Natural Shocks and Marriage Market: The Case of Arsenic Exposure in Rural Bangladesh

Shyamal Chowdhury [1] **and Prachi Singh** [2]

[1]University of Sydney

[2]Indian Statistical Institute, Delhi

Natural Shocks 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 for males 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. 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 contained in Demographic and Health Survey datasets for Bangladesh. We find that age at marriage for males reduces by 3.5 to 10.5 months and bride price also reduces by around 60 percent.

JEL Classification: I12 J12 R11

Corresponding author:

Prachi Singh
Economics and Planning Unit
Indian Statistical Institute
New Delhi
Email: prachi.singh12r@isid.ac.in

Natural Shocks and Marriage Market: The Case of Arsenic Exposure in Rural Bangladesh

1 Introduction

Arsenic contamination of drinking water causes a wide spectrum of health problems which range from skin lesions to various kinds of cancers. The health aspect of arsenic contamination has a direct bearing on longevity for males, and beauty and fertility for females. All of these factors play an important role in the marriage market (Buss, 1989). **(add more references on beauty premium etc.)** While direct health effects of arsenic contamination have been explored in the literature but the indirect behavioral changes which are likely to be present have largely not been explored as yet. The recent paper by Keskin et al. (2017) is an exception in this regard and it has explored the change in breastfeeding behaviour by mothers in response to the arsenic contamination information campaign. We contribute to this narrow literature by providing empirical evidence for behavioural change in marriage market in response to information about arsenic contamination in the neighbourhood.

(Discuss the public health interventions, few pre-intervention stats, e.g., child mortality, and post-interventions stats, e.g., its primary effects) For decades Bangladesh has been riddled with health problems arising from poor water quality, starting from 1970s when high mortality and morbidity rate was linked to diseases emanating from pathogen-laden surface water usage. A switch to groundwater was soon advocated and executed which lead to millions of tubewells being dug up all across the country.

(discuss what went wrong)During mid 1980s arsenic was detected in groundwater and by late 1990s it was discovered that 35-77 million people were inadvertently exposed to poisonous arsenic in their drinking water (Smith et al., 2000).

(Emphasize on effects associated with beauty, fertility and mortality)The health problems associated with arsenic exposure took various forms which have been explored in bio-medical literature. Skin lesions are among the first few symptoms of arsenic poisoning, this is also the most prevalent symptom which directly affects beauty of a person. 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. Another study (Argos et al.,2010) found that risk of all-cause and chronic-disease mortality was higher for people exposed to dangerous level of arsenic. 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., 2015) are among other devastating effects of arsenic exposure. Tragically arsenic problem is not just a health problem but a social problem as well. 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). Lastly arsenic exposure has also been found to have implications in the labour market in form of reduced labour supply (Carson et al., 2010).

(Discuss the intervention briefly, and discuss how we use it.)

The health outcomes associated with arsenic affect multiple attributes which matter in 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. All 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. (The above para needs to be more focused) We focus our attention on two outcomes in marriage market - age at marriage for males [3] and bride price (an integral part of marriages in Islamic culture), which 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. The age at marriage marks the transition to adulthood and has important consequences linked to family size, educational and labour market achievement. In Bangladesh, where the risk of being affected with arsenicosis is positive, the perceived risk (about own health) and age at marriage could possibly be related (Spivey points out risk aversion significantly affects time to marriage, with more risk averse males marrying sooner than their more risk-loving counterparts (Spivey, 2010)).**(This points out that only the risk averse people marry sooner, not much how a general increase in risk level induces people to change their behavior at the margin)**

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. We essentially link shocks emanating from local disease environment to marriage market. The information campaign which focused on arsenic contamination in Bangladesh 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.

Our study is primarily at the sub-district (*upzila*)level, given that arsenic contamination is quite random in nature **(We need to be careful here -the randomness is something that will be questioned.)**so we have few sub-districts which are arsenic affected while there are others which have a "safe" level of arsenic in the tubewell which is the main source of drinking water. We thus follow a difference-in-difference approach and compare changes in marriage market outcomes across marriage cohorts (a marriage cohort refers to people

[3]we don't focus our attention on age at marriage for females as we don't observe them in their original household where the decision regarding marriage took place. However we still provide some evidence in our appendix for age at marriage for females

getting married in the same year) belonging to arsenic affected areas to changes across the same marriage cohorts belonging to non-arsenic areas.

We look at all first marriages for males over a 20 year time period, with 1990-2001 forming the pre-treatment period while 2002-2011 forms the post-treatment period. We hypothesize 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 arsenicosis (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.

