215	5.42			0.648**	1.026**
		(4.527, 4.864)		(5.229, 5.606)			(0.058, 1.238)	(0.118, 1.934)
Mean Number of Examinations	1096	1.87	1215	2.29			0.415***	0.656***
		(1.781, 1.951)		(2.207, 2.375)			(0.154, 0.676)	(0.261, 1.051)
Mean Number of Medicines	1096	2.94	1215	2.82			-0.114	-0.181
		(2.869, 3.016)		(2.746, 2.889)			(-0.325, 0.096)	(-0.507, 0.145)
PANEL B: DICHOTOMOUS OUTCOME VARIABLES FOR NON-DESIRABLE OUTCOMES								
Antibiotics Given	1096	0.48	1215	0.49	1.066	0.016	0.016	0.025
		(0.455, 0.514)		(0.464, 0.52)	(0.802, 1.416)	(-0.055, 0.087)	(-0.055, 0.087)	(-0.087, 0.137)
Injection Given	1096	0.09	1215	0.08	0.826	-0.015	-0.015	-0.024
		(0.078, 0.113)		(0.063, 0.093)	(0.55, 1.242)	(-0.047, 0.017)	(-0.047, 0.018)	(-0.074, 0.027)
IV Given	1096	0.00	1215	0.00	0.658	-0.001	-0.001	-0.001
		(0.001, 0.008)		(0, 0.006)	(0.048, 9.046)	(-0.007, 0.005)	(-0.007, 0.005)	(-0.011, 0.008)

Notes: All regressions control for the age of the attending provider, and standard errors are clustered at the level of the provider. 95% confidence intervals are presented below the means and estimated coefficients. Estimates that are significant at the 95% level of confidence are marked (**) and at the 99% level of confidence (***). Wilson intervals without continuity correction are used for dichotomous variables.

Table 3.4: Clinical practice of control, treatment, and public providers assessed using standardized patients

	(A)	(B)	(C)	(1)	(2)	(3)	(4)
	Means			Control Group - PHC		Treatment Group - PHC	
	PHC (N=33)	Control (N=396)	Treatment (N=394)	Linear Regression, ME	Logistic Regression, OR	Linear Regression, ME	Logistic Regression, OR
PANEL A: CONTINUOUS OUTCOME VARIABLES FOR DESIRABLE AND NON-DESIRABLE OUTCOMES							
Mean History and Exam Checklist %	0.202	0.273	0.313	0.071**		0.111***	
	(0.158, 0.246)	(0.259, 0.287)	(0.299, 0.327)	(0.017, 0.125)		(0.055, 0.166)	
Mean Consultation Length (min)	1.735	3.252	3.495	1.519***		1.762***	
	(1.385, 2.084)	(3.078, 3.426)	(3.312, 3.677)	(1.077, 1.961)		(1.3, 2.223)	
Mean Number of Medicines	2.758	2.162	2.208	-0.595**		-0.548	
	(2.272, 3.243)	(2.032, 2.291)	(2.07, 2.346)	(-1.158, -0.031)		(-1.123, 0.026)	
PANEL B: DICHOTOMOUS OUTCOME VARIABLES FOR DESIRABLE OUTCOMES							
Correct Case Management	0.667	0.520	0.594	-0.147	0.519	-0.073	0.724
	(0.496, 0.802)	(0.471, 0.569)	(0.545, 0.641)	(-0.304, 0.01)	(0.249, 1.082)	(-0.23, 0.085)	(0.353, 1.482)
Average Quality Case Management or Better	0.182	0.114	0.174	-0.068	0.580	-0.007	0.954
	(0.086, 0.344)	(0.086, 0.149)	(0.14, 0.215)	(-0.177, 0.042)	(0.267, 1.26)	(-0.118, 0.104)	(0.443, 2.054)
Referred Case (Asthma/Angina)	0.182	0.285	0.344	0.104	1.843	0.162**	2.435*
	(0.073, 0.385)	(0.234, 0.343)	(0.289, 0.403)	(-0.05, 0.258)	(0.665, 5.106)	(0.003, 0.32)	(0.87, 6.816)
PANEL C: DICHOTOMOUS OUTCOME VARIABLES FOR NON-DESIRABLE OUTCOMES							
Lowest Quality Case Management	0.000	0.142	0.103	0.142***	–	0.102***	–
	(0, 0.104)	(0.111, 0.18)	(0.076, 0.137)	(0.103, 0.181)	–	(0.071, 0.134)	–
Antibiotics (All)	0.667	0.477	0.480	-0.191**	0.371**	-0.188**	0.372**
	(0.496, 0.802)	(0.429, 0.526)	(0.431, 0.529)	(-0.348, -0.034)	(0.156, 0.885)	(-0.346, -0.029)	(0.156, 0.885)
Antibiotics (Asthma/Angina)	0.636	0.331	0.332	-0.306***	0.244***	-0.304***	0.242***
	(0.43, 0.803)	(0.277, 0.39)	(0.278, 0.391)	(-0.502, -0.11)	(0.095, 0.63)	(-0.503, -0.106)	(0.094, 0.621)
Offered Injection (Asthma/Angina)	0.045	0.011	0.019	-0.034	0.233	-0.026	0.403
	(0.008, 0.218)	(0.004, 0.033)	(0.008, 0.044)	(-0.123, 0.054)	(0.02, 2.674)	(-0.116, 0.063)	(0.04, 4.06)
Any Unnecessary or Harmful Medicine	0.879	0.707	0.695	-0.171***	0.268**	-0.183***	0.276**
	(0.727, 0.952)	(0.66, 0.75)	(0.648, 0.739)	(-0.278, -0.063)	(0.095, 0.757)	(-0.289, -0.077)	(0.103, 0.738)

Notes: Mean history and exam checklist completion measures the number of items that were completed from the structured questionnaire. Correct case management is defined as: an inhaler, corticosteroid, or referral for asthma; asking to see the child or recommending ORS for child diarrhea; and referral or aspirin for angina. Unnecessary or harmful medicines are any other class of medication. All regressions control for case presentation fixed effects, and standard errors are clustered at the level of the village. 95% confidence intervals are presented below the means and estimated coefficients. Estimates that are significant at the 95% level of confidence are marked (**) and at the 99% level of confidence (***). Odds ratios could not be computed for Lowest Quality Case Management as no interactions with PHCs fell into that category. Wilson intervals without continuity correction are used for dichotomous variables.

Bibliography

[1] Abramitzky, R., Delavande, A., and Vasconcelos, L. (2009). Marrying up: the role of sex ratio in assortative matching. Available at SSRN 1233882.

[2] Abuya, T., Fegan, G., Rowa, Y., Karisa, B., Ochola, S., Mutemi, W., Marsh, V. (2009) Impact of ministry of health interventions on private medicine retailer knowledge and practices on anti-malarial treatment in Kenya. The American Journal of Tropical Medicine and Hygiene 80(6): 905-913.

[3] Adu-Sarkodie, Y., Steiner, M., Attafuah, J., Tweedy, K. (2000) Syndromic management of urethral discharge in Ghanaian pharmacies. Sexually Transmitted Infections 76(6): 439-442.

[4] Ahsan, H., Y. Chen, F. Parvez, M. Argos, I. Hussain, H. Momotaj, D. Levy, A. VanGeen, G. Howe, and J. Graziano (2006). Health effects of arsenic longitudinal study (heals): Description of a multidisciplinary epidemiologic investigation. Journal of Exposure Science and Environmental Epidemiology 16, 191.

[5] Akbulut-Yuksel, M. and Rosenblum, D. (2012). The Indian Ultrasound Paradox. Working Paper.

[6] Almond, D., Hongbin, L., and Meng, L. 2010. Son Preference and Early Childhood Investments in China. Working Paper.

[7] Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Journal of Political Economy, 113(1), 151-184.

[8] Angrist, J. (2002). How do sex ratios affect marriage and labor markets? Evidence from America's second generation. The Quarterly Journal of Economics, August 2002. 997-1038.

[9] Anukriti, S. (2013). Sex Ratios, Schooling, and the Marriage Market. Working Paper.

[10] Argos, M., T. Kaira, P. Rathouz, Y. Chen, B. Pierce, F. Parvez, T. Islam, A. Ahmed, M. Rakibuz-Zaman, R. Hasan, G. Sarwar, V. Slavkovich, A. van Geen, J. Graziano, and H. Ahsan (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): A prospective cohort study. The Lancet 376, 267.

[11] Banerjee, A., Deaton, A., Duflo, E. (2004) Health care delivery in rural Rajasthan. Economic and Political Weekly: 944-949.

[12] Baqui, A., R. Black, A. Mitra, H. Chowdhury, K. Zaman, V. Fauveau, and R. Sack (1994). Matlab: Women, Children and Health. Dhaka: ICDDR,B.

[13] Bellows, J. and Miguel, E. (2008). War and local collective action in Sierra Leone. Journal of Public Economics, 93(11), 1144-1157.

[14] Bhalotra, S. and Cochrane, T. (2010). Where Have All the Young Girls Gone? Identification of Sex Selection in India. IZA Discussion Paper No. 5381.

[15] Bloom, G., Standing, H., Lucas, H., Bhuiya, A., Oladepo, O., Peters, D. H. (2011) Making health markets work better for poor people: the case of informal providers. Health Policy and Planning 26: i45-i52.

[16] BMOH (2004). National policy for arsenic mitigation 2004. Bangladesh Ministry of Health Compendium of Environment Statistics of Bangladesh 2004, 388.

[17] Caldwell, B., J. Caldwell, S. Mitra, and W. Smith (2003). Tubewells and arsenic in Bangladesh: Challenging a public health success story. International Journal of Population Geography 9, 23.

[18] Caldwell, B., W. Smith, K. Lokuge, G. Ranmuthugala, K. Dear, A. Milton, M. Sim, J. Ng, and S. Mitra (2006). Access to drinking water and arsenicosis in Bangladesh. Journal of Health Population and Nutrition 24(3), 336.

[19] Chakraborty, S., D'Souza, S. A., Northrup, R. S. (2000) Improving private practitioner care of sick children: testing new approaches in rural Bihar. Health Policy and Planning 15(4): 400-407.

[20] Chalker, J., Chuc, N., Falkenberg, T., Tomson, G. (2002) Private pharmacies in Hanoi, Vietnam: a randomized trial of a 2-year multi-component intervention on knowledge and stated practice regarding ARI, STD and antibiotic/steroid requests. Tropical Medicine International Health 7(9): 803-810.

[21] Chen, L., Evans, T., Anand, S., Boufford, J. I., Brown, H., Chowdhury, M., Cueto, M., Dare, L., Dussault, G., Elzinga, G., others (2004) Human resources for health: overcoming the crisis. The Lancet 364: 1984-1990.

[22] Chen, Y., J. Graziano, F. Parvez, I. Hussain, H. Momotaj, A. van Geen, G. Howe, and H. Ahsan (2006). Modification of risk of arsenic-induced skin lesions by sunlight exposure, smoking, and occupational exposures in Bangladesh. Epidemiology 17(4), 459.

[23] Chiappori, P. A., Fortin, B., and Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. Journal of Political Economy, 110(1), 37-72.

[24] Das, J., Holla, A., Mohpal, A., Muralidharan, K. (2015) Quality and Accountability in Healthcare Delivery: Audit Evidence from Primary Care Providers in India. World Bank, Washington D.C.

[25] Das, J., Kwan, A., Daniels, B., Satyanarayana, S., Subbaraman, R., Bergkvist, S., Das, R., Das, V., Pai, M. (Forthcoming) First use and validation of the standardized patient methodology to assess quality of tuberculosis care: a pilot, cross-sectional study. Lancet Infectious Diseases.

