Information Campaign on Water Quality and Marriage Market: The Case of Arsenic Exposure in Rural Bangladesh

Shyamal Chowdhury* Prachi Singh†

September, 2019

*University of Sydney
†Indian Statistical Institute, Delhi & Brookings India

Information Campaign on Water Quality and Marriage Market: The Case of Arsenic Exposure in Rural Bangladesh

Abstract

Arsenic contamination of drinking water has caused a major health emergency in Bangladesh owing to multiple health problems associated with it, which range from skin lesions to various types of cancers. However, it remained largely unknown and became a public knowledge only in 2002 through a nationwide information campaign. We study marriage patterns in Bangladesh and associate them with the information campaign which informed people about harmful health effects of arsenic and marked water sources into safe or unsafe categories. Using difference-in-difference we analyse the age at marriage and the bride price agreed at the time of marriage and find that both of them reduce in arsenic affected areas in comparison to non-arsenic affected areas in response to information campaign. We find that age at marriage for males reduced by 1 to 3.5 percent for males and by 0.67 to 1 percent for females. Bride price also reduced by around 60 percent. Our results are indicative of a behavioural change in marriage market found using census data, and have additionally been replicated using precise spatial contamination information constructed by using Demographic and Health Survey datasets for Bangladesh.

JEL Classification: I12 J12 R11

Corresponding author:

Prachi Singh
Economics and Planning Unit
Indian Statistical Institute
New Delhi
Email: prsingh@brookingsindia.org

Information Campaign on Water Quality and Marriage Market: The Case of Arsenic Exposure in Rural Bangladesh

1 Introduction

Arsenic contamination of drinking water is a major public health problem all over the world. While it is fairly wide-spread, having being detected in atleast 70 countries [1], it is especially severe in South Asia[2], and in particular in Bangladesh. In order to address this problem, while various solutions have been suggested: for example, developing alternative water sources and methods for arsenic testing and removal (Ravenscroft et al (2009)) ; educating the population through information campaigns still remains one of most important public policy tools. Given the gravity of the problem, Bangladesh was one of the first countries to inform it's exposed population about water quality and the various harmful effects of consuming arsenic on human health. In this paper, we seek to estimate the effect of the information campaign on marriage market outcomes, focusing on two outcomes – the age at marriage and the bride price (an integral part of marriages in Islamic culture).

Our study is motivated by the recognition by researchers and commentators that while the health effect of arsenicosis (arsenic poisoning) have been extensively studied, there is relatively sparser literature on the socio-demographic (and mental) effects of such poisoning (Brinkel et al. (2009)). If such effects are indeed important, then information campaigns increasing awareness may have important demographic consequences, some unintended. For example, while an information campaign telling communities that they are at risk of arsenic exposure may help people take precautions to prevent exposure, it may also change social perceptions towards adults who have already been exposed, irrespective of whether they show any symptoms of poisoning. These changed perceptions would then cause changes in

[1] Ravenscroft (2007)

[2] Discussed in Mukherjee et al. (2006)

market marriage equilibrium, some of which may be socially undesirable. Bride price is one such outcome and is a subject of our study. To elaborate further, bride price is the amount which is agreed upon at the time of the marriage, which has to be paid by the groom's family to the bride's family in the event of a divorce. Since perceptions of future health, future physical appearance and longevity, informed by information campaigns, may affect the equilibrium price, we focus on this important outcome of the marriage market. Further, for those who live in affected regions, greater awareness that they may be affected by arsenic, may cause behavioural changes: for example, it is plausible that it would make them more risk averse in the marriage market. Such changes in risk perception may significantly affect age at marriage, with more risk averse males marrying sooner than their more risk-loving counterparts (Spivey (2010)) fearing the onset of their symptoms. Hence one may expect changes in the age of adults in the marriage market as a consequence of greater awareness due to information campaigns . Our paper is one of the first to look at the impact of such an information campaign in Bangladesh, aimed towards health on these two important outcomes of the marriage market.

The information campaign was implemented from 1999 to 2005[3] and it was designed to create awareness about ill effects of arsenic in drinking water, visually complementing the information by painting tubewells into red (dangerous) and green (safe) categories and pursuing users to switch to safe sources of drinking water. It also suggested mitigation strategies such as shifting to a safer well in the neighbourhood or well-sharing and finally it informed people about harmful health effects of arsenic exposure via public forums. The information campaign was successful in generating awareness and various studies (Chen et al. (2007); Opar et al. (2007); Bennear et al. (2013); Keskin et. al (2017); Jakariya, M (2007) etc) have found evidence on reduction of usage of contaminated wells in response to the campaign.

The health outcomes associated with arsenic affect multiple attributes which matter in

[3]World Bank Completion report 2007 mentions that during the initial 2.5 years the programme didn't take off but gained steam only later that is in 2002

marriage market. Kalmijn discusses in his survey that marriages exhibit sorting of prospective matches along many attributes such as age, education, income, race, height, weight, and other physical traits indicative of health status (Kalmijn (1998)). Buss in his study which spanned 37 cultures found that females value *resource acquisition* in males, while males place high value on *reproductive capacity* in females (Buss (1989)). Other research in marriage market has also established that physical attractiveness and BMI (a rough measure of health) is also valued in spousal match process. Most of these "valued" characteristics are adversely affected by arsenic exposure and thus we explore how individuals looking for a prospective match in marriage market, tend to react to a negative information about health outcomes associated with arsenic exposure.

Information campaigns, especially the ones which have gain framing in their messaging like "exercising regularly can help you lose weight" are found to be more effective when it comes to inducing behaviour change (Gallagher et al. (2011)). Studies have documented how information campaigns nudge people to change behaviours. For example, Kirby et al. using a randomised control trial find that information about human immunodeficiency virus (HIV) and sexually transmitted disease (STD) had an effect in terms of reducing risky sexual behaviour in males who previously indulged in unprotected sex (Kirby et al. (2004)). Duflo analysis effect of HIV related curriculum in schools in Kenya and finds that girls switch to committed relationships and they are significantly more likely to report faithfulness as a way they protect themselves from HIV (Duflo et al. (2015)).

To evaluate the impact of the information campaign, the paper uses a Difference in Difference (DID) strategy. As mentioned above, the information campaign was implemented all over Bangladesh in 2002 [4]. However, since it informed people about arsenic contamination, the campaign was more relevant for those areas which had high level of arsenic contamination. Given this, we define sub-districts with arsenic levels in excess of 50ug/litre as the treatment group and those below as the control group. These arsenic levels were based on British

[4] Keskin et al. (2017) also use the same cutoff for their analysis due to late actual implementation of the programme

Geological Survey (BGS) which tested a representative sample of tubewells in 61 districts of Bangladesh. The BGS was conducted in 1999 before the information campaign was implemented. Hence our DID strategy considers the period 1990-2001 as the pre-intervention period and the years 2002-2011 as the post intervention period and compares the difference in the pre-post period outcomes for the treatment group relative to an analogous difference for the control group. In other words, we use the differential impact of information campaign in sub-districts with differential level of arsenic contamination at the sub-district level. In our specifications, we account for sub-district level fixed effects to account for initial differences in outcomes due to location (and differing levels of arsenic contamination) and year of marriage fixed effects to account for secular trends in outcomes. Further we include district level linear trends in our specifications to remove the impact of district level time varying variables.

Some of the crucial assumptions behind this empirical strategy are important to point out: first, while the information campaign is implemented all over Bangladesh in 2002, it really has greater relevance in sub-districts with high arsenic contamination. If the campaign also had a similar impact in low arsenic contamination sub-districts, this will tend to attenuate the impact of the intervention. Second, since the initial levels of arsenic are different between the treatment and control group (by design), the assumption is that the trends in outcomes, if there was no information campaign, would not vary by these initial arsenic levels. To be more precise, our identifying assumption is that conditional on sub-district and year of marriage fixed effects and district linear trends, our marriage market outcomes do not changing differentially in arsenic versus non-arsenic contaminated sub-districts other than reasons related to information campaign. To assuage concerns on this front, we conduct an event study to confirm that this identifying assumption does hold true, that is the trends in age at marriage (and bride price) are not statistically different between contaminated and uncontaminated areas in the period before information campaign was executed.

Data for age at marriage analysis is sourced from Census Data of Bangladesh for 2011. We use data on all married individuals captured in the census. For data on bride price primary data from Palli Karma-Sahayak Foundation (PKSF) Survey conducted by researchers from

University of Sydney in December 2010-January 2011. This survey contained modules on marriage, divorce, bride price and dowry. Additionally this survey contains sub-district identifiers which allow us to conduct the DID analysis. Using the sample of individuals who got married between 1990 and 2011[5], we find that both males and females got married earlier and bride price reduced after the information about arsenic in drinking water was made public. In particular, we find that age at marriage for males reduced by 2.6 months (a reduction of 1%) and for females it reduced by 1.4 months (a reduction of 0.67%) in arsenic affected areas in response to information campaign. We also find that bride price reduced by almost 64% in arsenic contaminated areas in response to the information campaign. These results survive a battery of robustness checks.

We also complement our results related to age at marriage by using other demographic surveys such as Demographic Health Surveys for Bangladesh (BDHS) which uses rich spatial information about arsenic contamination of drinking water by relying on GPS coordinates of the interviewed cluster. Additionally, we provide further support for our results by matching arsenic to non-arsenic sub-districts (using Integrated Public Microdata Series - IPUMS data) based on their demographic and other characteristics and then comparing changes in marriage market outcomes between these two groups. Results using these alternate data-set confirm the earlier results: the effect of information campaign on age at marriage is found to be negative with a even larger magnitude.

Our results are consistent with the hypothesis that individuals who live in arsenic areas learn about their likelihood of developing skin lesions and other health problems due to information campaign and tend to get married earlier in order to avoid discovery of symptoms of diseases related to arsenic exposure like skin lesions, cancer etc. (which in an alternative scenario, where they do get discovered would make their chances of finding a mate quite difficult). Similarly in case of bride price, public knowledge about arsenic contamination in an area generates concerns about beauty and fertility of a prospective female match leading to a lower bride price which gets offered. While our results are consistent with these

[5]Period of analysis for bride price is 1990 to 2010.

explanations, a more detailed investigation of these mechanism is outside the scope of this paper due to limitations of data.