We follow four approaches (how are they different approaches?) to establish our results for age at marriage for males. Our baseline specification controls for time-invariant differences between sub-districts (such as geography) and location-invariant difference between marriage cohorts. Secondly we address presence of underlying trends in the economy by controlling for district (*zila*) level trends. Third, we restrict our sample to arsenic affected sub-districts and matched non-arsenic sub-districts that are comparable on pre-treatment characteristics. Finally we use rich spatial information about arsenic contamination of drinking water by relying on Demographic Health Surveys for Bangladesh which capture GPS coordinates of the interviewed cluster. Our bride price analysis uses two primary surveys and relies on approaches one and two above to arrive at results. Our identifying assumption is that conditional on sub-district and year of marriage fixed effects and district trends, marriage market trends were not changing differently between arsenic and non-arsenic areas other than for reasons related to the information campaign. An event study provides support for this assumption, where we find that prior to the campaign arsenic and non-arsenic areas were indistinguishable but they behaved differently after information campaign.

Using Census Data of Bangladesh (2011) which records details about marriage unions we find that age at marriage for males reduced by 3.5 to 5 months in arsenic affected areas

in response to information campaign, relying on precise spatial information the effect of information campaign turns to be much higher in magnitude (a reduction by 9.5 to 10.5 months in age at marriage). We also find that the information campaign had a dampening effect on bride price in arsenic affected areas.

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 and Section 7 concludes.

2 Background

Provide more precise stats on access to safe drinking water, infant mortality, and penetration of TWs in 1970, and what how penetration reached to almost 100% by 1990 and what happened to the mortality etc. During 1970s Bangladesh was struggling with disease burden due to water borne diseases. Soon millions of tubewells were dug all across the country. In rural areas the switch to groundwater source was almost universal with almost 97% of households using tubewells. Most of the tubewells were of shallow depth and unfortunately in few areas these tubewells contained dangerous levels of arsenic (shallow tubewells have a greater likelihood of containing arsenic) as was found in British Geological Survey (BGS) conducted in 1999.

The Bangladesh government (Department of Public Health Engineering(DPHE)) along with UNICEF and Non-Profit Organisations initiated a water quality information campaign during early 2000s. 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. Keskin et al., show in their paper that 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, they also reported awareness about symptoms of arsenicosis after the information campaign (Keskin at al., 2017). Other papers (mention them in a footnote) also find that people did switch to safer sources of water after the information campaign.

Information campaigns (especially the ones which disseminate negative information) have been found to have an effect on behavioural patterns of the focus groups. 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 et al. in their paper which analysis effect of HIV related curriculum in schools in Kenya also find 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).

3 Data

Our analysis combines arsenic contamination information with the demographic data. We describe below our various sources of data:

3.1 Demographic data:

a) Long Census Survey Data - 2011

The individual and household data are from Long Census Survey Data which was conducted by Bangladesh Bureau of Statistics (BBS) in March 2011. We use detailed segregated information at the sub-district level. This dataset has details about place of residence, duration of residence in the current district, type of construction of household, 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. The 2011 census has a sample of 167,293 households which covers all 64 districts of Bangladesh. We collapse our data to sub-district-year of marriage level to carry out our analysis. Our data shows that males in Bangladesh tend to get married earlier than the world average with mean age at first marriage being 24 years along with 70 percent of males being married by 25 years of age (Refer to Figure 1).

b) Integrated Public Microdata Series (IPUMS) Census Data - 2001

We match arsenic affected sub-districts with unaffected sub-districts from the long census data using sub-district level characteristics from IPUMS data. 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 25 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 [4]. We plot mean age at marriage for males for our unmatched sample and matched sample in Figure 3. We notice that the general trend was decreasing and in both the plots after information campaign the gap between age at marriage for males in arsenic and non-arsenic areas narrowed, which is specially true for our matched sample with sharp dip in the treatment year (2002).

c) Demographic Health Surveys

In order to use rich spatial data about arsenic contamination we rely on Bangladesh Demographic and Health Surveys (BDHS), conducted in 1999, 2004, 2007 and 2011. These surveys were conducted by National Institute of Population Research and Training (NIPORT) and

[4]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

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 4800 men sampled from 1067 clusters. We use this dataset to replicate and complement the results which we find using Long Census data.