[26] De Costa, A. Diwan, V. (2007) Where is the public health sector? Public and private sector healthcare provision in Madhya Pradesh, India. Health Policy 84: 269-276.

[27] Di Renzo, G. C., Rosati, A., Sarti, R. D., Cruciani, L., and Cutuli, A. M. (2007). Does fetal sex affect pregnancy outcome? Gender medicine, 4(1), 19-30.

[28] Dunn, G., Maracy, M., Tomenson, B. (2005) Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: the role of instrumental variable methods. Statistical Methods in Medical Research 14(4): 369-395.

[29] Edlund, L. 1999. Son Preference, Sex Ratios, and Marriage Patterns. Journal of Political Economy,107(6).

[30] Edlund, L. and Lee, C. (2009). Son preference, sex selection and economic development: theory and evidence from South Korea. Working Paper.

[31] Esrey, A. (1996). Water, waste, and well-being: A multicountry study. American Journal of Epidemiology 143(6), 608.

[32] Esrey, S. and J. Habicht (1986). Epidemiologic evidence for health benefits from improved water and sanitation in developing countries. Epidemiologic Reviews 8, 117.

[33] Feroze Ahmed, M. (2002). Alternative water supply options for arsenic affected areas of Bangladesh. Conference Paper, ITN-Bangladesh, Centre for Water Supply and Waste Management .

[34] Gautham, M., Shyamprasad, K., Singh, R., Zachariah, A., Singh, R., Bloom, G. (2014) Informal rural healthcare providers in North and South India. Health Policy and Planning 29: i20-i29.

[35] George, A. Iyer, A. (2013) Unfree markets: socially embedded informal health providers in northern Karnataka, India. Social Science Medicine 96: 297-304.

[36] Government of India. (2005) Report of the National Commission on Macroeconomics and Health. Report of the National Commission on Macroeconomics and Health.

[37] Hoque, B., S. Huttly, K. Aziz, M. Patwary, and R. Feachem (1989). Tubewell water consumption and its determinants in a rural area of Bangladesh. Journal of Tropical Medicine and Hygiene 92(3), 197.

[38] Hu, L. and Schlosser, A. (2011). Prenatal Sex Selection and Girls Well-Being: Evidence from India. IZA Discussion Paper No. 5562.

[39] Hvistendahl, M. (2011). Unnatural Selection: Choosing Boys Over Girls, and the Consequences of a World Full of Men. Public Affairs.

[40] Inoue, A., and Solon, G. (2010). Two-sample instrumental variables estimators. The Review of Economics and Statistics, 92(3), 557-561.

[41] Integrated Management of Neonatal and childhood illness: Physician Chart Booklet. (http://mohfw.nic.in/nrhm /stg/Contents.htm).

[42] Ischaemic Heart Disease: Acute Myocardial Infarction. IMNCI Physician Chart Booklet.

[43] Jacoby, J. (2010). 100 million 'missing' girls. The Boston Globe. March 14, 2010.

[44] Jakariya, M. (2007). Arsenic in tubewell water of Bangladesh and approaches for sustainable mitigation. Doctoral Thesis, KTH-International Groundwater Arsenic Research Group.

[45] Jayachandran, S. and Kuziemko, I. (2011). Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. The Quarterly Journal of Economics, 126(3), 1485-1538.

[46] Jindal, S. K., Gupta, D., Aggarwal, A. N., Agarwal, R. (2005). Guidelines for management of asthma at primary and secondary levels of health care in India. The Indian Journal of Chest Diseases Allied Sciences, 47(4): 309.

[47] Johnston, R. (2006). Arsenic in Bangladesh. UNICEF WES Section Briefing.

[48] King, G., Gakidou, E., Imai, K., Lakin, J., Moore, R. T., Nall, C., Ravishankar, N., Vargas, M., Téllez-Rojo, M. M., Ávila, J. E. H., others (2009) Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. The Lancet 373(9673): 1447-1454.

[49] Kozul, C., H. Ely, R. Enelow, and J. Hamilton (2009). Low-dose arsenic comprises the immune response to influenza A infection in vivo. Environmental Health Perspectives 117 (9), 441.

[50] Kremer, M., J. Leino, E. Miguel, and A. Zwane (2010). Spring cleaning: Rural water impacts, valuation and property rights institutions. Quarterly Journal of Economics.

[51] Lafortune, J. (2010). Making yourself attractive: Pre-marital investments and the returns to education in the marriage market. Unpublished mimeo available at http://sites. google. com/site/jeannelafortune/papers/Draft-01-2010. pdf.

[52] Leonard, K. L. Masatu, M. C. (2010) Using the Hawthorne effect to examine the gap between a doctor's best possible practice and actual performance. Journal of Development Economics 93(2): 226-234.

[53] Liaw, J., G. Marshall, Y. Yuan, C. Ferreccio, C. Steinmaus, and A. Smith (2008). Increased childhood liver cancer mortality and arsenic in drinking water in northern Chile. Cancer Epidemiology, Biomarkers and Prevention 17(8), 1982.

[54] Lindskog, U., P. Lindskog, and S. Wall (1987). Water supply, sanitation and health education programmes in developing countries: Problems of evaluation. Scandinavian Journal of Social Medicine 15, 123.

[55] Lokuge, K., W. Smith, B. Caldwell, K. Dear, and A. Milton (2004). The effect of arsenic mitigation interventions on disease burden in Bangladesh. Environmental Health Perspectives 112(11), 1172.

[56] Madajewicz, M., A. Pfaff, A. van Geen, J. Graziano, I. Hussein, H. Momotaj, R. Sylvi, and H. Ahsan (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. Journal of Development Economics 84 (2), 731.

[57] MAQARI Team (2011) Mapping Medical Providers in Rural India: Four Key Trends. Center for Policy Research Policy Briefs. New Delhi: Centre for Policy Research.

[58] Mazumdar, D., J. Das Gupta, A. Sandra, A. Pal, A. Ghose, and S. Sarkar (1998). Chronic arsenic toxicity in West Bengal - the worst calamity in the world. Journal of Indian Medical Association 96(1), 4.

[59] McClellan, M., McNeil, B. J., Newhouse, J. P. (1994) Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. JAMA 272(11): 859-866.

[60] Melamed, N., Yogev, Y., and Glezerman, M. (2009). Effect of fetal sex on pregnancy outcome in twin pregnancies. Obstetrics and Gynecology, 114(5), 1085-1092.

[61] Meng, L., (2008). Prenatal Sex Selection and Missing Girls in China: Evidence from the Diffusion of Diagnostic Ultrasound. Working Paper.

[62] Milton, A., W. Smith, B. Rahman, Z. Hasan, U. Kulsum, K. Dear, M. Rakibuddin, and A. Ali (2005). Chronic arsenic exposure and adverse pregnancy outcomes in Bangladesh. Epidemiology 16(1), 82.

[63] Mohanan, M., Vera-Hernández, M., Das, V., Giardili, S., Goldhaber-Fiebert, J. D., Rabin, T. L., Raj, S. S., Schwartz, J. I., Seth, A. (2015) The Know-Do Gap in Quality of Health Care for Childhood Diarrhea and Pneumonia in Rural India. JAMA Pediatrics 169(4): 349-357.

[64] Morris, S., R. Black, and L. Tomaskovic (2003). Predicting the distribution of under-five deaths by cause in countries without adequate vital registration systems. International Journal of Epidemiology 32, 1041.

[65] Peltzer, K., Mngqundaniso, N., Petros, G. (2006) A controlled study of an HIV/AIDS/STI/TB intervention with traditional healers in KwaZulu-Natal, South Africa. AIDS and Behavior 10(6), 683-690.

[66] Pop-Eleches, C. (2006). The impact of an abortion ban on socioeconomic outcomes of children: evidence from Romania. Journal of Political Economy, 114(4), 744-773.

[67] Pruss, A., D. Kay, L. Fewtrell, and J. Bartram (2002). Estimating the burden of disease from water, sanitation, and hygiene at a global level. Environmental Health Perspectives, 537.

[68] Research Council, N. (2001). Arsenic in drinking water: 2001 update. Washington, DC: National Academic Press.

[69] Rossman, T., A. Uddin, F. Burns, and M. Bosland (2002). Arsenite cocarcinogenesis: an animal model derived from genetic toxicology studies. Environmental Health Perspectives 110(Suppl 5), 749.

[70] Rowe, A. K., Onikpo, F., Lama, M., Deming, M. S. (2012) Evaluating health worker performance in Benin using the simulated client method with real children. Implementation Science 7(1): 95.

[71] Schoenfeld, A. (2005). Area, village, and household response to arsenic testing and labeling of tubewells in Araihazar, Bangladesh. Doctoral Thesis, Columbia University (80).

[72] Shah, N. M., Brieger, W. R., Peters, D. H. (2010) Can interventions improve health services from informal private providers in low and middle-income countries? A comprehensive review of the literature. Health Policy and Planning 26: 275-287.

[73] Stenson, B., Syhakhang, L., Lundborg, C. S., Eriksson, B. (2001) Private pharmacy practice and regulation: A Randomized Trial in Lao P.D.R. International Journal of Technology Assessment in Health Care 17: 579-589.

[74] Sudhinaraset, M., Ingram, M., Lofthouse, H. K., Montagu, D. (2013) What is the role of informal healthcare providers in developing countries? A systematic review. PloS one 8(2): e54978.

[75] Suzuki, S., L. Arnold, T. Ohnishi, and S. Cohen (2008). Effects of inorganic arsenic on the rat and mouse urinary bladder. Toxicological Sciences 106 (2), 350.

[76] Sylvia, S., Shi, Y., Xue, H., Tian, X., Wang, H., Liu, Q., Medina, A., Rozelle, S. (2014) Survey using incognito standardized patients shows poor quality care in China's rural clinics. Health Policy and Planning 30(3): 322-333.

[77] Tofail, F., M. Vahter, J. Hamadani, B. Nermell, S. Huda, M. Yunus, M. Rahman, and S. Grantham-McGregor (2009). Effect of arsenic exposure during pregnancy on infant development at 7 months in rural Matlab, Bangladesh. Environmental Health Perspectives 117(2), 288.

[78] Trivedi, A. and Timmons, H. (2013). India's Man Problem. New York Times: India Ink.

[79] Tucker, S., F. Li, R. Wilson, D. Byrd, S. Lai, Y. Tong, and L. Loo (2001). Relationship between consumption of arsenic-contaminated well water and skin disorders in Huhhot, Inner Mongolia. Final Report, University of Texas, Department of Dermatology.

[80] UNICEF (2008). Arsenic mitigation in Bangladesh. UNICEF Brief.

[81] Vachter, M. (2008). Health effects of early life exposure to arsenic. Basic and clinical pharmacology and toxicology 102, 204.

[82] VanGeen, A., H. Ahsan, A. Horneman, R. Dhar, Y. Zheng, I. Hussain, K. Ahmed, A. Gelman, M. Stute, H. Simpson, S. Wallace, C. Small, F. Parvez, V. Slavkovich, N. Lolacono, M. Becker, Z. Cheng, H. Momotaj, M. Shahnewaz, A. Seddique, and J. Graziano (2002). Promotion of well-switching to mitigate the current arsenic crisis in Bangladesh. Bulletin of the World Health Organization (80), 732.

[83] VanGeen, A., Y. Zheng, R. Versteeg, M. Stute, A. Horneman, R. Dhar, M. Steckler, A. Gel- man, C. Small, H. Ahsan, J. Graziano, I. Hussain, and K. Ahmed (2003). Spatial variability of arsenic in 6000 tube wells in a 25 km sq area of Bangladesh. Water Resources Research 39(5), 1.

[84] Varghese, J. (2010) The new rural doctor: qualified quack or appropriate healthcare provider? Indian Journal of Medical Ethics 7(2).