Our paper contributes to the literature of how information campaigns about health cause behavior changes in health and other socio-economic and demographic outcomes. Impact of information campaigns, aimed at health, that cause changes in how people behave, in health seeking behavior and other socio-economic dimensions, have been studied in other contexts. Oster et al. (2013) find that in case of Huntington disease which is a hereditary disease with ramifications on life expectancy, individuals who are not tested for the disease are optimistic about their health and make decisions as if they weren't diagnosed with Huntington. However, individuals with confirmed diagnoses (to whom true picture about their health status has been revealed) behave differently. Goldstein et al. (2008) find that HIV-positive mothers who learn their status are more likely to receive medication to prevent transmission to their children. The effect of provision of information on behaviour has also been documented in other studies like Altman and Traxler (2014) (reminders about dental health increase check-up appointments with dentists), Calzo- lari and Nardotto (2014) (reminders induce gym users to increase physical activity), Alaii et al. (2003); Habluetzel et al. (1997); Binka et al. (1996) (information about benefits of using bed-nets to prevent malaria increased usage of bed-nets), Cairncross et al. (2005); Curtis et al. (2001); Luby et al.(2010) (dissemination of information about importance of hand washing to reduce infections increased frequency of hand washing).

In the case of Bangladesh, the literature of the impact of the information campaign we are focusing on is limited to Keskin et al. (2017) that looks at how mothers in Bangladesh increased breastfeeding duration for infants in response to the information about water quality. However none of the studies look at how such information campaigns affect the marriage market. In particular, this paper is the first to examine the causal effect of information about health outcomes related to poor water quality in the neighbourhood on marriage market outcomes. In the process, our paper links information about local disease environment to marriage market outcomes.

Given that we look at the impact of the information campaign on marriage market outcomes, it is important to acknowledge that individuals may already have responded to visible symptoms of arsenic poisoning without knowing the cause. Literature has documented effect of arsenic contamination on longevity for adults, beauty and fertility for females (Hassan et al. (2005); Milton et al. (2005); Rahman et al. (2007); Sohel et al. (2009); Argos et al. (2010)). All of these factors play an important role in the marriage market (Buss (1989)). An information campaign can make these aspects more salient by removing the uncertainty on cause and may therefore still have an effect. However, the pre-existing recognition of some these effects does imply that the estimated impacts of the information campaign in our paper are likely to be an underestimate.

The paper is organized as follows, the following section provides background on water quality and related policies undertaken in Bangladesh. Section 3 describes various data sources that we use in our analysis. Section 4 describes our empirical strategy and Section 5 presents the corresponding results. Section 6 provides a discussion about the patterns we observe for marriage market and Section 7 concludes.

2 Background

2.1 Effects of Exposure to Arsenic

Exposure to arsenic causes many health problems which have been explored in bio-medical literature. Skin lesions are among the first few symptoms of arsenic poisoning, a longitudinal study (Ahsan et al., (2006)) has shown that higher dosage of arsenic is positively associated with higher probability of appearance of skin lesions. In addition to skin lesions, other studies have found that risk of all-cause and chronic-disease mortality is higher for people exposed to dangerous level of arsenic (Argos et al. (2010) & Sohel et al. (2009)). Adverse pregnancy outcomes like still birth, spontaneous abortion have also been linked to arsenic exposure (Milton et al. (2005); Rahman et al. (2006)). Various types of cancers (skin, kidney, bladder) and adverse effect on mental health (Chowdhury et al. (2016)) are among other devastating effects of arsenic exposure. The arsenic problem is not just a health problem but a social problem as well. The physical manifestation of early symptoms of arsenicosis like skin lesions is more than just a health effect. Given its visibility on the body, it has larger manifestations on the marriage market, as it is linked to the "beauty" of a person. Due to lack of information and illiteracy skin lesions are often confused with leprosy, which is considered a contagious killer by rural people. The early symptoms of arsenicosis which includes formation of black spots and warts thus leads to ostracism and social isolation (Alam et al. (2002)). Additionally, beyond the social problem arsenic exposure has also been found to have implications in the labour market in form of reduced labour supply (Carson et al. (2010)).

2.2 Arsenic in Water: Bangladesh

During 1970s Bangladesh was struggling with high disease burden due to water borne diseases related to surface water usage. Soon millions of tubewells were dug all across the country. In rural areas the switch to groundwater source was almost universal with almost 95% of households using tubewells (Caldwell et al. (2003)). A nation wide survey conducted by British Geological Survey (BGS) in 1999 found that most of the tubewells being used were of shallow depth and in few areas these tubewells contained dangerous levels of arsenic (shallow tubewells have a greater likelihood of containing arsenic)[6]. The BGS and Department of Public Health Engineering of Bangladesh (DPHE) report (2001) estimated that around 35 Million people were inadvertently exposed to harmful levels of arsenic by sourcing water from tubewells. Smith et al. (2000) also documents that between 35-77 Million people were exposed to arsenic due to contaminated tubewells usage.

[6] Arsenic was first detected in groundwater in 1985, however no comprehensive study ensued after this discovery until late 1990s, that is when British Geological Survey (BGS) conducted a nation wide survey of tubewells.

2.3 Intervention: Information Campaign on Arsenic

The Bangladesh government (Department of Public Health Engineering(DPHE)) along with UNICEF and Non-Profit Organisations implemented a water quality information campaign between 1999 and 2005 called the Bangladesh Arsenic Mitigation and Water Supply Programme (BAMWSP). The campaign didn't pick up any steam during the initial 2.5 years (World Bank (2007) - project completion report) with major roll-out happening during later years especially after 2002. In a paper which looks at the impact of same intervention on a different outcome (Keskin et al. (2017) uses 2002 as the cut-off for treatment period. The campaign had few distinct features: it disseminated information about contamination of tubewells by color coding tubewells into red (unsafe) and green (safe) categories; it also suggested mitigation strategies such as shifting to a safer well in the neighbourhood or well-sharing and finally it informed people about harmful health effects of arsenic exposure via public forums.

The constant visual reminders along with negative information about adverse health effects of arsenic exposure did have an impact on the intended population. The information campaign was quite effective in terms of generating awareness in terms of people reporting that they had heard of arsenic after information campaign and creating awareness about symptoms of arsenicosis after the information campaign (Keskin at al., 2017). Other papers (Chen et al. (2007); Opar et al. (2007); Bennear et al. (2013); Balasubramanya et al. (2014) etc) also find that people did switch to safer sources of water after the information campaign.

3 Data

In order to analyse the causal impact of information campaign on marriage market outcomes our analysis combines arsenic contamination information with the demographic data. We describe below our various sources of data:

3.1 Contamination data

British Geological Survey (BGS)- 1999

We source data on arsenic contamination, as present in 1999 from British Geological Survey (BGS) conducted in 1998-1999. In this survey, BGS tested a geographically representative sample of 3534 wells all across the country and also recorded the GPS coordinates of the tubewells along with other details such as the depth of the tubewell and the year in which it was constructed. Bangladesh has 64 districts in total but BGS sampled only 61 of them. Hence our study is restricted to these 61 districts. This survey found that more than a quarter of total sampled tubewells contained harmful level of arsenic (Bangladesh Government recognizes >50ug/litre as the dangerous level of arsenic contamination, however WHO prescribes >10ug/litre as the dangerous cutoff).

We categorize sub-districts into arsenic and non-arsenic affected group using mean arsenic contamination for each sub-district. This is calculated by averaging arsenic contamination level for all tubewells located in a sub-district[7]. To decide the threshold level of contamination that may be deemed dangerous, we use the value of 50ug/litre as recognized by the Bangladesh government. If the mean arsenic figure for a sub-district was above >50ug/litre then we code it as an arsenic affected sub-district and a sub-district with contamination <50ug/litre is coded as non-arsenic affected sub-district. Figure 1 shows the geographical variation in arsenic contamination at the sub-district level. The light shaded (green) sub-districts represent non-arsenic affected sub-districts (i.e. arsenic contamination was lower than 50ug/litre) while the dark shaded (red) show arsenic affected sub-districts. As is evident from the figure, contaminated tubewells are more likely to be in the south.

[7] On an average this survey tested 8 tubewells in each sub-district.

3.2 Demographic data

3.2.1 Long Census Survey Data - 2011

The individual and household data are obtained from Long Census Survey Data which was conducted by Bangladesh Bureau of Statistics (BBS) in March 2011. This dataset contains details about place of residence (sub-district), duration of residence in the current district, type of construction of house, education of household members and details related to nuptiality which includes our main variable of interest, that is the age at which an individual first got married. While the 2011 census surveyed 167,293 households across all 64 districts of Bangladesh, for the purpose of analysis in this paper, we focus our attention on the rural sample where the tubewell use was almost universal before the implementation of information campaign (Caldwell et al. (2003)). We combined the census data with arsenic contamination data (BGS data) to sub-districts to their arsenic contamination status. Our final estimation sample comprises of married individuals (75243 males and 86985 females) from 420 sub-districts in Bangladesh. Out of the total 420 sub-districts, 134 (around 32 percent) were found to have arsenic levels above >50ug/litre. It should also be noted that we don't observe female's original household, that is the household they were born in. We only observe females in the households they got married into, and for this reason we primarily focus on males and suggest interpretation of analysis for females to be done with caution [8].

We provide summary statistics for the male and female estimation sample in Table 1 (panels A and B). A greater proportion of both males and females are literate in arsenic areas in comparison to non-arsenic areas [9] In arsenic areas, a greater proportion of individuals (both males and females) belong to households who own assets (land and house). However, this does not mean that they are necessarily richer as only a small proportion of these households have cemented walls. Most of individuals in both arsenic and non-arsenic have

[8] In discussion we describe how males from arsenic areas get matched to females from arsenic areas in marriage market, which mitigate some concerns when interpreting results for females.

[9] Table 1 doesn't show statistical difference between the two groups (arsenic and non-arsenic), however differences are significant for all variables in both male and female sample.

their religion as Islam. The difference in characteristics in individuals imply that we cannot compare outcomes across arsenic and non-arsenic regions. Hence as we describe later, we resort to looking at changes overtime.

We describe next our outcome variable of interest, age at marriage. When we plot age at which males and females get married in Figure 2a), we observe that males in Bangladesh tend to get married earlier than the world average with mean age at first marriage being 24 years with 70 percent of males being married by 25 years of age (70 percent of females were married by the age of 18.3 years, Figure 2b). In Figure 3 we plot mean age at marriage overtime for arsenic and non-arsenic areas. Not surprisingly, we observe that individuals tend to get married later in arsenic affected areas (in comparison to non-arsenic areas). That arsenic contamination may have a link with marriage market is suggested by this observation. We observe that the mean age of marriage is higher for both males and females in arsenic affected areas. The age at marriage has been decreasing for males overtime (Figure 3a) while it has been increasing for females (Figure 3b). The gap (or difference) in age at marriage between arsenic and non-arsenic areas has narrowed especially after the implementation of information campaign (represented by green solid vertical line for year 2002). This relationship is shown in Figures 3c) and 3d) where we plot the difference in age at marriage between arsenic and non-arsenic areas overtime for males and females. We observe that the difference in age at marriage is lower in the post treatment period that is after the implementation of information campaign.