d) Primary Surveys

For our bride price analysis we rely on two surveys: Bangladesh Rural Urban Linkage Survey(BRULS) and Palli Karma-Sahayak Foundation (PKSF) Suvey. Both the surveys contained identical modules on marriage, divorce, bride price and dowry. The BRUL Survey was commissioned by International Food Policy Research Institution between December 2004 and January 2005, while the second survey(PKSF) was a primary survey conducted between December 2010 and January 2011. BRULS contains detailed information for 1482 marriage unions while PKSF survey contains information on 2400 marriages. We combine both the datasets to arrive at our final estimation sample for our period of study (1990-2011), which contains 1699 marriages from 80 (can you check the number? looks too large to me.) sub-districts. These datasets contain 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. Since there are fewer observations per year in the combined dataset that we use for bride price analysis, so in Figure 4 we plot a smooth quadratic fitted graph which looks at mean bride price in arsenic and non-arsenic areas (This essentially plots fitted values from the regression of mean bride price on a trend variable and its square). We find that in non-arsenic areas bride price was increasing overtime, while in case of arsenic areas it reached a peak and then it steadily deviated from upward trend and started falling in post treatment (information campaign) period.

3.2 Contamination data:

British Geological Survey (BGS)- 1999

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. In this survey, BGS tested over 3500 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. 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). Figure 2 plots all the tubewells which were surveyed by BGS, with the red ones belonging to the unsafe category. As is evident from the figure there is geographical variation in arsenic contamination in Bangladesh, with contaminated tubewells mostly lying in the southern region. However it should be noted that the BGS only sampled 61 out of total 64 sub-district and we restrict our study to these 61 sub-district. In addition, no tube-well labeling or information campaign followed the BGS survey.

Our measures of contamination constructed using BGS data are at the sub-district level except in case of BDHS data where we construct contamination figures for each individual cluster, which are smaller than sub-districts?. If the mean arsenic figure for a sub-district was above >50ug/litre then we code it as an arsenic sub-district (and sub-district with contamination <50ug/litre are coded as non-arsenic affected sub-districts), while in case of BDHS data these mean figures were calculated in the 5 mile radius around the cluster location to construct a "local" level of contamination with the same cutoff for contamination, that is 50ug/litre. 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.

Final Estimation Sample

not clear to meOur final estimation sample for age at marriage for males merges sub-district level data from IPUMS dataset to individual level data from Long census data. Finally we collapse the data to sub-district-year of marriage level and combine it with BGS data to get contamination information at the sub-district level. For age at marriage results which use BDHS data we directly combine individual level data from BDHS to BGS data to get cluster level contamination information. Lastly for bride price results, individual level data from the two surveys (BRULS and PKSF) are directly combined with BGS data to retrieve contamination information at sub-district level.

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), beginning with the following baseline specification:

$$A_{ut} = \alpha + \theta Arsenic_u * post_t + \beta X_{ut} + \gamma_u + \delta_t + \varepsilon_{ut} \quad (1)$$

in the above equation, you should make a distinction between individual level variables and sub-district level variables by adding an additional subscript. where A_{ut} is the marriage market outcome - mean age at first marriage for males at sub-district - year of marriage level for individuals who reside in sub-district u and who belong to marriage cohort t (i.e. who got married in year t) (the previous sentence is not clear - divide it into two and make it clearer.). $Arsenic_u$ is the dummy variable which takes value 1 if the mean arsenic contamination level is greater than 50ug/litre for a sub-district, $post_t$ is the dummy variable for post treatment period (2002-2011) and X_{ut} includes control for mean figures for religion being Islam, literacy, ownership of land and house and a dummy for cemented wall of house. To control for time-invariant differences between sub-district and location-invariant differences between marriage

cohorts we include sub-district fixed effects, γ_u and marriage cohort fixed effects, δ_t. We thus estimate θ by comparing 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. We next control for district level trends to control for underlying trends present in the economy (add a notation for it in the equation, which will increase a subscript as well.).

Table 1 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 address this concern we follow a matching procedure by using sub-district level characteristics from IPUMS data. We match sub-districts using pre-treatment characteristics from 2001 which includes controls for status of employment, type of employment, educational attainments, household level characteristics like number of families, number of children, gender mix of households, mean age of household etc. We also control for mean age of adult unmarried males and females, their sex-ratio and sex-ratio for children below age one. Our matching exercise results in 83 arsenic affected sub-districts getting matched to unique 83 non-arsenic sub-districts, and column 6 in table 2 shows that the matched treated and matched controls are now comparable [5]. We then combine our DID with propensity score matching by restricting our study to matched treated and control sub-districts .