[85] Wasserman, G., V. Liu, F. Parvez, H. Ahsan, P. Factor-Litvak, A. vanGeen, V. Slavkovich, N. Lolacono, Z. Cheng, I. Hussain, H. Momotaj, and J. Graziano (2004). Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. Environmental Health Perspectives 112(13), 1329.

[86] Wei, S., and Zhang X. 2009. The Competitive Saving Motive: Evidence From Rising Sex Ratios and Savings Rates in China. NBER Working Paper 15093.

[87] WHO (2008). An interview with Mahmuder Rahman: Bangladesh's arsenic agony. Bulletin of the World Health Organization 86 (1), 11.

[88] WorldBank (2007). Implementation completion and results report on a credit in the amount of SDR 24.2 million to Bangladesh for arsenic mitigation water supply. Environment and Water Resources Unit, South Asia Region.

[89] Wright, J., S. Gundryl, and R. Conroy (2004). Household drinking water in developing countries: A systematic review of microbiological contamination between source and point- of-use. Tropical Medicine and International Health 9 (1), 106.

[90] Yu, W., C. Harvey, and C. Harvey (2003). Arsenic in groundwater in Bangladesh: A geostatistical and epidemiological framework for evaluating health effects and potential remedies. Water Resources Research 39, 1146.

[91] (2010). Gendercide: The worldwide war on baby girls. The Economist.

[92] (2010). Haryana's lonely bachelors. The Economist.

[93] Govt Plans to Train Quacks. The Telegraph. 2 Aug 2014.

Appendices for Chapter 1

Appendix Figures

Figure A1.1: Payoff Matrix for Parental Sex Selection

Payoff Matrix					
Poor/Rich	US_m	US_f	$noSS$	PS_m	PS_f
US_m	$U(m_P) - CR$ $U(m_R) - C$	$U(m_P) - CR$ $U(f_R) - C$	$U(m_P) - CR$ $0.5[U(m_{RR}) + U(f_{RR})]$	$U(m_P) - CR$ $U(m_R) - 0.5E$	$U(m_P) - CR$ $U(f_R) - 0.5E$
US_f	$U(f_{PR}) - CR$ $U(m_{RP}) - C$	$U(f_P) - CR$ $U(f_R) - C$	$U(f_P) - CR$ $0.5[U(m_{RR}) + U(f_{RR})]$	$U(f_{PR}) - CR$ $U(m_{RP}) - 0.5E$	$U(f_P) - CR$ $U(f_R) - 0.5E$
$noSS$	$0.5[U(m_P) + U(f_{PR})]$ $0.5[U(m_{RP}) + U(m_R)] - C$	$0.5[U(m_{PP}) + U(f_{PP})]$ $U(f_R) - C$	$0.5[U(m_{PP}) + U(f_{PP})]$ $0.5[U(m_{RR}) + U(f_{RR})]$	$0.5[U(m_P) + U(f_{PR})]$ $0.5[U(m_{RP}) + U(m_R)] - 0.5E$	$0.5[U(m_{PP}) + U(f_{PP})]$ $U(f_R) - 0.5E$
PS_m	$U(m_P) - 0.5E$ $U(m_R) - C$	$U(m_P) - 0.5E$ $U(f_R) - C$	$U(m_P) - 0.5E$ $0.5[U(m_{RR}) + U(f_{RR})]$	$U(m_P) - 0.5E$ $U(m_R) - 0.5E$	$U(m_P) - 0.5E$ $U(f_R) - 0.5E$
PS_f	$U(f_{PR}) - 0.5E$ $U(m_{RP}) - C$	$U(f_P) - 0.5E$ $U(f_R) - C$	$U(f_P) - 0.5E$ $0.5[U(m_{RR}) + U(f_{RR})]$	$U(f_{PR}) - 0.5E$ $U(m_{RP}) - 0.5E$	$U(f_P) - 0.5E$ $U(f_R) - 0.5E$

Figure A1.2: Birth order and son preference

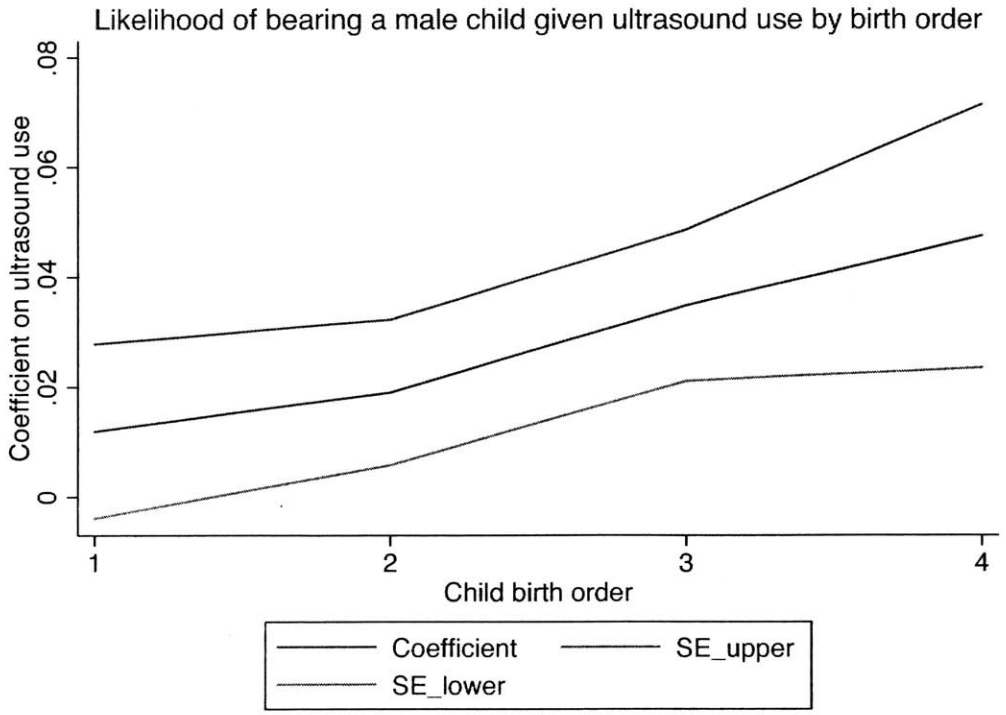
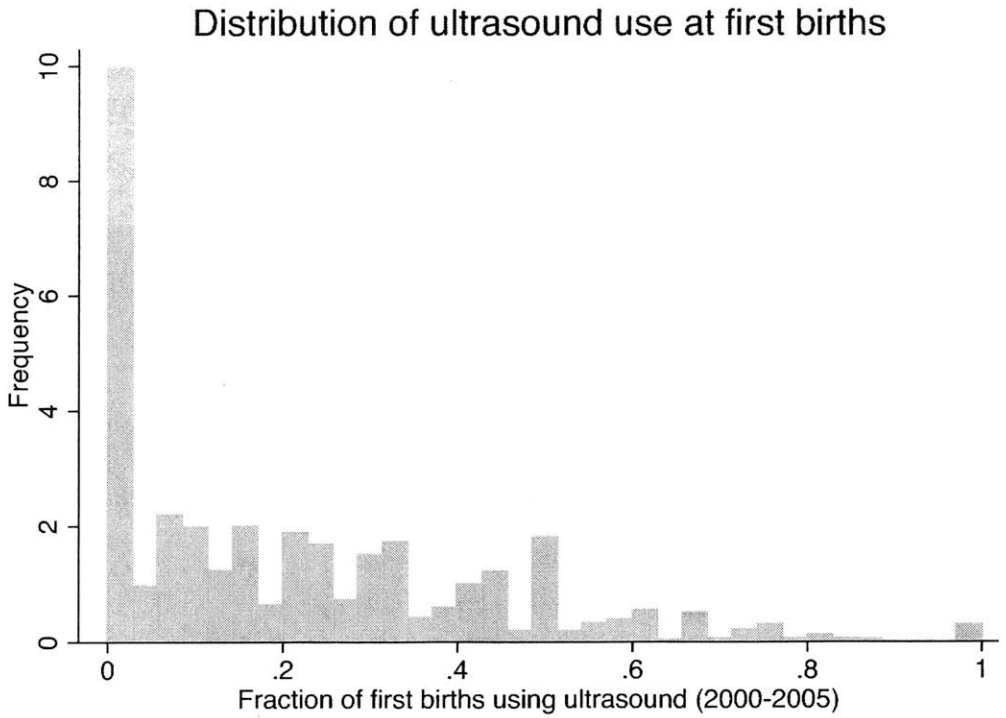


Figure A1.3: Distribution of ultrasound



Appendix Tables

Table A1.1: Population differences across wealth

VARIABLES	(1) Sex ratio in poor strata	(2) Sex ratio in wealthy strata	(3) Gap in sex ratios: wealthy - poor strata
Ultrasound	-1.052***	0.0988	0.245
	[0.384]	[0.248]	[0.896]
Constant	1.559***	-0.343	-2.768***
	[0.248]	[0.265]	[0.663]
Observations	1825	1955	458

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Regressions include state fixed effects and controls for village level averages of wealth, education, and religion, as well as whether the village is rural

Appendices for Chapter 2

Appendix Tables

Table A2.1: Sample means without village fixed effects

VARIABLES	High Conc.	N	Low Conc.	T-statistic
Panel I: Exogenous Variables				
Age of mother	38.48	1928/889	38.23	0.79
Age of mother at earliest birth	20.87	1928/889	21.47	-2.59
Education of mother	3.59	1928/889	3.11	3.73
Education of father	3.97	1928/889	3.40	0.001
Mean birth interval	2.30	1928/889	2.24	1.31
Solvency	0.65	1928/889	0.69	-1.91
Land size	0.85	1928/889	0.85	0.08
Number of rooms in house	2.79	1928/889	2.67	2.7
Electricity	0.44	1928/889	0.30	7
Muslim	0.96	1928/889	0.98	-1.67
Fraction of children living in household	0.81	1654/775	0.82	-0.95
Respondent's age	42.63	1928/889	42.31	0.79
Male respondent	0.16	1928/889	0.16	0.32
Sufficiency of food per week	0.91	1928/889	0.95	-3.2
Outstanding loan	0.56	1928/889	0.51	2.26
Years living in house	26.21	1928/889	24.20	3.67
Years living in village	30.61	1547/729	30.02	0.88
Mean monthly income of household	11.55	1928/889	10.97	1.47
Head of household works in agriculture	0.40	1928/889	0.47	-3.71
Head of household works in business	0.16	1928/889	0.15	1.2
Panel II: Endogenous Variables				
Arsenic concentration (ppm)	96.26	1928/889	28.60	65.34
Whether closest well tested	0.68	1700/836	0.71	-1.78
Whether closest well painted	0.67	1542/772	0.66	0.54
Fraction of deaths under 12 mo.	0.06	1928/889	0.06	1.59
Fraction of deaths under 24 mo.	0.07	1928/889	0.06	1.9
Fraction of deaths under 60 mo.	0.09	1928/889	0.08	1.15
M:F sex ratio	0.43	1928/889	0.46	-3.32
Number of offspring in family	4.18	1928/889	4.14	0.5
Number of drinking sources used	2.05	1928/889	2.06	-0.93
Number of cooking sources used	2.25	1928/889	2.23	1.28
Value of house ($)	2190.39	1835/864	2011.29	2.24

Notes:

(1) All averages are calculated controlling for village means.

(2) Sufficiency of food defined as family members taking at least two meals a day last week; solvency defined as last week's expenses being within the budget.