3.2.2 Primary Survey

For our bride price analysis we rely on Palli Karma-Sahayak Foundation (PKSF) Survey. This survey contained modules on marriage, divorce, bride price and dowry. It was conducted by researchers from University of Sydney in December 2010-January 2011. We combine this survey with contamination data (from BGS) to arrive at our final estimation sample for the period of study (1990-2010)[10], with information on 875 marriages from 20 sub-districts. 9 out of these 20 sub-districts were arsenic affected (arsenic contamination $> 50\mathtt{ug/litre}$). This dataset contains detailed information about marriage unions which includes year of marriage, the amount of dowry paid, bride price agreed upon, age of bride and groom, education of bride groom and income status of bride and groom's family at the time of marriage. In Table 1, Panel C we summarize our data from this survey, education of both bride and groom is more for couples belonging to the arsenic group (the difference is statistically significant for when compared with non-arsenic group). The age at which brides and grooms get married is not statistically different between arsenic and non-arsenic group. However, the age at marriage is particularly low for females, with mean age at marriage being 16.75 years. We also observe that a small proportion of marriages are those where the groom was chosen by a bride rather than by her family. Lastly the mean bride price amount is close to 53300 taka (difference between arsenic and non-arsenic group is not statistically significant) and the mean dowry amount is 22700 taka. The mean dowry in non-arsenic areas is around 20000 taka while it is much higher in arsenic areas (dowry amount is 22 percent higher in arsenic affected areas, the difference is statistically significant between the two groups). In Table 2 we provide summary about mean bride price before and after the intervention, we observe that mean bride price increased from 42611 taka in pre-treatment period to 66619 taka in non-arsenic sub-districts (an increase of 56 percent). However the arsenic affected areas saw an increase in bride price by only 30 percent [11].

4 Estimation Methodology

To estimate the effect of information campaign about arsenic contamination on age at marriage we follow a difference-in-difference approach (DID) using the following specification:

[10]Since the survey took place from December 2010 to January 2011, so we only observe marriages till year 2010

[11]We don't provide detailed year by year figures for this dataset as they are imprecisely estimated due to small sample size for each year

$$\mathsf{Y}_{ijdt} = \alpha + \theta \mathsf{Arsenic}_j * \mathsf{post}_t + \beta \mathsf{X}_{ijdt} + \gamma_j + \delta_t + \rho_d t + \varepsilon_{ijdt} \tag{1}$$

where Y_{ijdt} is the marriage market outcome - age at first marriage (for males or females) or log of bride price for an individual i who resides in sub-district j, belonging to district d and who belongs to marriage cohort t (i.e. who got married in year t). $\mathsf{Arsenic}_j$ is the dummy variable which takes value 1 if the mean arsenic contamination level is greater than 50ug/litre for a sub-district and 0 otherwise, post_t is the dummy variable which takes value 1 for post treatment period (2002-2011) and 0 for pre-treatment period (1990-2001). X_{ijdt} includes control for individual characteristics which matter in marriage market, since they differ for the two outcomes, we describe them in detail below. To control for time-invariant differences between sub-district and location-invariant differences between marriage cohorts we include sub-district fixed effects, γ_j and marriage cohort fixed effects, δ_t. We also control for district level trends, $\rho_d t$ to account for any underlying trends present at the district level. These fixed effects and trends account for any omitted variables at sub-district and year level; and any linearly time varying factors at the district level.

Our coefficient of interest is θ which captures the effect of information campaign. Essentially we compare difference in outcomes between marriage cohorts from arsenic affected sub-districts before and after the information campaign to the difference in outcomes in the same marriage cohorts from non-arsenic sub-districts. The individual level controls (X_{ijdt}) for age at marriage regressions based on Census data include religion being Islam, ownership of land, ownership of a house, literacy status and type of quality of house (that is whether it has cemented walls). For bride price regressions we rely on PKSF survey data and individual level controls here include education level of bride, difference in education level between bride and groom, age of bride, difference in age of bride and groom, a dummy for choosing the marital partner by themselves(rather than being chosen by family), a dummy for brides family being richer than groom' family and a dummy for groom's family being richer than bride' family (omitted category being two families belonging to same income class).

Identification

Our identifying assumption is that trends in age at marriage (and bride price) were not differential across arsenic and non-arsenic sub-districts, other than because of information campaign. The district specific trends also help in establishing validity of our results, as our estimates are identified off deviations from district trends. The possible threats to our identification thus comes from sub-district level trends which might be correlated with arsenic contamination.

For each of our DID specifications we test whether the parallel trend assumption is satisfied by conducting the following estimation exercise:

$$Y_{ijdt} = \alpha + \Sigma_{2001}^{1990}\lambda_t(Arsenic_j * I_t) + \beta X_{ijdt} + \gamma_j + \delta_t + \rho_d t + \varepsilon_{ijdt} \qquad (2)$$

where Y_{ijdt} is our outcome of interest: age at first marriage or bride price and j is the sub-district of residence. I_t is an indicator variable for each of the pre-treatment years. In presence of sub-district and year of marriage fixed effect, the interaction terms between year dummies and arsenic dummy reveal whether control and treatment group followed different trend overtime. We look for individual significance of all these interaction terms. If these terms are individually insignificant then that reveals that the parallel trend assumption is satisfied as the two groups followed similar trend in the pre-treatment period.

We estimate specification 2 for all our outcome variables for the pre-treatment period and present our results in a visual format in Figure 4 (panels a, b and c). We notice that for almost all our models we satisfy parallel trend assumption as all the interaction terms are insignificant. We also test for joint significance of interaction terms and reject it at 10 percent level for all our results (the only exception is the female sample where parallel trend assumption may not be satisfied as individual and joint significance of interaction terms is statistically different from zero). In Table 3 we test a slightly different version of specification 2 to test whether trends differed between arsenic and non-arsenic areas. We do this by replacing individual year interacted arsenic variables ($\Sigma_{2001}^{1990}(Arsenic_j * I_t)$ with a single variable for arsenic dummy interacted with a continuous trend variable $Arsenic * Trend$. An insignificant coefficient on this new variable reflects that trends were not different between the two groups before the intervention. We find in Table 3 (columns 1 to 3) that parallel trend assumption does hold true for all our outcome variables.

5 Results

5.1 Age at Marriage

In Table 4 we present results from specification 1. All columns include sub-district and marriage cohort (year of marriage) fixed effects. Columns 1 and 2 refer to results for the male sample while columns 3 and 4 present results for the female sample. Our main results in columns 2 and 4 control for district level trends. The coefficient of $\mathsf{Arsenic} * \mathsf{Post}$ Table 2 shows that both males and females from arsenic affected areas are get married earlier after implementation of information campaign. For the male sample the reduction in age at marriage is 0.22 years (around 2.6 months), while for the female sample the magnitude is slightly lower, around 0.12 years (1.4 months)[12] These results translate into a reduction of age at marriage after information campaign by around 1 percent for males and by 0.67 percent for females. Results on other controls reveal that being literate increases the age at marriage, also religion of an individual being Islam is associated with getting married earlier. Ownership of assets is also associated with higher age at marriage.

5.2 Bride Price

Table 5 presents our results for bride price. Our dependent variable here is log of bride price, and thus coefficients of interaction variables can be interpreted in terms of the percentage decline in bride price for arsenic areas in the treatment period. Column 1 presents results with no district level trends and we find a 34% decrease in bride price (however this effect

[12] The results on female's age at marriage should be interpreted with some caution as we don't observe them in their original place of residence, however assuming that in equilibrium marriage matches happen between similar areas, i.e. males from arsenic areas get matched to females from arsenic areas and males from non-arsenic areas get matched to females from non-arsenic areas.

is not found to be significant). Column 2 accounts for district level trends and we find that arsenic areas witnessed a 64.4% decrease in bride price post information campaign. We also observe that bride price is positively associated with education level of the bride while it is negatively associated with groom's education level (not significant). Additionally getting married earlier is associated with lower bride price[13].

5.3 Robustness Checks

5.3.1 Matched sample analysis

We complement our findings from Census data by using additional demographic surveys. The arsenic affected sub-districts can be systematically different from non-arsenic sub-districts, so we additionally wanted to analyze whether arsenic affected areas which are *similar* to non-arsenic areas exhibit similar marriage patterns after information campaign. We conduct this analysis by matching arsenic affected sub-districts with unaffected sub-districts by using sub-district level characteristics from Integrated Public Microdata Series (IPUMS) Census Data for 2001. The IPUMS Survey is a huge census and for matching purposes we use pre-treatment characteristics which come from 2001 IPUMS survey which covered over 12 million individuals (2.6 million households) residing in all 64 districts of Bangladesh. This dataset was also compiled by Bangladesh Bureau of Statistics (BBS) and has individual and household information available at sub-district level. We use 19 such variables for our matching exercise. In particular, we use details about employment, education, household characteristics, sex-ratios for the unmarried population and sex ratio for children below one year of age, collapsed at sub-district level for our matching purpose [14]. Table A1 compares the mean 2001 characteristics of arsenic (treated) and non-arsenic (control) sub-districts. Column 1 and 2 reveal that these two groups are considerably different from each other. To

[13] Bride price results rely on a small small due to which many estimates have large standard errors because of which we observe insignificant results.

[14] We don't use IPUMS data for our main analysis as it doesn't contain information about our main dependent variable i.e. the age at which individuals got married.

address this concern we follow the matching procedure by using sub-district level characteristics from IPUMS data. Our matching exercise results in 79 arsenic affected sub-districts getting matched to unique 79 non-arsenic sub-districts, and column 6 in table 2 shows that the matched treated and matched control sub-districts are statistically similar in terms of various demographic and non-demographic characteristics [15].

To establish consistency of our original results we use a sub-district-year panel for all 158 matched sub-districts where the outcomes and control variables were aggregated at sub-district-year level by constructing means. In Table 6, we observe that for both male and female sub-sample mean age at marriage reduced after information campaign. The magnitude is higher than our original results with age at marriage getting reduced for males by 1.7 percent and for females by 1 percent.

5.3.2 Local contamination measures

We next conduct analysis for age at first marriage for males using Bangladesh Demographic Health Surveys (BDHS). We turn to BDHS data as it provides rich spatial information about contamination of drinking water[16]. We use four rounds of these surveys conducted in 1999, 2004, 2007 and 2011. These surveys were conducted by National Institute of Population Research and Training (NIPORT) and followed identical questionnaires overtime. For each year around 10000 households were sampled from 350 clusters, except for year 2011 when 17000 households were sampled from 600 clusters. BDHS collected GPS information for all sampled clusters. For our rural sample we focus on the men's questionnaire which contains questions regarding their age at marriage, education and assets. Our final estimation sample has around 5000 men sampled from 1067 clusters. We use this dataset to replicate and

[15] The probit regression for generating propensity scores for matching is presented in appendix Table A2 and Figure A1 in appendix plots the number of treated and control sub-districts matched over p-score. We were able to match 410 sub-districts out of a total of 420 sub-districts which are included in our main analysis using Census data of 2011

[16] We don't use female sample here as we don't observe the original village from which she got married, we only observe females in the households they got married to.

complement the results which we find using Census data.