For our final analysis for age at first marriage for males we turn to BDHS data which uses rich spatial information about contamination of drinking water. We replicate our results from census data using the following analysis:

$$A_{ict} = \alpha + \theta Arsenic_c * post_t + \beta X_{ict} + \gamma_c + \delta_t + \varepsilon_{ict} \tag{2}$$

here the unit of analysis is an individual i, from cluster c who belongs to marriage cohort t. We have arsenic contamination, $Arsenic_c$ defined at cluster level as the dummy variable which takes value 1 if mean arsenic contamination is greater than 50ug/litre in the 5 mile

[5]The probit regression for generating propensity scores for matching is presented in appendix table 2 and figure 1 in appendix plots the number of treated and control sub-districts matched over p-score

radius around the cluster location. The controls are similar to the ones used in equation (1) above. We additionally also control for district level trends.

We now turn our attention to other marriage market outcome which is bride price. We use data from two surveys (BRULS and PKSF) to do this analysis. We again follow a DID methodology to identify effect of information campaign on bride price.

$$BP_{iut} = \alpha + \phi Arsenic_u * post_t + \beta X_{iut} + \gamma_u + \delta_t + \varepsilon_{iut} \tag{3}$$

where BP_{iut} is the log of bride price for marriage union i which took place in sub-district u in year t. All other variables are same as the ones defined for equation (1) except for control variables included in X_{iut}, which includes 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). ϕ is our coefficient of interest which identifies the effect of information campaign by comparing difference in bride price before and after information campaign for marriages taking place in arsenic areas which the same difference for marriages happening in non-arsenic areas. As an additional approach to control for underlying trends we introduce district level trends as well in our analysis.

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

$$Y_{ijt} = \alpha + \Sigma_{2001}^{1990}\lambda_t(Arsenic_j * I_t) + \beta X_{ijt} + \gamma_j + \delta_t + \varepsilon_{ijt} \tag{4}$$

(do we control for district trend in the above?) where Y_{ijt} is our outcome of interest: age at first marriage for males or bride price and j is the level at which arsenic contamination is measured, so it is a sub-district for equation (1) and (3), while its a cluster for equation (2). 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 and joint significance of all these interaction terms. If these terms are individually (and jointly) insignificant then that reveals that the parallel trend assumption is satisfied as the two groups followed similar trend in the pre-treatment period.

Any possible threats against identification strategy followed so far? Some discussion that decrease in marriage age and bride price can happen through a third factor such as income?

5 Results

Marriage Market Response to Information Regarding Arsenic Contamination:

A) Age at Marriage for Males

Table 2 presents results from specification 1. Each column presents results corresponding to multiple approaches which we employ to estimate our model. All columns include sub-district and marriage cohort (year of marriage) fixed effects. In column 2 we additionally control for district level trends. The next two columns 3 and 4 presents results from our restricted sample of matched treated and control sub-districts. Table 2 shows that individuals from arsenic affected areas are getting married earlier after implementation of information campaign. The coefficients are relatively stable across specifications and they are somewhat bigger in magnitude for the matched sample. For the unmatched sample the reduction in mean age at marriage for males is around 3.5 months, while for the matched sample the magnitude is slightly higher, around 4.8 months (discuss what it means in % terms, e.g., x% reduction in marriage age etc.) [6]. Other control variables reveal that higher literacy increases the age at marriage, also sub-districts with greater population of Muslims tend to

[6]We don't focus on female's age at marriage 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, we estimate a similar model for age at marriage for females and find that mean age at marriage for females also reduced by 2 months. Results for female age at first marriage are provided in appendix table 1

get married earlier.

We also did an event study, where we ran specification 4 for our entire study period (1990-2011) and plotted coefficients of interaction between year dummies and arsenic dummy. Figure 6 captures our event study (for the matched sample) in which most of interaction coefficients are insignificant in the pre-treatment period in comparison to coefficients in post-treatment period where they are significant. We also test whether the average of all coefficients in pre-treatment period [1990-2001] is equal to the average of all coefficients in the post-treatment period [2002-2011] and we reject this hypothesis with p-value = 0.001 (corresponding p-value for unmatched sample is 0.002).

We next present our results using BDHS data which uses spatial information about arsenic contamination at cluster level. Table 3 presents results from specification 2, 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 our variable of interest, that is interaction term $\texttt{Arsenic}_c * \texttt{post}_t$ is negative and significant in both the columns [7], with the order of magnitude being 9.5 to 10.5 months, which is much larger in comparison to the effect seen using long census data. We find similar results as before for our control variables as well, with education and religion significantly affecting age at marriage.