(3) In columns 1-4 ("measured contamination"), high concentration versus low concentration defined according to field test of shallow tubewell closest to residence. High concentration households are those with tubewells that contain arsenic concentrations greater than 60ppb. In columns 5-8 ("reported concentration"), high concentration households are those who report that their well tested positive for arsenic concetration, or (if household has no recollection of well being tested or test result) if closest

(4) High concentration and low concentration are defined according to measured concetration of arsenic in nearest shallow tubewell, with cutoff point of 60ppb.

(5) Sample means and mean differences do not account of village fixed effects.

Table A2.2: Measured arsenic contamination and early life exposure, household-level estimation

	Fraction of deaths per HH under 12 mo.			Fraction of deaths per HH under 24 mo.			Fraction of deaths per HH under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00695	0.00640	0.00407	0.0117	0.0110	0.00782	0.00831	0.00702	0.00354
	[0.00775]	[0.00774]	[0.00771]	[0.00815]	[0.00798]	[0.00793]	[0.00928]	[0.00873]	[0.00861]
Fraction born after 2000	-0.0476***	-0.0190	-0.0130	-0.0596***	-0.0247	-0.0163	-0.0844***	-0.0269	-0.0188
	[0.0151]	[0.0176]	[0.0176]	[0.0154]	[0.0175]	[0.0175]	[0.0176]	[0.0193]	[0.0192]
High conc. * Fraction born after 2000	0.0297*	0.0276	0.0317*	0.0256	0.0233	0.0286	0.0330	0.0304	0.0359*
	[0.0177]	[0.0183]	[0.0188]	[0.0181]	[0.0187]	[0.0192]	[0.0200]	[0.0208]	[0.0211]
M:F sex ratio		0.0116	0.00713		0.0128	0.00685		0.00947	0.00191
		[0.0109]	[0.0104]		[0.0110]	[0.0104]		[0.0116]	[0.0109]
Number of offspring in family		0.0155***	0.0176***		0.0177***	0.0199***		0.0202***	0.0239***
		[0.00171]	[0.00205]		[0.00184]	[0.00226]		[0.00219]	[0.00258]
Earliest child birth year in family		0.000356	-0.00108		0.000284	-0.00155*		-0.000694	-0.00258**
		[0.000623]	[0.000771]		[0.000688]	[0.000874]		[0.000789]	[0.000997]
Age of mother at earliest birth			0.00185***			0.00197***			0.00293***
			[0.000700]			[0.000727]			[0.000775]
Education of mother			0.00288*			0.00365**			0.00462**
			[0.00171]			[0.00170]			[0.00179]
Education of father			-0.00184*			-0.00201*			-0.00236**
			[0.00108]			[0.00113]			[0.00114]
Mean birth interval			-0.0103***			-0.0132***			-0.0154***
			[0.00273]			[0.00287]			[0.00311]
Solvency			0.000184			0.000536			0.00336
			[0.00531]			[0.00571]			[0.00624]
Land size (acres)			0.00376**			0.00339*			0.00325
			[0.00176]			[0.00176]			[0.00201]
Number of rooms in house			-0.0114***			-0.0121***			-0.0138***
			[0.00280]			[0.00293]			[0.00309]
Electricity			-0.00494			-0.00625			-0.00407
			[0.00562]			[0.00587]			[0.00663]
Muslim			0.00667			0.0130			0.00849
			[0.0112]			[0.0110]			[0.0129]
Monthly income per capita			0.000228			0.000267			0.000441
			[0.000281]			[0.000278]			[0.000321]
Mean among offspring with zero exposure in households with low arsenic	0.076566	0.076566	0.076566	0.083527	0.083527	0.083527	0.12297	0.12297	0.12297
	[0.26621]	[0.26621]	[0.26621]	[0.27700]	[0.27700]	[0.27700]	[0.32878]	[0.32878]	[0.32878]
Observations	2817	2817	2817	2817	2817	2817	2817	2817	2817

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

(4) An observation is a household. Outcome is fraction of live births that died before given age cutoff.

(5) High concentration defined by arsenic test of closest well to household ("measured contamination").

Table A2.3: Measured arsenic contamination and early life exposure, binary-exposure

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00785	0.00606	0.00461	0.0135*	0.0110	0.00905	0.0128	0.00865	0.00661
	[0.00723]	[0.00722]	[0.00712]	[0.00777]	[0.00771]	[0.00756]	[0.00850]	[0.00840]	[0.00821]
Born after 2000	-0.0403***	-0.00116	0.000706	-0.0530***	-0.00814	-0.00578	-0.0692***	-0.00460	-0.00272
	[0.00879]	[0.0127]	[0.0127]	[0.00931]	[0.0135]	[0.0135]	[0.0106]	[0.0152]	[0.0152]
High concentration * Born after 2000	0.0142	0.0163	0.0203*	0.0158	0.0184	0.0233*	0.0162	0.0200	0.0255*
	[0.0117]	[0.0117]	[0.0117]	[0.0124]	[0.0124]	[0.0124]	[0.0138]	[0.0136]	[0.0137]
Sex		0.0292***	0.0279***		0.0287***	0.0270***		0.0269***	0.0243***
		[0.00499]	[0.00499]		[0.00528]	[0.00527]		[0.00588]	[0.00587]
Parity		0.00308*	0.00329		0.00442***	0.00425*		0.00620***	0.00335
		[0.00157]	[0.00216]		[0.00167]	[0.00221]		[0.00176]	[0.00236]
Birth year		-0.00636***	-0.00693***		-0.00842***	-0.00904***		-0.0134***	-0.0145***
		[0.00204]	[0.00207]		[0.00218]	[0.00222]		[0.00243]	[0.00246]
Birth year squared		0.00006	0.00009		0.000103	0.000131		0.000188**	0.000241***
		[0.00007]	[0.00008]		[0.00008]	[0.00008]		[0.00009]	[0.00009]
Age of mother at birth			-0.000243			-0.000192			0.00009
			[0.000540]			[0.000576]			[0.000622]
Mother's education			0.00117			0.00175			0.00253
			[0.00148]			[0.00155]			[0.00164]
Father's education			-0.00146			-0.00164			-0.00212*
			[0.00107]			[0.00114]			[0.00126]
Years since birth of last child			-0.0107***			-0.0135***			-0.0173***
			[0.00178]			[0.00188]			[0.00204]
Solvency (acres)			0.000313			-0.00165			0.00170
			[0.00556]			[0.00599]			[0.00633]
Land size			0.00474***			0.00420**			0.00387**
			[0.00164]			[0.00170]			[0.00193]
No. of rooms in house			-0.0133***			-0.0143***			-0.0152***
			[0.00268]			[0.00298]			[0.00317]
Electricity			-0.00936			-0.0105			-0.0125*
			[0.00615]			[0.00668]			[0.00734]
Muslim			0.0115			0.0211			0.0172
			[0.0145]			[0.0148]			[0.0157]
Monthly income per capita			0.000214			0.000227			0.000313
			[0.000313]			[0.000328]			[0.000356]
Mean among offspring with zero exposure in households with low arsenic concentration	0.07649	0.07649	0.07649	0.09012	0.09012	0.09012	0.11739	0.11739	0.11739
	[0.26582]	[0.26582]	[0.26582]	[0.28641]	[0.28641]	[0.28641]	[0.32194]	[0.32194]	[0.32194]
Observations	11766	11766	11766	11766	11766	11766	11766	11766	11766

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

Table A2.4: Measured arsenic contamination and early life exposure based on when well tested

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00733	0.00541	0.00422	0.0134*	0.0107	0.00905	0.0129	0.00844	0.00672
	[0.00707]	[0.00706]	[0.00697]	[0.00760]	[0.00754]	[0.00740]	[0.00826]	[0.00815]	[0.00797]
Born after well tested	-0.0499***	-0.0200*	-0.0191	-0.0607***	-0.0235*	-0.0222*	-0.0766***	-0.0235	-0.0224
	[0.00914]	[0.0120]	[0.0120]	[0.00986]	[0.0132]	[0.0132]	[0.0110]	[0.0146]	[0.0146]
High concentration * Born	0.0244**	0.0273**	0.0307**	0.0236*	0.0273**	0.0317**	0.0233	0.0289**	0.0339**
after well tested	[0.0124]	[0.0124]	[0.0125]	[0.0132]	[0.0132]	[0.0133]	[0.0143]	[0.0143]	[0.0145]
Sex		0.0292***	0.0279***		0.0287***	0.0270***		0.0269***	0.0243***
		[0.00499]	[0.00499]		[0.00529]	[0.00527]		[0.00589]	[0.00587]
Parity		0.00311**	0.00330		0.00444***	0.00426*		0.00623***	0.00336
		[0.00158]	[0.00216]		[0.00167]	[0.00221]		[0.00176]	[0.00236]
Birth year		-0.00728***	-0.00801***		-0.00912***	-0.00991***		-0.0144***	-0.0157***
		[0.00194]	[0.00197]		[0.00210]	[0.00214]		[0.00234]	[0.00236]
Birth year squared		0.000109	0.000141**		0.000137*	0.000174**		0.000238***	0.000299***
		[0.00007]	[0.00007]		[0.00007]	[0.00007]		[0.00008]	[0.00008]
Age of mother at birth			-0.000254			-0.000201			0.00008
			[0.000542]			[0.000577]			[0.000623]
Mother's education			0.00116			0.00174			0.00252
			[0.00148]			[0.00155]			[0.00165]
Father's education			-0.00142			-0.00161			-0.00208*
			[0.00107]			[0.00114]			[0.00126]
Years since birth of last child			-0.0107***			-0.0135***			-0.0173***
			[0.00178]			[0.00188]			[0.00204]
Solvency			0.000327			-0.00161			0.00173
			[0.00556]			[0.00599]			[0.00633]
Land size (acres)			0.00470***			0.00418**			0.00384**
			[0.00164]			[0.00170]			[0.00192]
No. of rooms in house			-0.0132***			-0.0142***			-0.0151***
			[0.00268]			[0.00298]			[0.00317]
Electricity			-0.00908			-0.0102			-0.0122*
			[0.00615]			[0.00667]			[0.00733]
Muslim			0.0111			0.0207			0.0167
			[0.0145]			[0.0148]			[0.0157]
Monthly income per capita			0.000212			0.000223			0.000309
			[0.000313]			[0.000327]			[0.000356]
Mean among offspring with zero exposure in households with low arsenic concentration	0.075572 [0.26435]	0.075572 [0.26435]	0.075572 [0.26435]	0.088429 [0.28396]	0.088429 [0.28396]	0.088429 [0.28396]	0.11477 [0.31879]	0.11477 [0.31879]	0.11477 [0.31879]
Observations	11766	11766	11766	11766	11766	11766	11766	11766	11766

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

(4) Whether individual born after well tested based on survey reports of year that drinking water tested for arsenice.