We are able to construct arsenic contamination figures for each individual sampled cluster in BDHS data. Arsenic contamination for a cluster was calculated by averaging contamination level for all tubewells lying in the 5 mile radius around the cluster location. This is a "local" level of contamination with clusters with mean contamination figure >50ug/litre being classified as arsenic affected and those with <50ug/litre being classified as non-arsenic affected cluster. This local measure of contamination in BDHS data provides important information about contamination in the *neighbourhood* of a household which reflects the local visual clues (in form of red and green painted tubewells and people getting affected by arsenicosis symptoms) which people observe in their locality.

We now present our results using BDHS data which uses rich spatial information about arsenic contamination at cluster level. Table 7 presents results for males, both column 1 and 2 have cluster and marriage cohort fixed effects. Column 2 additionally controls for district level trends. Analysis using BDHS data shows that the our variable of interest, that is interaction term $\mathsf{Arsenic} * \mathsf{Post}$ is negative and significant in both the columns[17], with the order of magnitude being 0.84 years (10 months) which is a reduction in age of marriage by as much as 3.5 percent. This is much larger in comparison to the effect seen using census data[18]. We find similar results as before for our control variables as well, with education and religion significantly affecting age at marriage.

5.3.3 Marriage pool before and after information campaign

The results which we observe can potentially arise if the information campaign affected the marriage pool (that is mix of males and females in the marriageable age group). We explicitly test for this by looking at the sexratio at sub-district level before and after the implementation of information campaign. We use IPUMS Census 2001 data and Census

[17] Coefficient of interaction term in column 2 of table 3 is significant at 6% level.

[18] In Appendix Figure 2 we provide evidence for parallel trend assumption for models discussed in sections 5.3.1 and 5.3.2.

2011 data to create sub-district level mean sexratio between males and females in age group 15 to 25 years for years 2001 (before intervention) and 2011 (after intervention)[19]. Table 8 presents our results for this analysis, we observe (in columns 1 to 3 which correspond to different specifications) that sexratio did not change due to information campaign. This essentially cements our original theory that the effects we see in marriage market of information campaign are driven by behavioural factors rather than changes in marriage pool.

5.3.4 Shorter analysis window

Our original analysis was based on marriages which took place between 1990 to 2011, we shorten this analysis window to span from 1996 to 2006 to test whether our original results still hold. In Table 9 (columns 1 and 2), we observe that for the male sample estimate remains almost unchanged, while for the female sample although the estimate still retains the negative sign but the magnitude is much smaller and it is no-longer significant.

5.3.5 Placebo tests

Lastly, we conduct two placebo tests to establish validity of our results. First we replicate our analysis for urban sample where we don't expect to see any effect of the information campaign as the urban places don't rely on tubewells as a source of water (which were the focus of information campaign). In Table 9 (column 3 and 4), we observe that information campaign had no effect on age at marriage for both males and females which is in line with our expectation.

We also conduct another placebo test by randomly shuffling the arsenic contamination status of sub-districts. This basically randomly assigns sub-districts into new treatment and control groups. We then estimate the coefficient of (Arsenic * Post) variable by following our original specification. We repeat this exercise 1000 times, each time randomly assigning treatment and control status to sub-districts. This gives rise to a distribution of θ based on 1000 simulations. In Figure 5, we observe that for both male and female sample true

[19] 410 sub-districts are matched between IPUMS Census 2001 data and Census 2011 data.

estimate of θ lies beyond the 95% confidence interval which shows that the results we observe are not there by chance, which strengthens the validity of our results.

5.3.6 Other contamination measures and additional data on bride price

In our analysis we have categorized sub-districts into arsenic and non-arsenic sub-districts using $50ug/litre$ as the cut-off for dangerous level of arsenic in drinking water. We now provide additional results where alternate contamination variables are used for analysis. Table 9 (age at marriage results) and Table 10 (bride price results in columns 1 and 2) presents results based on two alternate measures of contamination. $MeanArsenic$ is a continuous measure of arsenic contamination while $Arsenic10$ is a dummy variable which takes value 1 if mean arsenic contamination in a sub-district is above $10ug/litre$ which is the WHO safety standard for arsenic in water. In Table 9 columns 1 and 2 we present results for the male sample while in columns 3 and 4 we present results for the female sample, Table 10 (columns 1 and 2) provides results for bride price. We observe that for both age at marriage and bride price the effect of information campaign based on alternate measures of arsenic contamination is negative. The effect of information campaign using continuous measure of arsenic contamination is significant for both age at marriage (males and females) and bride price. When we use $10ug/litre$ cut-off for defining arsenic contamination then we observe that the point estimate (of $Arsenic10 * Post$) is negative for all outcome variables. The estimate for age at marriage for males is -0.18 which is close of the original estimate of -0.22, while the estimate for females is -0.07 (not significant). The corresponding bride price estimate using $10ug/litre$ cut-off is negative and smaller in magnitude (in comparison to results based on $50ug/lire$ cut-off) but not significant.

Our bride price analysis uses only PKSF survey uptil now. Since the sample is small for bride price analysis so we incorporate additional data on more marriages using Bangladesh Rural Urban Linkage (BRUL) Survey. This survey was conducted by International Food Policy Research Institution between December 2004 and January 2005 and it used identical modules on marriage, divorce, bride price and dowry. The combined sample of PKSF and

BRULS for marriages between 1990 and 2010 contains 1699 marriages from 80 sub-districts. In Table 10 column 3 we replicate our original bride price result using this expanded dataset. We find that bride price reduced by 60% in arsenic affected areas after information campaign, which is pretty close to our original estimate of 64% reduction in bride price post intervention.

5.4 Heterogeneity

The distribution of arsenic contamination is not uniform across all sub-districts in Bangladesh. There were few sub-districts which also had high mean level contamination that is arsenic level greater than 100ug/litre. We expect that the intensity information campaign could have been higher for these sub-districts as the arsenic contamination problem was extremely intense for these sub-districts. Hence the behaviour change expected in these sub-districts could also be different. We analyze whether the effect of information campaign was more pronounced in these sub-districts. We do this by comparing individuals from sub-districts with mean arsenic contamination being less than 50/ug/litre (same as our old control group) to sub-districts with mean arsenic contamination greater than 100/ug/litre (new treatment group). In Table 12, we present our results based on this analysis. We find that for both males and females the magnitude of the effect of information campaign on age at marriage is larger than our original results, that is the age at marriage reduces by 1.2% and 1.3% for males and females respectively in response to information campaign.

6 Discussion and Conclusion

The patterns that we observe can be explained by the matching mechanism in marriage market. Suppose in the marriage market the matching is based on sorting which is driven by preferences over prospective match's traits like beauty, life-expectancy, fertility, health standard and income generating prospects. Lets assume that all of these traits can be subsumed in a single index value over which the sorting for partners takes place. A higher value of the index represents a more desirable partner. Now we know that for an individual belonging to an arsenic affected area the expected value of this index will take a lower value than the one for an individual who hails from a non-arsenic area. A higher index value is desired more in the market, hence in a stable match the highest ranked woman gets matched to the highest ranked man, the second highest ranked woman gets matched to second highest ranked man and so on. The Gale-Shapley algorithm (Gale and Shapley (1962) will thus give an equilibrium outcome where high ranked females from non-arsenic areas get matched to high ranked males from non-arsenic areas and low ranked females from arsenic areas get matched to low ranked males from arsenic areas. We call this arsenic-arsenic matching for the purpose of our discussion.

We hypothesize that the information campaign made the contamination information public, people now knew their own likelihood of being affected by arsenic poisoning symptoms. Also people from other areas (who are looking for a match) could take in the visual information in form of red and green painted tubewells and could ascertain probabilistically the chance of a prospective match being affected by arsenicosis symptoms. The information campaign thus complemented the index value in the marriage market. For an individual from an arsenic area given arsenic-arsenic matching and faced with a possibility of low life-expectancy and higher probability of developing skin lesions, which adversely affects his prospects of finding a good match in marriage market leads him to get married earlier before his symptoms get discovered. Following a similar argument (for bride price), given arsenic-arsenic matching groom's family is aware about their matching with a female from arsenic affected area. Since the social structure is such that stigma attached to a female with skin lesions (and possible low fertility) is way more than the one attached to a male, hence females disproportionately suffer more than males in the marriage market. Thus concerns about prospective bride's beauty and possible low fertility coupled with asymmetrical costs associated with females developing skin lesions leads to an agreed bride price which is lower [20]. While this mechanism is consistent with our evidence, we are unable to provide evidence

[20] Dowry is also an important aspect of marriages in Bangladesh, however the effect of information campaign on dowry can be quite ambiguous. Dowry unlike bride price is an amount which is exchanged at the

for the same in this paper.

The behavioural change induced by an information campaign has mainly been explored in the health domain (for example: increase in breastfeeding time and switch to safer sources of water for concerns related to health). We have gone a step further and shown that the information campaign had a spill-over effect in marriage market as well. Males and females (and their respective households) in marriage market learned about their likelihood of being affected by arsenicosis symptoms in the future and reacted to this information by reducing their age at marriage. Essentially people in marriage market were entering into marriage unions earlier than before in order to avoid an adverse scenario of not finding match if they develop arsenicosis symptoms in the future. The bride price also suffered a dampening effect (due to information campaign) owing to less valued brides in the market, where less value was attributed to them due to information garnered about their adverse fertility and beauty outcomes in the future. The campaign owes its success in terms of generating awareness to its unique design. Constant visual reminders, multiple strategies to avoid contaminated water and public forums for information disbursal seem to have worked in Bangladesh. We find evidence that campaigns with negative information, that is one which informs people about adverse outcomes, in this case warns them of ill-health tend to have an impact on the intended population by motivating them to change their behaviour in the marriage market.

time of marriage when the groom may not have started showing symptoms of arsenicosis. Given arsenic-arsenic matching the brides family know that they get matched to a groom from arsenic area with greater probability of being affected with arsenicosis which ideally drives down the dowry amount. However if we look at things from the supply side then the set of possible matches for a bride, is now also smaller since it now excludes non-arsenic grooms, this drives up the dowry amount. Hence we get an ambiguous result (insignificant result) for dowry.

References

Ahmad, J., Goldar, B., Misra, S., and Jakariya, M. (2003). Willingness to pay for arsenic-free, safe drinking water in bangladesh. *World Bank Water and Sanitation Program-South Asia*.