Table A2.5: Control test: 100 ppb arsenic contamination cutoff

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00824	0.00700	0.00693	-0.00172	-0.00306	-0.00298	0.00259	0.00105	0.000459
	[0.00905]	[0.00905]	[0.00870]	[0.0102]	[0.0102]	[0.00987]	[0.0122]	[0.0121]	[0.0118]
Early life exposure	-0.0273***	-0.00425	0.00209	-0.0420***	0.0111	0.0122	-0.0684***	0.0306	0.0225
	[0.00897]	[0.0152]	[0.0152]	[0.00861]	[0.0166]	[0.0167]	[0.00932]	[0.0212]	[0.0213]
High concentration*Early life exposure	-0.000970	-0.00178	-0.00143	0.0247	0.0233	0.0259	0.0125	0.0105	0.0147
	[0.0181]	[0.0183]	[0.0181]	[0.0195]	[0.0196]	[0.0194]	[0.0201]	[0.0201]	[0.0199]
Sex		0.0315***	0.0302***		0.0304***	0.0287***		0.0288***	0.0265***
		[0.00627]	[0.00628]		[0.00651]	[0.00653]		[0.00727]	[0.00727]
Parity		0.00224	0.00246		0.00327	0.00291		0.00535**	0.00265
		[0.00197]	[0.00272]		[0.00203]	[0.00275]		[0.00215]	[0.00294]
Birth year		-0.00863***	-0.00853***		-0.00941***	-0.00982***		-0.0132***	-0.0145***
		[0.00252]	[0.00257]		[0.00260]	[0.00267]		[0.00254]	[0.00261]
Birth year squared		0.000175*	0.000179*		0.000127	0.000156		0.000152	0.000236**
		[0.00009]	[0.0001]		[0.000100]	[0.000104]		[9.89e-05]	[0.000103]
Age of mother at birth			0.0001			0.000392			0.000522
			[0.000659]			[0.000709]			[0.000755]
Mother's education			0.000779			0.00150			0.00153
			[0.00171]			[0.00179]			[0.00192]
Father's education			-0.00242**			-0.00203			-0.00194
			[0.00120]			[0.00130]			[0.00144]
Years since birth of last child			-0.0109***			-0.0135***			-0.0167***
			[0.00206]			[0.00216]			[0.00236]
Solvency			-0.00623			-0.00791			-0.000402
			[0.00669]			[0.00715]			[0.00755]
Land size (acres)			0.00476***			0.00361**			0.00350*
			[0.00178]			[0.00178]			[0.00206]
No. of rooms in house			-0.00974***			-0.0106***			-0.0114***
			[0.00318]			[0.00343]			[0.00367]
Electricity			-0.00867			-0.0122			-0.0160*
			[0.00704]			[0.00764]			[0.00838]
Muslim			0.0120			0.0230			0.0130
			[0.0167]			[0.0176]			[0.0189]
Monthly income per capita			0.000368			0.000329			0.000405
			[0.000341]			[0.000359]			[0.000403]
Mean among offspring with zero exposure in households with low arsenic concentration	0.081413	0.081413	0.081413	0.099345	0.099345	0.099345	0.136142	0.136142	0.136142
	[0.27349]	[0.27349]	[0.27349]	[0.29915]	[0.29915]	[0.29915]	[0.34298]	[0.34298]	[0.34298]
Observations	8120	8120	8120	8120	8120	8120	8120	8120	8120

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

(4) High concentration defined by arsenic test of closest well to residence being above 100 ppb.

(5) Sample restricted to households with arsenic concentrations greater than 60 ppb.

Table A2.6: Measured arsenic contamination and early life exposure excluding recent wells

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00537	0.00397	0.00133	0.00578	0.00373	0.000384	-0.00142	-0.00491	-0.00915
	[0.00959]	[0.00963]	[0.00941]	[0.0108]	[0.0109]	[0.0105]	[0.0129]	[0.0129]	[0.0125]
Early life exposure	-0.0450***	-0.0122	-0.00896	-0.0639***	-0.0165	-0.0177	-0.0939***	0.00948	-0.00245
	[0.0125]	[0.0171]	[0.0170]	[0.0133]	[0.0197]	[0.0195]	[0.0158]	[0.0248]	[0.0246]
High concentration*Early life exposure	0.0208	0.0228	0.0263*	0.0305*	0.0339**	0.0379**	0.0306*	0.0355*	0.0413**
	[0.0155]	[0.0154]	[0.0154]	[0.0161]	[0.0162]	[0.0160]	[0.0185]	[0.0185]	[0.0183]
Sex		0.0270***	0.0261***		0.0258***	0.0243***		0.0225***	0.0205***
		[0.00617]	[0.00615]		[0.00653]	[0.00650]		[0.00736]	[0.00734]
Parity		0.00280	0.00281		0.00367*	0.00371		0.00537**	0.00382
		[0.00188]	[0.00258]		[0.00202]	[0.00269]		[0.00218]	[0.00287]
Birth year		-0.00780***	-0.00833***		-0.00887***	-0.00996***		-0.0139***	-0.0161***
		[0.00258]	[0.00261]		[0.00264]	[0.00269]		[0.00260]	[0.00264]
Birth year squared		0.000121	0.000133		0.000116	0.000155		0.000154	0.000253**
		[0.0001]	[0.0001]		[0.000102]	[0.000104]		[0.0001]	[0.000102]
Age of mother at birth			0.000113			0.000117			0.00005
			[0.000642]			[0.000689]			[0.000747]
Mother's education			0.00370**			0.00462**			0.00528***
			[0.00181]			[0.00188]			[0.00200]
Father's education			-0.00194			-0.00260*			-0.00307**
			[0.00133]			[0.00139]			[0.00153]
Years since birth of last child			-0.00966***			-0.0123***			-0.0160***
			[0.00219]			[0.00232]			[0.00254]
Solvency			-0.00211			0.000383			0.00653
			[0.00702]			[0.00749]			[0.00783]
Land size (acres)			0.00475***			0.00367**			0.00292
			[0.00174]			[0.00176]			[0.00203]
No. of rooms in house			-0.0183***			-0.0194***			-0.0193***
			[0.00332]			[0.00353]			[0.00369]
Electricity			-0.00798			-0.0123			-0.0171**
			[0.00729]			[0.00791]			[0.00869]
Muslim			-0.00472			0.00661			0.000819
			[0.0175]			[0.0181]			[0.0195]
Monthly income per capita			0.000355			0.000349			0.000449
			[0.000389]			[0.000397]			[0.000423]
Mean among offspring with zero exposure in households with low arsenic concentration	0.083718	0.083718	0.083718	0.103159	0.103159	0.103159	0.145201	0.145201	0.145201
	[0.27704]	[0.27704]	[0.27704]	[0.30426]	[0.30426]	[0.30426]	[0.35245]	[0.35245]	[0.35245]
Observations	7746	7746	7746	7746	7746	7746	7746	7746	7746

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

(4) Households with closest wells that were, according to survey data, constructed less than 8 years ago, are dropped from the sample.

Table A2.7: Measured arsenic contamination and early life exposure, 1999 campaign cutoff

	Death under 12 mo.			Death under 24 mo.			Death under 60 mo.		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
High concentration	0.00503	0.00328	0.00186	0.00940	0.00691	0.00487	0.00389	-0.00009	-0.00255
	[0.00745]	[0.00745]	[0.00735]	[0.00835]	[0.00831]	[0.00817]	[0.0101]	[0.0100]	[0.00986]
Early life exposure	-0.0460***	-0.00514	-0.00597	-0.0567***	0.00241	-0.00287	-0.0965***	-0.00272	-0.0150
	[0.00852]	[0.0123]	[0.0123]	[0.00898]	[0.0132]	[0.0133]	[0.0112]	[0.0173]	[0.0173]
High concentration*Early life exposure	0.0241**	0.0264**	0.0296***	0.0229**	0.0260**	0.0294**	0.0289**	0.0337**	0.0376***
	[0.0112]	[0.0112]	[0.0111]	[0.0115]	[0.0115]	[0.0115]	[0.0137]	[0.0137]	[0.0136]
Sex		0.0290***	0.0277***		0.0280***	0.0264***		0.0265***	0.0242***
		[0.00499]	[0.00498]		[0.00534]	[0.00533]		[0.00594]	[0.00592]
Parity		0.00308*	0.00331		0.00439***	0.00431*		0.00626***	0.00341
		[0.00158]	[0.00216]		[0.00167]	[0.00221]		[0.00176]	[0.00236]
Birth year		-0.00622***	-0.00706***		-0.00765***	-0.00888***		-0.0138***	-0.0156***
		[0.00194]	[0.00198]		[0.00193]	[0.00199]		[0.00200]	[0.00203]
Birth year squared		0.00005	0.00009		0.00005	0.000109		0.000174***	0.000276***
		[0.00007]	[0.00007]		[0.00007]	[0.00007]		[0.00007]	[0.00007]
Age of mother at birth			-0.000248			-0.000198			0.00008
			[0.000541]			[0.000576]			[0.000622]
Mother's education			0.00121			0.00177			0.00257
			[0.00148]			[0.00155]			[0.00164]
Father's education			-0.00146			-0.00163			-0.00207
			[0.00107]			[0.00114]			[0.00126]
Years since birth of last child			-0.0107***			-0.0134***			-0.0172***
			[0.00177]			[0.00188]			[0.00203]
Solvency			0.000331			-0.00157			0.00176
			[0.00556]			[0.00599]			[0.00633]
Land size (acres)			0.00472***			0.00418**			0.00382**
			[0.00164]			[0.00170]			[0.00193]
No. of rooms in house			-0.0133***			-0.0143***			-0.0153***
			[0.00268]			[0.00298]			[0.00317]
Electricity			-0.00926			-0.0103			-0.0123*
			[0.00615]			[0.00667]			[0.00732]
Muslim			0.0113			0.0206			0.0169
			[0.0145]			[0.0148]			[0.0157]
Monthly income per capita			0.000217			0.000238			0.000321
			[0.000314]			[0.000328]			[0.000356]
Mean among offspring with zero exposure in households with low arsenic concentration	0.079649	0.079649	0.079649	0.096261	0.096261	0.096261	0.145877	0.145877	0.145877
	[0.27080]	[0.27080]	[0.27080]	[0.29401]	[0.29401]	[0.29401]	[0.35308]	[0.35308]	[0.35308]
Observations	11766	11766	11766	11766	11766	11766	11766	11766	11766

Notes:

(1) Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1. OLS regressions, linear probability models. Data from the 2007 Kishoree Kontha Baseline survey in Barisal subdistrict. An observation is a live birth. All specifications use village fixed effects and are clustered at the household level.

(2) High concentration equal to 1 if closest shallow tubewell to residence revealed in field test to have arsenic concentration above 60ppb.

(3) Early life exposure measures the fraction of offspring's life before the specified age of death cutoff (12 mos, 24 mos, 60 mos) in which he/she was potentially exposed to water from the new source (fraction of years below cutoff lived post-2000).

(4) Early life exposure defined according to campaign date of 1999 rather than 2000.