Ahmad, S. A., Sayed, M., Barua, S., Khan, M. H., Faruquee, M., Jalil, A., Hadi, S. A., and Talukder, H. K. (2001). Arsenic in drinking water and pregnancy outcomes. *Environmental health perspectives*, 109(6):629–631.

Ahsan, H., Chen, Y., Parvez, F., Zablotska, L., Argos, M., Hussain, I., Momotaj, H., Levy, D., Cheng, Z., Slavkovich, V., et al. (2006). Arsenic exposure from drinking water and risk of premalignant skin lesions in bangladesh: baseline results from the health effects of arsenic longitudinal study. *American Journal of Epidemiology*, 163(12):1138–1148.

Ahsan, H., Perrin, M., Rahman, A., Parvez, F., Stute, M., Zheng, Y., Milton, A. H., Brandt-Rauf, P., Van Geen, A., and Graziano, J. (2000). Associations between drinking water and urinary arsenic levels and skin lesions in bangladesh. *Journal of occupational and environmental medicine*, 42(12):1195–1201.

Alaii, J. A., Van Den Borne, H., Kachur, S. P., Mwenesi, H., Vulule, J. M., Hawley, W. A., Meltzer, M. I., Nahlen, B. L., and Phillips-Howard, P. A. (2003). Perceptions of bed nets and malaria prevention before and after a randomized controlled trial of permethrin-treated bed nets in western kenya. *The American journal of tropical medicine and hygiene*, 68(4_suppl):142–148.

Alam, M., Allinson, G., Stagnitti, F., Tanaka, A., and Westbrooke, M. (2002). Arsenic contamination in bangladesh groundwater: a major environmental and social disaster. *International Journal of Environmental Health Research*, 12(3):235–253.

Altmann, S. and Traxler, C. (2014). Nudges at the dentist. *European Economic Review*, 72:19–38.

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., Islam, T., Ahmed, A., Rakibuz-Zaman, M., Hasan, R., et al. (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in bangladesh (heals): a prospective cohort study. *The Lancet*, 376(9737):252–258.

Balasubramanya, S., Pfaff, A., Bennear, L., Tarozzi, A., Ahmed, K. M., Schoenfeld, A., and van Geen, A. (2014). Evolution of households' responses to the groundwater arsenic crisis in bangladesh: information on environmental health risks can have increasing behavioral impact over time. *Environment and development economics*, 19(5):631–647.

Bank, W. (2007). Implementation completion and results report (ida-31240 swtz- 21082) on a credit in the amount of sdr 24.2 million (usd 44.4 million equivalent) to bangladesh for arsenic mitigation water supply.

Bennear, L., Tarozzi, A., Pfaff, A., Balasubramanya, S., Ahmed, K. M., and Van Geen, A. (2013). Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in bangladesh. *Journal of environmental economics and management*, 65(2):225–240.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G., Armah, G., Kajihara, B., Adiamah, J., and Smith, P. G. (1996). Impact of permethrin impregnated bednets on child mortality in kassena-nankana district, ghana: a randomized controlled trial. *Tropical Medicine & International Health*, 1(2):147–154.

Bjerk, D. (2009). Beauty vs. earnings: Gender differences in earnings and priorities over spousal characteristics in a matching model. *Journal of Economic Behavior & Organization*, 69(3):248–259.

Brinkel, J., Khan, M., and Kraemer, A. (2009). A systematic review of arsenic exposure and its social and mental health effects with special reference to bangladesh. *International journal of environmental research and public health*, 6(5):1609–1619.

Buchmann, N., Field, E. M., Glennerster, R., and Hussam, R. N. (2019). Throwing the baby out with the drinking water: Unintended consequences of arsenic mitigation efforts in bangladesh. *National Bureau of Economic Research*.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and brain sciences*, 12(1):1–14.

Cairncross, S., Shordt, K., Zacharia, S., and Govindan, B. K. (2005). What causes sustainable changes in hygiene behaviour? a cross-sectional study from kerala, india. *Social science & medicine*, 61(10):2212–2220.

Caldwell, B. K., Caldwell, J. C., Mitra, S., and Smith, W. (2003). Tubewells and arsenic in bangladesh: challenging a public health success story. *International Journal of Population Geography*, 9(1):23–38.

Calzolari, G. and Nardotto, M. (2011). Nudging with information: A randomized field experiment. *Available at SSRN 1924901*.

Carson, R. T., Koundouri, P., and Nauges, C. (2010). Arsenic mitigation in bangladesh: a household labor market approach. *American Journal of Agricultural Economics*, 93(2):407–414.

Chen, Y., van Geen, A., Graziano, J. H., Pfaff, A., Madajewicz, M., Parvez, F., Hussain, A. I., Slavkovich, V., Islam, T., and Ahsan, H. (2007). Reduction in urinary arsenic levels in response to arsenic mitigation efforts in araihazar, bangladesh. *Environmental Health Perspectives*, 115(6):917–923.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2012). Fatter attraction: anthropometric and socioeconomic matching on the marriage market. *Journal of Political Economy*, 120(4):659–695.

Chowdhury, S., Krause, A., and Zimmermann, K. F. (2016). Arsenic contamination of drinking water and mental health. *DEF-Discussion Papers on Development Policy*, (222).

Curtis, V., Kanki, B., Cousens, S., Diallo, I., Kpozehouen, A., Sangare, M., and Nikiema, M. (2001). Evidence of behaviour change following a hygiene promotion programme in burkina faso. *Bulletin of the World Health Organization*, 79:518–527.

de Walque, D. (2007). How does the impact of an hiv/aids information campaign vary with educational attainment? evidence from rural uganda. *Journal of Development Economics*, 2(84):686–714.

Duflo, E., Dupas, P., and Kremer, M. (2015). Education, hiv, and early fertility: Experimental evidence from kenya. *American Economic Review*, 105(9):2757–97.

Dupas, P. (2011a). Do teenagers respond to hiv risk information? evidence from a field experiment in kenya. *American Economic Journal: Applied Economics*, 3(1):1–34.

Dupas, P. (2011b). Health behavior in developing countries. *Annu. Rev. Econ.*, 3(1):425–449.

Elder Jr, G. H. (1969). Appearance and education in marriage mobility. *American Sociological Review*, pages 519–533.

Field, E. and Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy*, 116(5):881–930.

Flanagan, S. V., Johnston, R. B., and Zheng, Y. (2012). Arsenic in tube well water in bangladesh: health and economic impacts and implications for arsenic mitigation. *Bulletin of the World Health Organization*, 90:839–846.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

Gallagher, K. M. and Updegraff, J. A. (2011). Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. *Annals of behavioral medicine*, 43(1):101–116.

Goldstein, M., Zivin, J. G., Habyarimana, J., Pop-Eleches, C., and Thirumurthy, H. (2008). Health worker absence, hiv testing and behavioral change: Evidence from western kenya. *Unpublished Working Paper, Department of Economics, Columbia University.*

Habluetzel, A., Diallo, D., Esposito, F., Lamizana, L., Pagnonl, F., Lengeler, C., Traore, C., and Cousens, S. (1997). Do insecticide-treated curtains reduce all-cause child mortality in burkina faso? *Tropical medicine & international health*, 2(9):855–862.

Hassan, M. M., Atkins, P. J., and Dunn, C. E. (2005). Social implications of arsenic poisoning in bangladesh. *Social Science & Medicine*, 61(10):2201–2211.

Jakariya, M. (2007). *Arsenic in tubewell water of Bangladesh and approaches for sustainable mitigation.* PhD thesis, KTH.

Jalan, J., Somanathan, E., and Chaudhuri, S. (2009). Awareness and the demand for environmental quality: survey evidence on drinking water in urban india. *Environment and Development Economics*, 14(6):665–692.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual review of sociology*, 24(1):395–421.

Keskin, P., Shastry, G. K., and Willis, H. (2017). Water quality awareness and breastfeeding: evidence of health behavior change in bangladesh. *Review of Economics and Statistics*, 99(2):265–280.

Kinniburgh, D. and Smedley, P. (2001). Arsenic contamination of groundwater in bangladesh. *British Geological Survey and Department of Public Health Engineering (BGS and DPHE)*, page http://www.bgs.ac.uk/arsenic/bangladesh/.

Kirby, D. B., Baumler, E., Coyle, K. K., Basen-Engquist, K., Parcel, G. S., Harrist, R., and Banspach, S. W. (2004). The "safer choices" intervention: its impact on the sexual behaviors of different subgroups of high school students. *Journal of Adolescent Health*, 35(6):442–452.

Luby, S. P., Kadir, M. A., Yushuf Sharker, M., Yeasmin, F., Unicomb, L., and Sirajul Islam, M. (2010). A community-randomised controlled trial promoting waterless hand sanitizer and handwashing with soap, dhaka, bangladesh. *Tropical medicine & international health*, 15(12):1508–1516.

Madajewicz, M., Pfaff, A., Van Geen, A., Graziano, J., Hussein, I., Momotaj, H., Sylvi, R., and Ahsan, H. (2007). Can information alone change behavior? response to arsenic contamination of groundwater in bangladesh. *Journal of development Economics*, 84(2):731–754.

Marini, M. M. (1978). The transition to adulthood: Sex differences in educational attainment and age at marriage. *American sociological review*, pages 483–507.

Milton, A. H., Smith, W., Rahman, B., Hasan, Z., Kulsum, U., Dear, K., Rakibuddin, M., and Ali, A. (2005). Chronic arsenic exposure and adverse pregnancy outcomes in bangladesh. *Epidemiology*, pages 82–86.

Mukherjee, A., Sengupta, M. K., Hossain, M. A., Ahamed, S., Das, B., Nayak, B., Lodh, D., Rahman, M. M., and Chakraborti, D. (2006). Arsenic contamination in groundwater: a global perspective with emphasis on the asian scenario. *Journal of Health, Population and Nutrition*, pages 142–163.

Opar, A., Pfaff, A., Seddique, A., Ahmed, K., Graziano, J., and van Geen, A. (2007). Responses of 6500 households to arsenic mitigation in araihazar, bangladesh. *Health & place*, 13(1):164–172.

Oster, E., Shoulson, I., and Dorsey, E. (2013). Optimal expectations and limited medical testing: evidence from huntington disease. *American Economic Review*, 103(2):804–30.

Rahman, A., Vahter, M., Ekström, E.-C., Rahman, M., Golam Mustafa, A. H. M., Wahed, M. A., Yunus, M., and Persson, L.-Å. (2007). Association of arsenic exposure during pregnancy with fetal loss and infant death: a cohort study in bangladesh. *American journal of epidemiology*, 165(12):1389–1396.

Ravenscroft, P., Brammer, H., and Richards, K. (2009). *Arsenic pollution: a global synthesis*, volume 28. John Wiley & Sons.

Smith, A. H., Lingas, E. O., and Rahman, M. (2000). Contamination of drinking-water by arsenic in bangladesh: a public health emergency. *Bulletin of the World Health Organization*, 78:1093–1103.