Table A2.8: DHS Sample Means

VARIABLES	Arsenic free	N	Contaminated	T-statistic
		Panel I: Exogenous variables		
Age of mother	29.72	7345/541	29.47	0.5
Age of mother at earliest birth	20.02	7345/541	19.58	1.27
Education of mother	3.38	7345/541	3.32	0.26
Education of father	4.47	7345/541	4.68	-0.83
Mean birth interval	3.04	7345/541	3.12	-0.37
Solvency	0.51	7345/541	0.51	0.12
Land size (acres)	0.67	7345/541	0.70	-0.14
Number of rooms in house	2.93	7345/541	2.98	-0.51
Electricity	0.43	7345/541	0.47	-1.74
Muslim	0.89	7345/541	0.89	0.59
Years living in house	26.90	3070/205	26.02	0.7
Head of household works in agriculture	0.33	3135/215	0.38	-1.24
Head of household works in business	0.43	3135/215	0.43	0.19
Sufficient earnings	2.17	2901/197	2.13	0.62
		Panel II: Endogenous variables		
Arsenic concentration (ppb)	1.54	7345/517	6.76	-94.86
Whether well painted	0.35	6250/520	0.40	-1.92
Fraction of deaths under 12 mo.	0.06	7345/541	0.06	-0.01
Fraction of deaths under 24 mo.	0.07	7345/541	0.07	-0.33
Fraction of deaths under 60 mo.	0.08	7345/541	0.08	-0.2
M:F sex ratio	0.52	7345/541	0.49	1.25
Number of offspring in family	0.25	7345/541	0.26	-0.51
Deep tubewell mentioned as a source for drinking	0.04	7341/540	0.02	2.45

Notes:

(1) All means account for village fixed effects

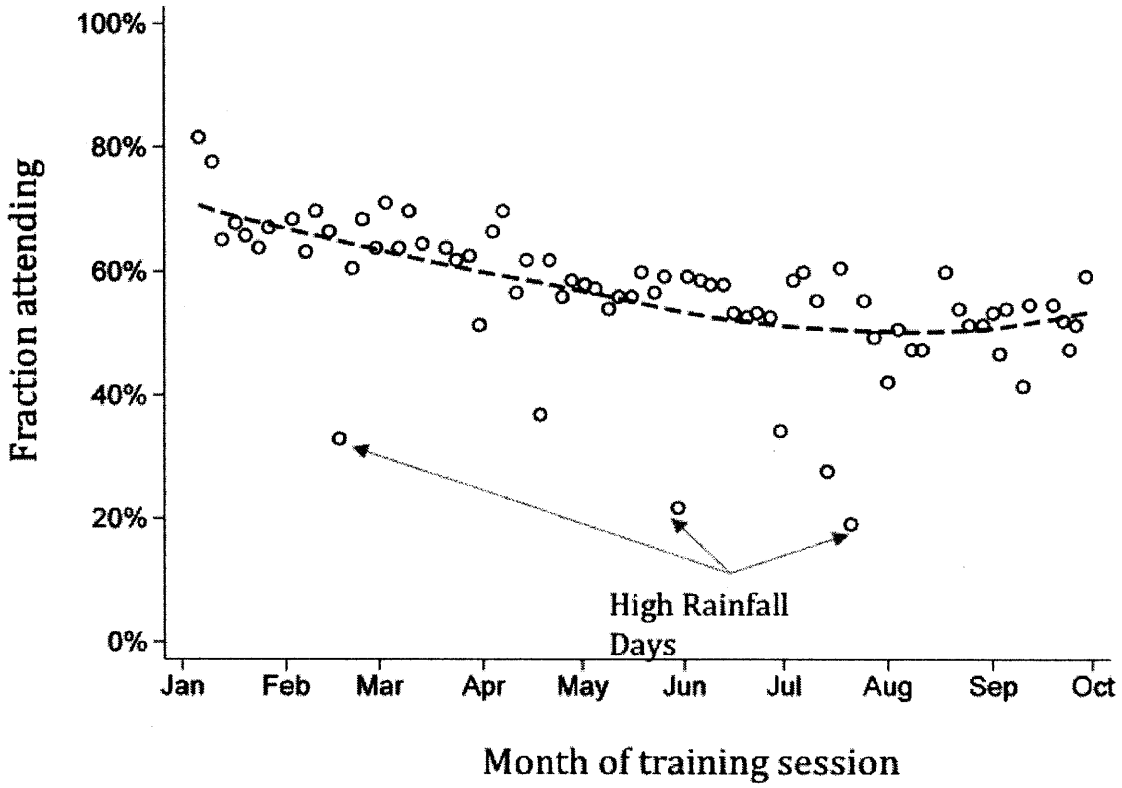
(2) Solvency defined sufficient food consumption in whole year

(3) Sufficient earnings defined as whether man believes his earnings sufficient for family; limited sample

Appendices for Chapter 3

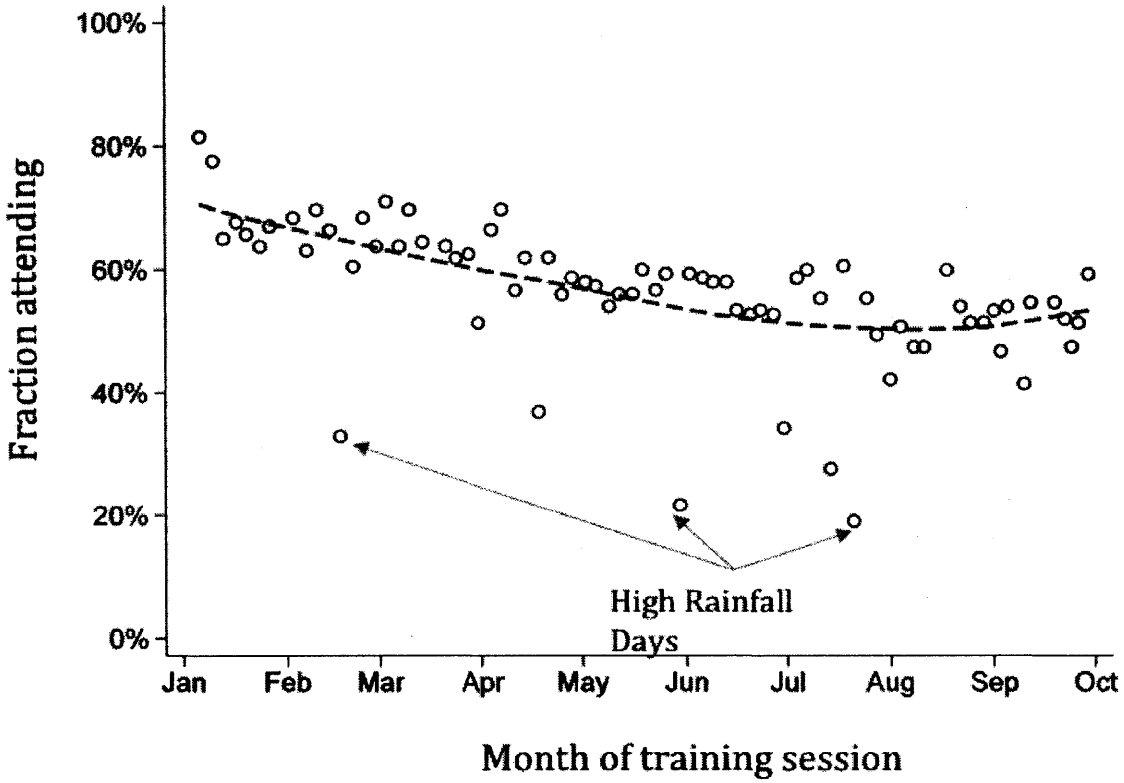
Appendix Figures

Figure A3.1: Attendance in training course by month



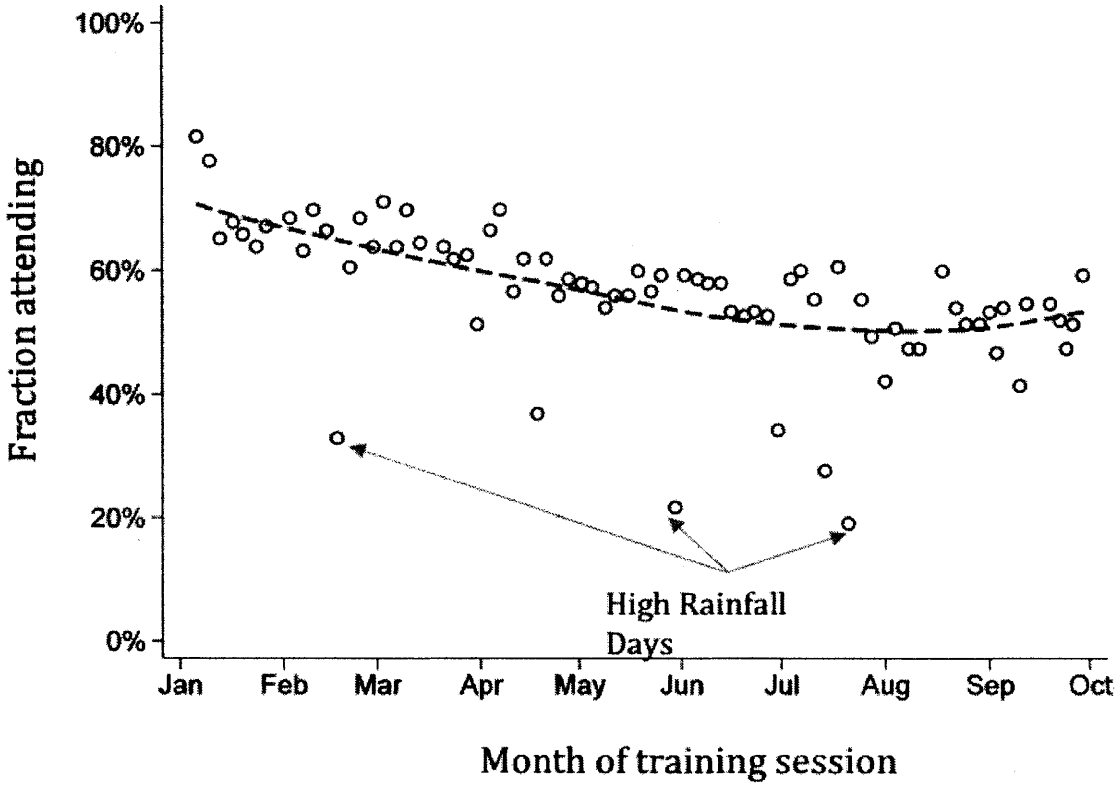
This figure shows attendance in the training course, where each scatter point shows the fraction of informal providers randomized into the treatment group who attended the session. Extreme low values correspond to days with heavy rainfall in the region with three examples marked. The dashed line shows the non-parametric fit relating attendance to the date of the session.

Figure A3.2: Impact of training on condition-specific components of correct case management, odds-ratios of the intention-to-treat estimator using a logistic regression



This figure shows the impact of training on the condition-specific components of correct case management. The odds-ratios are computed from a logistic regression model relating each outcome variable to treatment status, with condition-specific dummies and age of the provider as additional controls. Every component of correct case management improved, although the small sample size decreases the precision of the estimates. Estimates that are significant at the 95% level of confidence are marked (**) and at the 99% level of confidence (***).

Figure A3.3: Attendance over the duration of the training course simulated at different distances of providers from the training center



This figure shows the simulated attendance over the duration of the training course depending on the distance of the training center from providers. The figure uses the coefficient estimates of daily attendance from the multiple linear regression in Table S2. 95% confidence intervals are shaded in gray. If the coefficient on the effect of distance on attendance has a causal interpretation, this would imply that moving the training center next to the provider's clinic would increase attendance to above 85%. Conversely, if every provider was 40Km from the training center, attendance would drop to 45%.

Appendix Tables

Table A3.1: Baseline balance and assessment of differential loss to follow-up by treatment status

	(1) Control Mean (Full Sample) N = 152	(2) Treatment Mean (Full Sample) N = 152	(3) Control Mean (SP Followup Sample) N=133	(4) Treatment Mean (SP Followup Sample) N=134	(5) T-Test Difference (4) - (3)
Mean Age	39.29	41.10	40.37	42.00	1.63
	(37.87, 40.71)	(39.6, 42.6)	(38.94, 41.8)	(40.45, 43.55)	(-0.531, 3.794)
Mean Years of Experience	12.97	13.32	13.96	14.12	0.16
	(11.63, 14.31)	(12, 14.63)	(12.54, 15.38)	(12.72, 15.52)	(-1.709, 2.023)
Mean Consultation Fee (INR)	10.78	10.33	10.93	10.34	-0.60
	(9.07, 12.49)	(8.78, 11.88)	(9.05, 12.81)	(8.73, 11.94)	(-3.085, 1.893)
Mean Consultation Fee (USD)	0.17	0.17	0.17	0.17	-0.01
	(0.15, 0.2)	(0.14, 0.19)	(0.14, 0.21)	(0.14, 0.19)	(-0.049, 0.03)
Mean Village Literacy Rate	0.53	0.52	0.53	0.52	-0.01
	(0.52, 0.54)	(0.51, 0.53)	(0.51, 0.54)	(0.5, 0.53)	(-0.03, 0.008)
Proportion Male	0.95	0.96	0.99	0.97	-0.02
	(0.9, 0.97)	(0.92, 0.98)	(0.96, 1)	(0.93, 0.99)	(-0.055, 0.01)
Proportion Completed High School	0.61	0.63	0.60	0.63	0.03
	(0.53, 0.68)	(0.55, 0.7)	(0.52, 0.68)	(0.54, 0.7)	(-0.099, 0.15)
Proportion with No Formal Training	0.74	0.76	0.71	0.78	0.06
	(0.67, 0.81)	(0.68, 0.82)	(0.63, 0.78)	(0.7, 0.84)	(-0.046, 0.169)
Proportion with Minimal Formal Training	0.26	0.24	0.29	0.22	-0.06
	(0.19, 0.33)	(0.18, 0.32)	(0.22, 0.37)	(0.16, 0.3)	(-0.169, 0.046)

Notes: Fees in 2014 Indian rupees (INR) and 2014 U.S. dollars at ($1 = INR62.5). Purchasing-power-parity adjustment would result in an exchange rate of INR16.76 per dollar. "No Formal Training" includes providers with no formal medical training, although they may have degrees in community medicine. Literacy rates are from the 2011 Indian population census. 95% confidence intervals are presented below the means and estimated t-test differences. Wilson intervals without continuity correction are used for indicator variables, and t-test confidence intervals are adjusted for village-level clustering.