Sohel, N., Persson, L. Å., Rahman, M., Streatfield, P. K., Yunus, M., Ekström, E.-C., and Vahter, M. (2009). Arsenic in drinking water and adult mortality: a population-based cohort study in rural bangladesh. *Epidemiology*, pages 824–830.

Spivey, C. (2010). Desperation or desire? the role of risk aversion in marriage. *Economic Inquiry*, 48(2):499–516.

UNICEF (2008). Arsenic mitigation in bangladesh.

UNICEF et al. (2007). Predicting the global extent of arsenic pollution of groundwater and its potential impact on human health. *New York. Report Prepared by Peter Ravenscroft, Consultant Hydrogeologist, Cambridge, UK.*

Van Geen, A., Ahsan, H., Horneman, A. H., Dhar, R. K., Zheng, Y., Hussain, I., Ahmed, K. M., Gelman, A., Stute, M., Simpson, H. J., et al. (2002). Promotion of well-switching to mitigate the current arsenic crisis in bangladesh. *Bulletin of the World Health Organization*, 80:732–737.

Tables and Figures

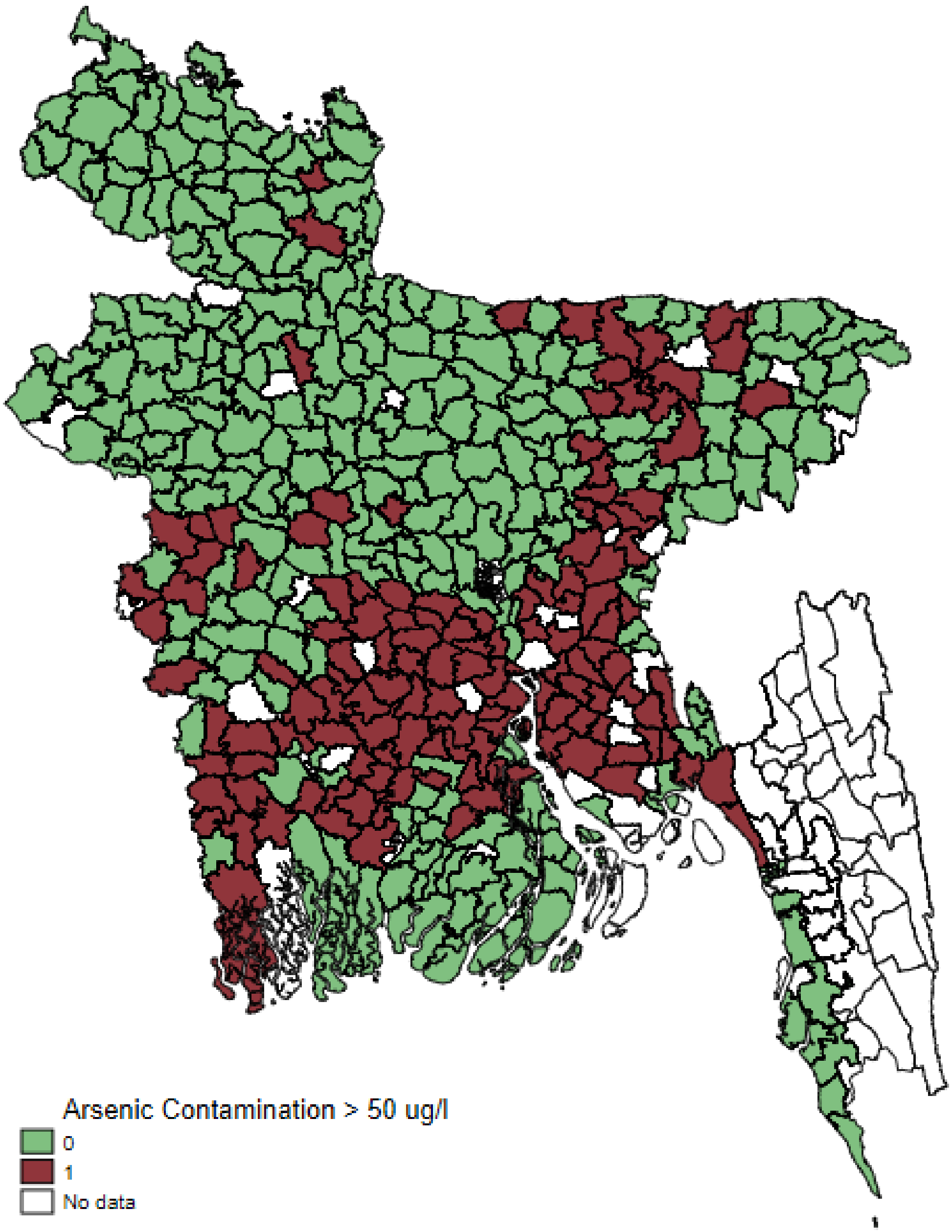
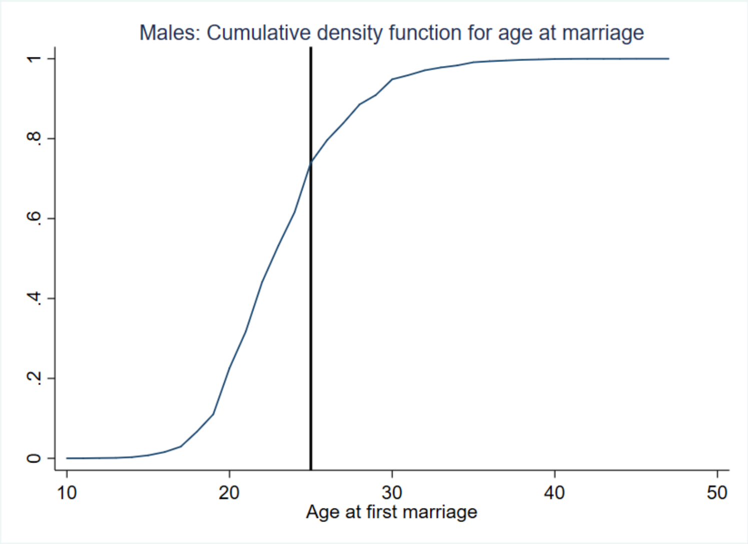
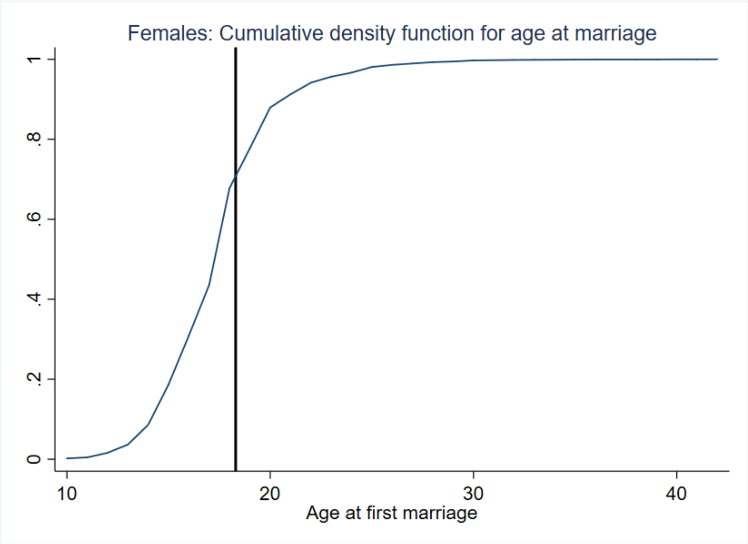


Figure 1: Arsenic Contamination Map for Bangladesh: Source of arsenic contamination data is British Geological Survey 1999.



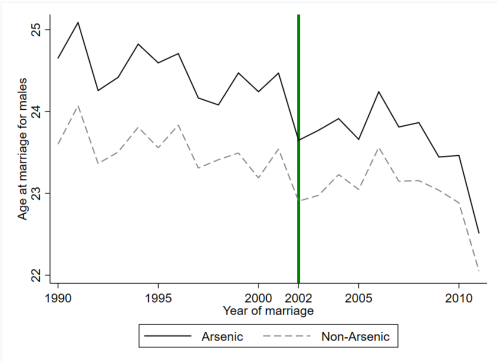
(a) Age at marriage for males



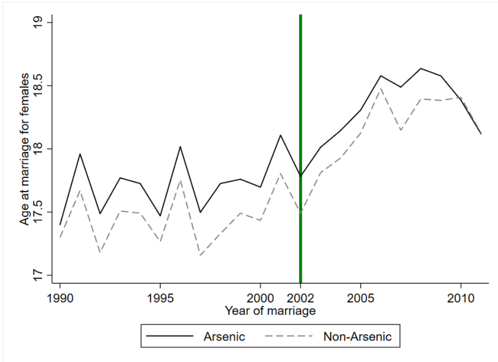
(b) Age at marriage for females

Figure 2: Cumulative Density Function for Age at Marriage for Males and Females.

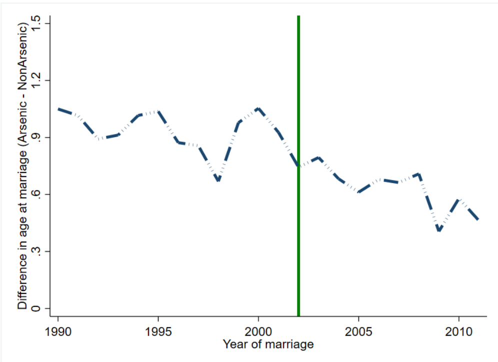
Source - Census Data 2011



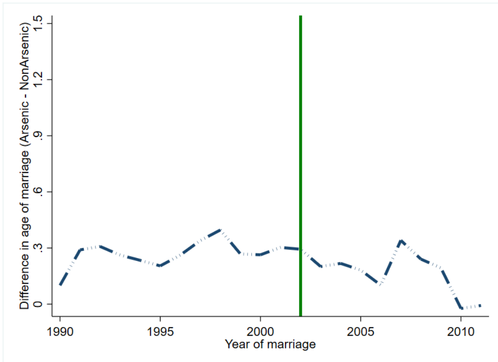
(a) Age at marriage for Males



(b) Age at marriage for Females

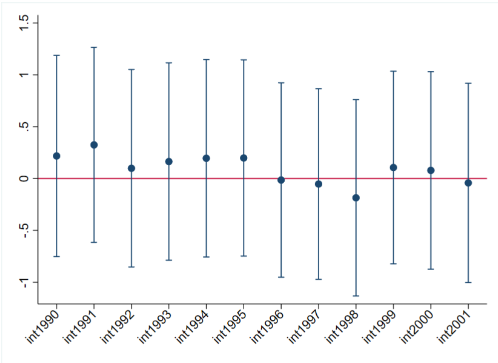


(c) Difference in Age at marriage for Males

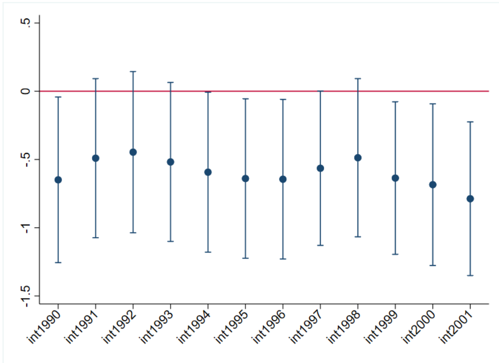


(d) Difference in Age at marriage for Females

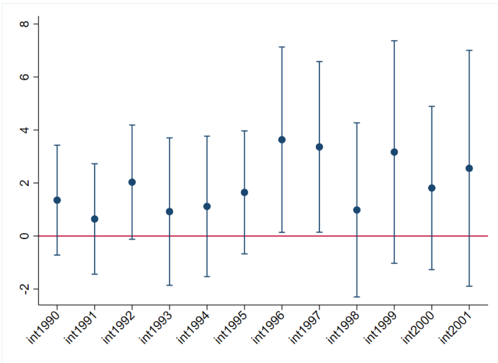
Figure 3: Age at marriage for males and females, differences in age of marriage between arsenic and non-arsenic areas over time using Census data for 2011.