Table A3.2: Correlates of daily training attendance

	(1) Bivariate	(2) Multivariate
Distance to Training (Haversine, km)	-0.013**	-0.012**
	(-0.024, -0.003)	(-0.023, -0.002)
Rainfall (mm)	-0.004***	-0.004***
	(-0.005, -0.003)	(-0.005, -0.003)
Male	0.031	-0.008
	(-0.224, 0.286)	(-0.262, 0.247)
Age	0.005	0.005
	(-0.001, 0.012)	(-0.002, 0.011)
No Formal Training	-0.017	-0.019
	(-0.142, 0.108)	(-0.14, 0.102)
Constant	–	0.595***
	–	(0.191, 0.999)
Block FE	Yes	Yes
N	11,400	11,250

Notes: All regressions are at the provider x day level and include block fixed effects. The distance to the training center is based on GPS coordinates for the training location and each provider and computed as the straight-line distance between the two points using the Haversine formula. 95% confidence intervals are presented below the means and estimated coefficients, and are adjusted for clustering at the provider level.

Table A3.3: Frequency distribution of training attendance

Fraction of Sessions Attended	Number of Providers	Percent
0-10%	33	21.7
10-20%	9	5.9
20-30%	7	4.6
30-40%	3	1.9
40-50%	5	3.29
50-60%	6	3.9
60-70%	10	6.58
70-80%	14	9.2
80-90%	40	26.32
90-100%	25	16.45

This figure shows the simulated attendance over the duration of the training course depending on the distance of the training center from providers. The figure uses the coefficient estimates of daily attendance from the multiple linear regression in Table S2. 95% confidence intervals are shaded in gray. If the coefficient on the effect of distance on attendance has a causal interpretation, this would imply that moving the training center next to the provider's clinic would increase attendance to above 85%. Conversely, if every provider was 40Km from the training center, attendance would drop to 45%.

Supplementary Methods

Description of Study Instruments: Clinical Observation Tool

The clinical observation tool assesses the behavior of health care providers with real patients using observers who remain in the providers' clinics for the entire duration of their practice for one day. Since providers saw patients at multiple times, and often at night, observer timings could vary, starting at 0800 when the clinic opened and finishing as late at 2300 if the provider closed the clinic late. During the interaction with each patient, a structured questionnaire was completed by the observer with details such as the number and type of examinations completed, the number of questions asked and the consultation time. Observers were not allowed to speak to the provider or patient at any time during the consultation. At the end of the consultation, the provider would list the medicines dispensed or prescribed for the patient, which would be noted in the structured observation form. The clinical observation tool was adapted for this study from previous use in a number of studies.[1]

Description of Study Instruments: Standardized Patient Tracer Conditions

Three tracer conditions were used to assess the impact of the multi-topic training for informal providers. As documented in the text, these were designed to ensure that we captured the range of conditions that the informal providers may be asked to manage as the only health care providers in their locations. The tracer conditions had been previously validated and used in the Indian rural context.[2] These were:

- Chest pain suggestive of unstable angina: The 40-45 year old SP begins his interaction with the provider with the opening statement: "Doctor, this morning I had pain in my chest."
- Respiratory distress suggestive of asthma: The 25-30 year old SP begins his interaction with the provider with the opening statement: "Doctor, last night I had a lot of difficulty with breathing."
- Diarrhea in a child sleeping at home: The father of the child begins his interaction with the provider with the opening statement: "My child has been having loose motion. Can you give me some medicines?"

For each of these tracer conditions, we used definitions of correct management appropriate for informal providers, which in practice implies that for conditions like asthma, "referral" is coded as appropriate. The table below documents the condition-specific definitions used in the manuscript:

Description of standardized patient cases

Case	Case description	Case was designed to assess quality of care for a person who	Expected correct case management
SP1	Classic case of unstable angina, with patient describing crushing chest pain. Upon further questioning, the patient will reveal that the pain radiated down his arm, that he is a regular smoker and that there is a history of heart disease in his family	Presents with an urgent case that requires triage to higher-level care with recommendations for a referral, an ECG and aspirin	Referral or the provision of aspirin
SP2	Classic case of presumed asthma with breathing difficulty. Upon further questioning, the patient will reveal and a previous history of similar episodes with environmental triggers.	Presents with a potentially chronic condition that needs to be diagnosed and potentially referred or managed	Referral, prescription of an inhaler, or prescription/provision of corticosteroids
SP3	Child with diarrhea. Upon further questioning the patient will reveal that the diarrhea is bloody with mucus, but that the child is not severely dehydrated	Presents with a common condition among children in these contexts that needs to be carefully managed; follow-up instructions as appropriate	Provision of ORS or request by the provider to see the child before providing treatment

As an alternative to these definitions of correct case management, the treatment grading assigned by the raters allows for finer distinctions as well as the consideration of additional medicines that may have been unnecessary. The results in Table 2 (main text) therefore suggest that the findings are robust to alternate methods of assessing the quality of the treatment. Nevertheless, robustness to alternate definitions was also assessed. For instance, for diarrhea the provision of ORS alone could be used as correct case management; for angina, referral or referral for an ECG could be used as the appropriate definition for correct case management. Depending on the specific definitions used, the most conservative estimate for the OR for the intention to treat estimator was 1.33 (95% CI: 0.949, 1.861) compared to 1.40 (95% CI: 1.017, 1.931) in the main text.

Checklist Completion

Compliance with the recommended condition-specific checklist of items was based on international and national guidelines referenced in the text. The table below presents each checklist item and the fraction of study participants who asked the question or performed the relevant examination, stratified by treatment assignment.

Recommended checklist items for angina

History Questions	Control	Treatment
Where is the pain?	77%	81%
When did this pain start?	38%	49%
What were you doing when the pain began?	7%	9%
Severity of Pain	43%	40%
Radiation (is the pain "walking"/mobile?)	10%	11%
Previous history of similar pain?	48%	61%
Since when have you had these pains?	34%	35%
Does the pain change? What do you do that makes it worse?	5%	5%
Quality of pain (heavy/dull vs. sharp)	21%	26%
Does the pain change/increase with inhalation or exhalation?	14%	19%
Shortness of breath	21%	18%
Nausea or vomiting	24%	25%
Excessive sweating	7%	9%
Age	25%	31%
Other Medical History	2%	3%
Whether suffering from diarrhea	1%	2%
Whether suffering from constipation	9%	2%
Whether suffering from pain in the stomach	17%	16%
Whether stools are normal?	54%	49%
Acidity/Gas-related questions	62%	68%
Whether suffering from fever	21%	20%
Whether patient smokes	4%	5%
Whether patient drinks alcohol	2%	2%
Patient's occupation	66%	71%
Did brother, sister, or parents have a similar problem?	2%	7%
Exams	**Control**	**Treatment**
Pulse	50%	60%
BP	59%	62%
Auscultation (Front)	75%	76%
Auscultation (Back)	24%	29%
Temperature (Thermometer)	2%	2%
Temperature (Touch)	12%	16%

Recommended checklist items for asthma

History Questions	Control	Treatment
Probes about breathing difficulty (current episode)	92%	95%
Enquires about cough	65%	71%
Probes about expectoration, i.e. does anything comes up such as mucus/blood or is this a dry cough?	16%	29%
Have you had breathing problems previously?	46%	58%
Probes about breathing difficulties (previous episodes)	19%	28%
Since when have you had these breathing problems?	40%	53%
How often does this happen?	9%	11%
Is the shortness of breath constant or episodic?	8%	11%
What triggers the episodes? (e.g. dust, pollution, bad air quality, cold)	25%	33%
How long does an attack last?	4%	2%
Did you eat anything that you had not taken before?	1%	2%
Childhood illnesses especially re: cough or breathing problems	4%	3%
Patient age	32%	33%
Whether suffering from fever	55%	68%
Whether suffering from chest pain	54%	60%
Whether suffering from weight loss	0%	0%
Whether suffering from night sweats	0%	2%
Whether suffering from throat or upper respiratory symptoms	7%	17%
Whether patient smokes	14%	16%
Patients' occupation	46%	40%
Questions regarding family history	6%	14%
Exams	**Control**	**Treatment**
Pulse	46%	54%
BP	30%	29%
Auscultation (Front)	64%	74%
Auscultation (Back)	34%	42%
Throat	9%	22%
Temperature (Thermometer)	3%	4%
Temperature (Touch)	29%	43%

Recommended checklist items for diarrhea

History Questions	Control	Treatment
Age of child	96%	98%
Quality of stool (including blood)	54%	56%
Frequency of stool	80%	89%
Quantity of stool	34%	40%
Questions about urination	13%	16%
Is child active/ playful? General behavior of child.	6%	9%
Whether suffering from fever	55%	60%
Whether suffering from abdominal pain	55%	68%
Enquires about vomiting	75%	74%
Previous health status	11%	17%
Source of drinking water? Boiling water? From tap? How is it stored?	3%	2%
Enquires about food preparation	6%	9%
What has the child eaten?	18%	38%
Whether the child is taking fluids	2%	6%
Any question about neighborhood/family background	0%	0%
Any questions about the physical Environment	0%	0%

Standardized Patient Recruitment and Training

A total of 16 SPs were recruited from an initial group that was extensively screened and trained for 150 hours by a multi-disciplinary team. Protocols developed for a past deployment in rural India were followed in this study.2 There were no adverse events for SPs during the fieldwork, reflecting the SPs ability to use evasion techniques developed during training to successfully avoid any such situations. During the interactions, SPs noted if the provider challenged the presentation, forcing them to disclose that the interaction was part of the SP study; no such challenges were recorded so that detection rates were effectively zero.

Previous studies have documented low inter-rater differences in quality of clinical practice assessed using SPs.3 Nevertheless, in order to minimize potential bias from variation in care across different SPs, we ensured that each SP visited multiple providers and each provider was visited by multiple SPs. The assignment allows us to use an additional full set of SP indicators to control for potential inter-rater differences that may be correlated with the treatment status. Including the full set of indicators does not alter any of the results, although there is a marginal decline in precision for some outcome variables (results available on request).

Multiple Regression Models

Regression models used to assess program impact can differ in assumptions regarding the error term and the estimator of interest. To assess robustness and present a range of estimates relevant for policy, we use both linear and non-linear models and present both the intention-to-treat estimator (the average treatment effect in the program evaluation literature) and the instrumental variables estimator (the treatment on the treated in the program evaluation literature).

For the intention-to-treat estimators, we present both the odds-ratio and the marginal effects in the logistic regression model and the marginal effects in the linear regression model; this allows for easier comparability with the marginal effects in the instrumental variables specifications. For the instrumental variables estimates, we instrument attendance over the course of the training program with treatment status. Estimates from the linear regression model are robust to misspecification in the functional form of the error term but have a higher root mean square error in smaller samples, leading to lower precision.[4] In our estimates, however, the confidence intervals are sufficiently precise that further structural assumptions that can lead to model misspecification (such as normally distributed errors in the Instrumental variables probit model) are not required. We therefore present marginal effects from the linear regression model only. Model specifications are as follows:

- **Intention-to-treat, Linear Regression:** $y_i = \alpha + \beta * T_i + \gamma * Age_i + \sum_1^3 Condition_i + \epsilon_i$
  - Where T_i is a dummy variable that takes the value 1 if the provider is in the treatment group and 0 if the provider is in the control group; Age_i is the age of the provider and $Condition_i$ are condition specific indicator variables.
- **Intention-to-treat, Logistic Regression:** The logit model is specified as $y*_i = \alpha + \beta * T_i + \gamma * Age_i + \sum_1^2 Condition_i + \epsilon_i$
  - Where ϵ_i $Logistic(0,1)$ and the outcome variable, y_i, is a dichotomous variable such that $y_i = 1$ if $y_i^* > 0$ and $y_i = 0$ otherwise. We present both the odds-ratio (OR) as a measure of the relative effect and the marginal effects (ME) as a measure of the absolute effect. The latter also facilitates direct comparison with the instrumental variables estimates.

- **Instrumental Variables, Linear Regression:** The instrumental variables specification recovers the impact of attendance on outcomes. Since attendance in the as-treated regression is endogenous and biased by selection, we use treatment status to instrument for attendance. Specifically, the "as-treated" regression is: $y_i = \alpha + \beta * Attendance_i + \gamma * Age_i + \sum_1^3 Condition_i + \epsilon_i$
  - Here, $Covariance(Attendance_i, \epsilon_i)! = 0$, if, for instance, more motivated providers have higher attendance, and therefore the estimate β is asymptotically inconsistent.
  - Let Z be an instrument such that the following two assumptions are satisfied for the conditional covariance:
    1. $Covariance(Attendance_i, Z_i|X_i) \neq 0$ and
    2. $Covariance(Z_i, \epsilon_i|X_i) = 0$
  - Then, the instrumental variables estimate $\beta_{IV} = \frac{Covariance(y_i, Z_i|X_i)}{Covariance(y_i, Z_i|X_i)}$ is a consistent estimator of β.
  - Using assignment to treatment as the instrument, Z_i, the two assumptions required for consistency of the instrumental variables estimator are satisfied:
    1. $Covariance(Attendance_i, T_i) \neq 0$ as assignment to treatment increased attendance by 56%. To test formally the significance of the result and check if there is a "weak instrument" problem, the F-statistic from this regression is used as a diagnostic, with F-statistics below 10 signaling such a problem. In our case, the F-statistic > 300 in all cases.
    2. $Covariance(T_i, \epsilon_i) = 0$ as T_i is randomly assigned, and therefore orthogonal to the error term by assumption.

References for Supplementary Methods

1. Das, J., Hammer, J. (2014) "Quality of primary care in low-income countries: facts and economics." Annu. Rev. Econ., 6(1): 525-553.

2. Das, J., Holla, A., Das, V., Mohanan, M., Tabak, D., Chan, B. (2012) "In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps." Health Affairs, 31(12): 2774-2784.

3. Das, J., Kwan, A., Daniels, B., Satyanarayana, S., Subbaraman, R., Bergkvist, S., Das, R., Das, V., Pai, M. (Forthcoming) "First use and validation of the standardized patient methodology to assess quality of tuberculosis care: a pilot, cross-sectional study." Lancet Infectious Diseases.

4. Angrist, J. D., Pischke, J. S. (2008) Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

The Liver Foundation Training Program Curriculum

Overall aim of the IP Training Program

To improve the quality of curative care provision, including primary, lifesaving, and referral services in rural areas, by training the existing and functioning health care human resources of the informal sector within the community.

Characteristics of Informal Providers

An IP is a self-employed healthcare provider who lacks formal medical training or qualifications but is engaged in the delivery of medical services and medicines. IPs have different degrees of formal education, ranging from secondary to graduate levels, and may have some level of informal training either through previous employment in a clinic or nursing home or through time spent as an apprentice in a medicine shop. IPs generally function within a defined community with whom they have considerable linkages but are not within the ambit of any existing formal healthcare programs or regulatory protocols.

Training Objectives

The training program seeks to:

- Primary
    1. Ensure IPs possess a basic scientific understanding of human health and disease, in order that they be able to decide "what not to do" and "what to do" in service delivery settings in a village.
    2. Decrease rash clinical acts and irrational drug usage by IPs by increasing cognizance of the potential of harm thereof.
    3. Improve referral of patients by IPs to the formal sector.
    4. Improve integration of IPs with the mainstream public health delivery program.
- Secondary
    1. Convert unqualified IPs to enriched health workers.
    2. Provide socially relevant cultural and behavioral inputs to improve social accountability.

After the training, IPs should be able to identify common ailments, provide early primary remedy, discern cases that require higher-level care and refer such cases to doctors and health facilities. IPs should also be able to manage emergencies and provide potentially lifesaving interventions within their locality, prior to stabilizing the patient for transport to a facility. IPs should understand their potential role in serving the community, while recognizing that they are not "doctors", but rather health workers who have the respect of the community.

Methods of Instruction

The training consists of primarily classroom instruction in the regional language and through bilateral interactions between students and teachers based on a structured curriculum. Simulations of common clinical settings encountered in rural areas along with real-life demonstrations are frequently incorporated into the training.

Training Duration and Structure

Liver Foundation's training program is a 9-month long program that is taught through two sessions a week, on two different days, with each session consisting of at least two hours of interaction. There are a total of at least 72 sessions and 150 hours of interaction. The trainees maintain their clinics as usual during the training period. Teachers are all medical doctors with extensive work experience in rural areas and the core teaching group consists of three doctors (Dr. Saibal Majumdar , Dr. Kajal Chatterjee and Dr. Rupak Ghosh), with periodic input from other doctors on specific topics. Training is divided into three phases:

Training phase	Duration	Teaching hours	Synopsis
Initiation	First 10 weeks	40 hours in 20 sessions	The primary objective is to introduce trainees to the habit of learning, inculcate medical discipline, and provide behavioural and ethical inputs. Clinical course content focuses on human anatomy and physiology.
Consolidation	Second 16 weeks	70 hours in 35 sessions	The primary objective is to provide a deeper understanding of human diseases and emergencies in the rural primary care setting. The training stresses how to identify potentially dangerous situations, restrict harmful practices, and diagnose and manage minor illnesses.
Enrichment	Final 10 weeks	40 hours in 17 sessions	The enrichment phase continues with greater exposure to clinical conditions, but the main focus is on existing public health programs and the role of IPs in such programs.

Course Content in Initiation Phase (20 sessions)

1. Introduction to Ethics (2 sessions): Health worker cultural norms and legal aspects safeguarding clinical practice in India. The distinction between social acceptability and regulatory legality.
2. Outline of Human Anatomy and Physiology (14 sessions): Taught by two teachers in each session with a focus on:

    1 Organization of the human skeleton and organs with charts and a demonstrative atlas

    i. Discussion of cells, tissues and organs

    ii. Discussion of blood and body fluid compartments

    iii. Discussion of how each organ is connected with other organs in structure and function

    2 Bones and joints of the body

    i. How movements are coordinated across different joints

ii. The impact of trauma and re-union process

iii. How the brain controls body movements and discussion of paralysis and disability

3 Structure and organization of heart, lung and vasculature and how these are coordinated. Major vessels of the body. Coronary circulation, cerebral circulation and impairment of flow. Circulations in critical regions - renal. Coronary artery disease - explanations of pathophysiology. Lung and pleura. Blunt trauma and collections in chest.

4 Structure of the brain and big nerves of the body and how the nerves are interconnected and controlled.

5 Kidney, gastrointestinal tract and liver structure outlines. Layers in the wall of viscera.

6 Pelvic organs and reproductive system in females with focus on birth passage structure.

7 Eye, nose and ear structure.

3. Drug Usage Principles and Adverse Effects (4 sessions):

1 What is a pharmaceutical agent? The nature of a pharmaceutical agent, its modes of administration and impacts of each. Principles of drug safety and the importance of drug dosages as well as exposition on the unpredictable nature of parenteral route of administration. Discussion of drug safety regulations in the country.

2 Principles of oral drug therapy and explanation of how these are guided by absorption, distribution and elimination.

3 Examples of unnecessary or harmful medication prescription including drawbacks of combination therapy. Unpredictability of drug behaviour in different individuals. Drug quality and sensitivity of active ingredients to temperature.

4 Life threatening side effects of drugs including the Steven Johnson syndrome, extra pyramidal syndromes, side-effects of steroids, and the differential impact of drugs on special populations, including children, pregnant women and the elderly. Discussion of the importance of drug pricing and client behavior.

Course content in Consolidation Phase (35 sessions)

1. How human diseases are caused, including the role of microbes (bacteria, virus, amoeba and helminths). How microbes cause disease through tissue invasion and inflammation. Major communicable and non-communicable diseases in India, their changing pattern and the role of IPs in this setting. (2 sessions)

2. Outlines of clinical decision making, history taking and examinations, including how much to question patients during emergencies before taking action and the role of a rational approach. (2 sessions)

3. Clinical condition specific approach: how different illnesses are caused, the environment-agent link, what to do in clinical settings, minimal curative care provisions and referral principles. The potential complications of delaying referrals. (10 sessions)

1 Fever

2 Diarrhea

3 Chest pain and respiratory distress

4 Abdominal pain and vomiting

5 Headache and unconsciousness

6 Cough, sore throat, runny nose and respiratory infections

7 Flatulence, gaseousness and acidity

8 Trauma and burn - primary care

9 How to identify a sick child and clinical actions for sick children

10 Complications of pregnancy and the role of an IP

4. Non-communicable diseases and how to manage such diseases, with a discussion of lifestyle changes including exercise, diet and quitting smoking. (6 sessions)

    1 The role of the IP in managing hypertension, diabetes, coronary artery disease, cancers and COPD.

5. Specific contextually relevant communicable diseases, including principles of management. (8 sessions)

    1 Malaria

    2 Tuberculosis

    3 Enteric Infections

    4 Pneumonia

    5 Skin and soft tissue infections

    6 Others, including discussions initiated by trainees

6. Pregnancy and child care. (5 sessions)

7. Review and recapitulation. (2 sessions)

Course content in Enrichment Phase (17 sessions)

1. Public health priorities in India: infections, malnutrition, education, poverty and cleanliness.

2. Safe drinking water and the role of improved sanitation and healthy behaviors.

3. Waterborne diseases and causes, early management focusing on Oral Rehydration Therapy and no use of antibiotics or anti-amoebic drugs. Cholera, dysentery and water purification measures in villages.

4. Immunization schedule and safe immunization.

5. Tuberculosis control: suspecting TB, program organization of TB control in India, contra-indicated actions in TB suspects.

6. Health indices and their role in health care: discussion of the Crude Death Rate, Maternal Mortality Rate, Infant Mortality Rate and Neo-natal Mortality Rate and how to maintain these data in a given locality.

7. Contraception methods

8. Child Nutrition
9. Existing vertical public health programs and the role of IPs through coordination.
10. Summary

Intramural Assessment

Three quarterly assessments are conducted, one at the end of each phase. At each assessment, trainees are given grades with assessments conducted by external examiners.