(a) Age at marriage for males

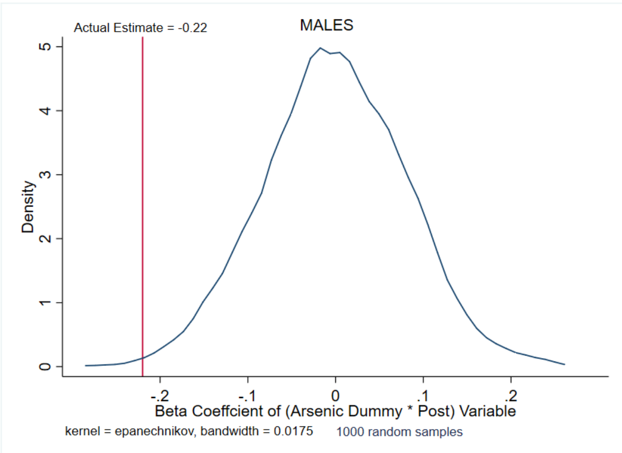


(b) Age at marriage for females

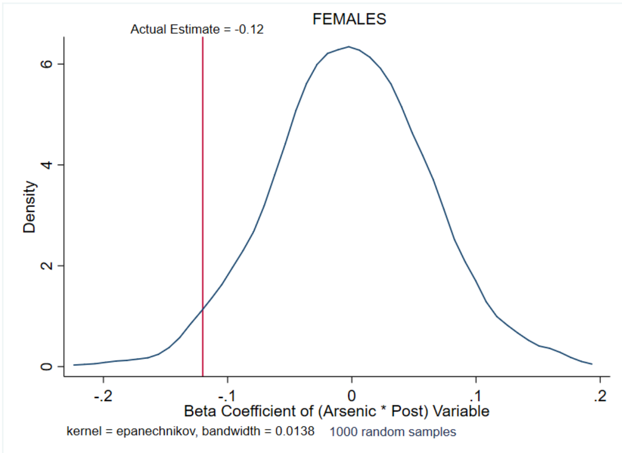


(c) Bride Price

Figure 4: Testing Parallel Assumption using specification 2: Plotting coefficients of interaction between year dummies and arsenic dummy for all pre-treatment years. Insignificance of coefficients points towards a similar trend being followed by the two groups in the pre-treatment period. The error bars are 95% confidence intervals, clustering at sub-district level.



(a) Males



(b) Females

Figure 5: Placebo Test: Distribution of θ using 1000 simulations which randomly assign arsenic contamination status of 1 (treatment) or 0 (control) to sub-districts.

Table 1: Summary Table

PANEL A. MALES

	All		*Non-Arsenic*		*Arsenic*	
variablename	Mean	SE	Mean	SE	Mean	SE
Literate (%)	0.51	0.002	0.5	0.002	0.54	0.003
Religion is Islam (%)	0.89	0.001	0.9	0.001	0.87	0.002
Owns Land (%)	0.92	0.001	0.91	0.001	0.94	0.001
Owns House (%)	0.89	0.001	0.88	0.001	0.91	0.002
House has cemeneted walls (%)	0.16	0.001	0.16	0.002	0.16	0.002
Age at marriage (years)	23.61	0.015	23.34	0.018	24.14	0.026
Observations	75243		50281		24962	

PANEL B. FEMALES

	All		*Non-Arsenic*		*Arsenic*	
	Mean	SE	Mean	SE	Mean	SE
Literate (%)	0.56	0.002	0.55	0.002	0.6	0.003
Religion is Islam (%)	0.89	0.001	0.9	0.001	0.88	0.002
Owns Land (%)	0.93	0.001	0.91	0.001	0.95	0.001
Owns House (%)	0.89	0.001	0.88	0.001	0.91	0.002
House has cemeneted walls (%)	0.17	0.001	0.17	0.002	0.16	0.002
Age at marriage (years)	17.88	0.01	17.8	0.012	18.03	0.017
Observations	86965		57243		29742	

PANEL C. PKSF Data

	All		*Non-Arsenic*		*Arsenic*	
	Mean	SE	Mean	SE	Mean	SE
Education bride (in years)	5.31	0.13	4.98	0.19	5.56	0.17
Education groom (in years)	5	0.14	4.74	0.21	5.19	0.18
Bride's age at the time of marriage	16.75	0.1	16.88	0.16	16.64	0.12
Groom's age at the time of marriage	23.59	0.15	23.59	0.22	23.6	0.2
Partner chosen by bride	0.08	0.01	0.08	0.01	0.08	0.01
Brides family richer	0.32	0.02	0.33	0.02	0.32	0.02
Groom's family richer	0.25	0.01	0.21	0.02	0.28	0.02
Mehr (in Taka)	53366	2230	53980	3236	52900	3060
Dowry (in Taka)	22709	1034	20168	1291	24642	1528
Observations	875		378		497	

Table 2: Average bride price (in Taka) : PKSF Data

	Pre-2002	**Post-2002**	**Observations**
Arsenic	45952	59600	497
	(4013)	(4569)	
NonArsenic	42611	66619	378
	(4226)	(4801)	

Table 3: Parallel trend assumption for DID

	Males ATM (1)	**Females ATM** (2)	**Bride Price** (3)
Arsenic * Trend	-0.029	-0.019	0.12
	(0.020)	(0.012)	(0.13)
Observations	40288	45855	443
Other HH & Individual Controls	Yes	Yes	Yes
Year of Marriage FE	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes
District trends	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other individual level controls as mentioned in the Table 4 (for age at marriage) and Table 5 (for bride price). ATM refers to Age at Marriage.

Table 4: Dependent Var: Age at Marriage (Census Data)

	Males		**Females**	
	(1)	(2)	(3)	(4)
Arsenic * Post	-0.25***	-0.22**	-0.073	-0.12*
	(0.086)	(0.085)	(0.065)	(0.066)
Literate	0.30***	0.29***	-0.13***	-0.15***
	(0.041)	(0.041)	(0.027)	(0.027)
Religion is Islam	-1.02***	-1.02***	-0.64***	-0.64***
	(0.10)	(0.10)	(0.063)	(0.063)
Owns land	0.36***	0.36***	0.26***	0.27***
	(0.088)	(0.088)	(0.061)	(0.060)
Owns house	0.22***	0.22***	0.17***	0.17***
	(0.074)	(0.074)	(0.062)	(0.061)
House has cemented walls	0.89***	0.89***	0.42***	0.42***
	(0.054)	(0.053)	(0.035)	(0.034)
Observations	75243	75243	86985	86985
Control Mean Age at Marriage	23.34	23.34	17.79	17.79
Year of Marriage FE	Yes	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes	Yes
District trends	No	Yes	No	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Table 5: Dependent Var: Log of Bride Price

	(1)	(2)
Arsenic * Post	-0.340	-0.644*
	(0.453)	(0.368)
Education bride (in years)	0.108***	0.108***
	(0.0235)	(0.0242)
Education groom (in years)	-0.0485*	-0.0361
	(0.0239)	(0.0215)
Bride's age at the time of marriage	-0.0453	-0.0532
	(0.0425)	(0.0407)
Groom's age at the time of marriage	-0.0116	-0.0139
	(0.0269)	(0.0276)
Partner chosen by bride	0.158	0.106
	(0.305)	(0.310)
Brides family richer	0.123	0.144
	(0.161)	(0.158)
Groom's family richer	0.239	0.266
	(0.167)	(0.158)
Observations	875	875
Control Mean Bride Price (in Taka)	53980	
Year of Marriage FE	Yes	Yes
Sub-District FE	Yes	Yes
District Trends	No	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Table 6: Robustness Check 1:
Matched Sample Analysis for Age at marriage (ATM)

	Males (1)	**Females** (2)
Arsenic * Post	-0.39**	-0.19*
	(0.16)	(0.10)
subdistrict-year level means		
Literate	0.85***	0.44***
	(0.18)	(0.14)
Religion is Islam	-1.43***	-0.70**
	(0.40)	(0.28)
Owns land	0.19	0.20
	(0.46)	(0.41)
Owns house	0.32	0.15
	(0.42)	(0.26)
House has cemented walls	0.49	0.61***
	(0.31)	(0.21)
Observations	3414	3441
Number of subdistricts	158	158
Control Mean Age at Marriage	23.63	17.94
Year of Marriage FE	Yes	Yes
Sub-district FE	Yes	Yes
District trends	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Table 7: Robustness Check 2:
Using alternate dataset - BDHS (for males only)

	(1)	(2)
Arsenic * Post	-0.80**	-0.84*
	(0.40)	(0.46)
Not Educated	-0.34*	-0.34*
	(0.19)	(0.20)
Religion is Islam	-1.82***	-1.87***
	(0.35)	(0.35)
Electricity connection	0.58**	0.58**
	(0.23)	(0.24)
HH has cemented walls	1.44***	1.42***
	(0.27)	(0.27)
Owns telephone	-0.71	-0.61
	(1.00)	(1.00)
Owns tv	-0.082	-0.072
	(0.24)	(0.24)
Owns bicycle	0.36*	0.36*
	(0.20)	(0.21)
Observations	4947	4947
Number of Clusters	1067	1067
Control Mean Age at Marriage	23.61	23.61
Year of Marriage FE	Yes	Yes
Cluster FE	Yes	Yes
District trends	No	Yes

Note: Standard errors in parentheses are clustered by cluster-year. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Table 8: Robustness Check 3:
Sexratio changes between 2001 and 2011

	(1)	(2)	(3)
Treatment*Post	0.031	0.031	0.021
	(0.022)	(0.021)	(0.034)
Post	0.149***	0.149***	0.000
	(0.011)	(0.011)	(.)
Treatment	-0.021		-0.016
	(0.013)		(0.010)
Constant	0.811***	0.804***	0.886***
	(0.006)	(0.005)	(0.007)
Observation	820	820	820
District FE	Yes	No	No
District*Year FE	No	No	Yes
Subdistrict FE	No	Yes	No

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Table 9: Robustness Check 4:
a) Shorter Analysis Window & b) Placebo Test Using Urban Sample

	Analysis window 1996 to 2006		**Urban sample**	
	Males	Females	Males	Females
	(1)	(2)	(3)	(4)
Arsenic * Post	-0.23**	-0.023	-0.042	-0.049
	(0.11)	(0.091)	(0.22)	(0.14)
Literate	0.31***	-0.15***	0.78***	0.21*
	(0.049)	(0.036)	(0.10)	(0.11)
Religion is Islam	-1.09***	-0.66***	-1.40***	-1.09***
	(0.11)	(0.067)	(0.21)	(0.20)
Owns land	0.47***	0.34***	0.67***	0.40***
	(0.11)	(0.071)	(0.16)	(0.14)
Owns house	0.18*	0.14**	0.075	0.10
	(0.092)	(0.068)	(0.14)	(0.11)
House has cemented walls	0.95***	0.48***	1.49***	0.99***
	(0.067)	(0.045)	(0.18)	(0.093)
Observations	40887	47162	14856	16742
Control Mean Age at Marriage	23.34	17.74	24.85	18.49
Year of Marriage FE	Yes	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes	Yes
District trends	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the Table 4.

Table 10: Robustness Check 5A: Age at Marriage - Other contamination measures

	Males		Females	
	(1)	(2)	(3)	(4)
MeanArsenic * Post	-0.0015**		-0.00077*	
	(0.00060)		(0.00042)	
Arsenic10 * Post		-0.18**		-0.077
		(0.086)		(0.069)
Observations	75243	75243	86985	86985
Other Individual Level Controls	Yes	Yes	Yes	Yes
Year of Marriage FE	Yes	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes	Yes
District trends	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the Table 4.

Table 11: Robustness Check 5B: Bride Price
Other contamination measures & Additional data

	(1) PKSF	(2) PKSF	(3) PKSF+BRULS
MeanArsenic * Post	-0.007**		
	(0.003)		
Arsenic10 * Post		-0.097	
		(0.266)	
Arsenic50 * Post			-0.605*
			(0.312)
Observations	875	875	1699
Other Individual Level Controls	Yes	Yes	Yes
Year of Marriage FE	Yes	Yes	Yes
Subdistrict FE	Yes	Yes	Yes
District trends	Yes	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the Table 5.

Table 12: Heterogeneity
Treatment Group = Arsenic Contamination > 100;
Control Group = Arsenic Contamination < 50

	Males (1)	**Females** (2)
Arsenic100 * Post	-0.28**	-0.23**
	(0.12)	(0.093)
Literate	0.29***	-0.15***
	(0.043)	(0.029)
Religion is Islam	-1.01***	-0.62***
	(0.12)	(0.069)
Owns land	0.35***	0.26***
	(0.097)	(0.066)
Owns house	0.23***	0.14**
	(0.081)	(0.067)
House has cemented walls	0.92***	0.43***
	(0.060)	(0.038)
Observations	63600	73351
Control Mean Age at Marriage	23.34	17.79
Year of Marriage FE	Yes	Yes
Subdistrict FE	Yes	Yes
District trends	Yes	Yes

Note: Standard errors in parentheses are clustered by sub-district. Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.

Appendix

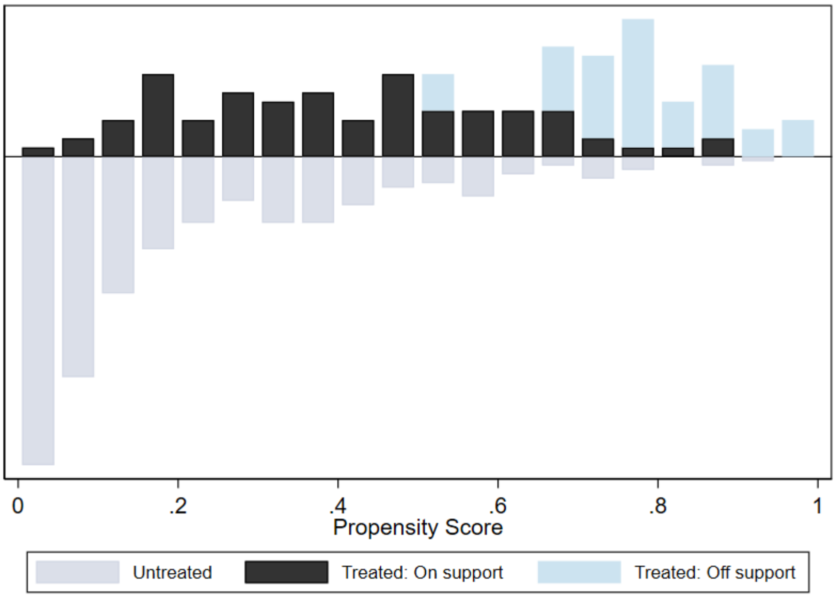
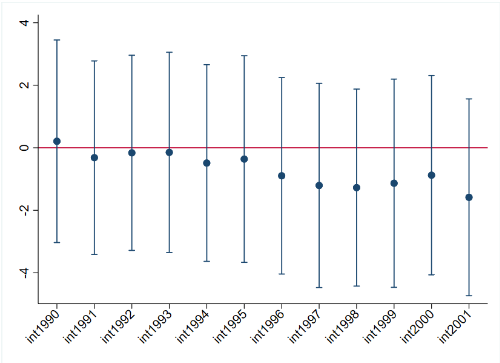
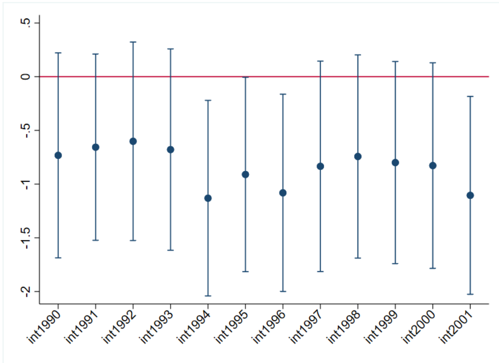


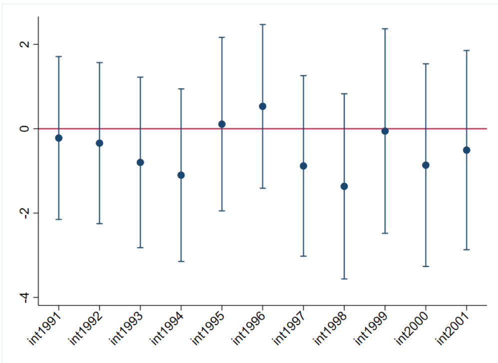
Figure A1: Density of sub-districts over propensity score. The matching is done using psmatch2 command in stata, with caliper value 0.05 and unique matching strategy with no replacement



(a) Age at marriage for males (Matched sample)



(b) Age at marriage for females (Matched sample)



(c) Age at marriage for males (BDHS)

Figure A2: Testing Parallel Assumption using specification 2: Plotting coefficients of interaction between year dummies and arsenic dummy for all pre-treatment years. Insignificance of coefficients points towards a similar trend being followed by the two groups in the pre-treatment period. The error bars are 95% confidence intervals, clustering at sub-district level.

Table A1: Sub-district characteristics, full sample and matched sample of sub-districts; means and t-stat for differences

	1	2	3	4	5	6
SUBDISTRICT LEVEL MEANS	**Control**	**Treated**	**t-stat ((1) vs (2))**	**Matched Control**	**Matched Treated**	**t-stat ((4) vs (5))**
employment category						
employed in agriculture (%)	0.56	0.63	-5.13	0.58	0.58	-0.01
employed in formal sector (%)	0.16	0.14	2.99	0.16	0.16	0.17
employed in business (%)	0.13	0.11	5.93	0.13	0.12	0.28
employed in others (%)	0.15	0.12	4.22	0.14	0.14	-0.30
education details						
literacy level (%)	0.44	0.40	4.02	0.43	0.43	-0.07
number of years of education	3.30	2.96	4.29	3.19	3.20	-0.08
completed primary education (%)	0.91	0.91	-1.41	0.91	0.91	-0.12
employment status details						
employed (%)	0.42	0.43	-4.50	0.43	0.43	0.84
unemployed (%)	0.02	0.02	3.18	0.02	0.02	-0.65
inactive (%)	0.15	0.14	3.15	0.15	0.15	0.39
involved in housework (%)	0.40	0.40	0.24	0.40	0.41	-1.42
household characteristics						
number of children	1.67	1.61	2.98	1.63	1.63	0.06
number of families	1.45	1.39	5.21	1.44	1.45	-0.36
electricity connection (%)	0.24	0.18	4.55	0.22	0.22	-0.10
ownership of house (%)	0.95	0.94	2.44	0.95	0.95	-0.71
religion is muslim (%)	0.87	0.89	-2.39	0.87	0.89	-1.18
other demographic characteristics						
ratio of males to females for adults	0.49	0.50	-5.44	0.50	0.49	1.65
ratio of males in children with age <1 year	0.52	0.52	-0.84	0.52	0.52	1.12
Number of Sub-districts	277	133		79	79	

*Source - IPUMS data for 2001

Table A2: Probit Estimation for Propensity Score Matching

Sub-district level means	**Arsenic Contamination = 1 or 0**
employment category	
employed in agriculture	0.11
	(0.44)
employed in formal sector	0.11
	(0.64)
employed in business	2.05**
	(0.88)
employed in others	-
education details	
literate	-1.88**
	(0.80)
number of years of education	0.39***
	(0.14)
completed primary education	6.55***
	(1.57)
employment status details	
employed	1.40
	(1.45)
unemployed	-2.83
	(5.85)
inactive	0.16
	(0.86)
involved in housework	-
household characteristics	
number of children	0.23**
	(0.11)
number of families	-0.080
	(0.27)
have electricity connection	0.77***
	(0.19)
ownership of house	2.12***
	(0.64)
religion is Islam	-0.93***
	(0.23)
No toilet facility	-0.59***
	(0.16)
other demographic characteristics	
ratio of males to females in adults >15 years	-7.58***
	(1.81)
ratio of males in children with age <1 year	-0.92
	(0.94)
Psuedo R-square	0.3089
Observations (number of subdistricts)	410

Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regression includes a constant term and other controls as mentioned in the table.