B) Bride Price

Table 4 presents our results from specification 3 which models bride price. We present results from just PKSF dataset [8] and from the combined dataset (i.e. PKSF + BRULS). All columns include sub-district and marriage cohort fixed effects. We provide results based on two measures of contamination: one is based on continuous measure of arsenic (variable name - meanarsenic) and another one which uses a dummy variable for mean arsenic contamination being greater than 50ug/litre (variable name - arsenic). All results have negative

[7] coefficient for the interaction term in column 2 of table 3 is significant at 6% level

[8] BRULS dataset doesn't contain any arsenic affected sub-district

coefficients, implying that bride price reduced in comparison to non-arsenic areas after information campaign.

C) Test for Parallel Trend Assumption

We estimate specification 4 for all our results for the pre-treatment period and present our results in a visual format in figure 5. We notice that for each of 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. In Panel d) of figure 5 for bride price results we notice that for years 1996 and 1997 arsenic and non-arsenic areas were significantly different, we re-estimate our model by dropping observations for these two years and our results still hold [9].

6 Discussion

most of it can go to the Introduction and some of it in the Results section

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. 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 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

[9] results are available on request

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 [10].

[10]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 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.

7 Conclusion

The behavioural change induced by an information campaign has mainly been explored in the health dimension (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.

Bibliography

Ahmad, J., B.N. Goldar, S. Misra and M. Jakariya (2003). "Willingness to Pay for Arsenic-Free, Safe Drinking Water in Bangladesh." Dhaka: World Bank Water and Sanitation Program-South Asia.

Ahmad, S.A., M.H. Sayed, S. Barua, M.H. Khan, M.H. Faruquee, A. Jalil, S.A. Hadi and H.K. Talukder. (2001). "Arsenic in Drinking Water and Pregnancy Outcomes." *Environmental Health Perspectives.*

Ahsan, H., Chen, Y., Parvez, F., Zablotska, L., Argos, M., Hussain, I., and Van Geen, A. (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.*

Ahsan H, Perrin M, Rahman A, et al. (2000). "Associations between Drinking Water and Urinary Arsenic Levels and Skin Lesions in Bangladesh." *Journal of Occupational Environmental Medicine.*

Alam, M. G. 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.*

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., and Sarwar, G. (2010). "Arsenic Exposure from Drinking Water, and All-cause and Chronic-Disease Mortalities in Bangladesh (HEALS): A Prospective Cohort Study." *The Lancet.*

Bjerk, D. (2009). "Beauty vs. Earnings: Gender Differences in Earnings and Priorities Over Spousal Characteristics in a Matching Model." *Journal of Economic Behavior and*

Organization.

British Geological Survey and Department of Public Health Engineering (BGS and DPHE) (2001). "Arsenic Contamination of Groundwater in Bangladesh." Edited by D.G. Kinniburgh and P.L. Smedley. Keyworth, U.K.: British Geological Survey.
http://www.bgs.ac.uk/arsenic/bangladesh/.

Buss, D. M. (1989). "Sex Differences in Human Mate Preferences: Evolutionary Hypotheses Tested in 37 cultures." *Behavioral and Brain Sciences.*

Caldwell, B., J. Caldwell, S.N. Mitra and W. Smith (2003). "Tubewells and Arsenic in Bangladesh: Challenging a Public Health Success Story." *International Journal of Population Geography.*

Carson, R. T., Koundouri, P., and Nauges, C. (2010). "Arsenic Mitigation in Bangladesh: A household Labor Market Approach." *American Journal of Agricultural Economics.*

Chen, Y., A. van Geen, J.H. Graziano, A. Pfaff, M. Madajewicz, F. Parvez, A.Z. Hussain, V. Slavkovich, T. Islam and H. Ahsan (2007). "Reduction in Urinary Arsenic Levels in Response to Arsenic Mitigation Efforts in Araihazar, Bangladesh." *Environmental Health Perspectives.*

Chiappori, P. A., Oreffice, S., and Quintana-Domeque, C. (2012). "Fatter Attraction: Anthropometric and Socioeconomic Matching in the Marriage Market." *Journal of Political Economy.*

Chowdhury, S., Krause, A., and Zimmermann, K. F. (2016). "Arsenic Contamination of Drinking Water and Mental Health." SSRN Working Paper

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.*

Duflo, E., Dupas, P., and Kremer, M. (2015). "Education, HIV, and early fertility: Experimental Evidence from Kenya." *The American Economic Review.*

Dupas, P. (2011a). "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics*

Dupas, P. (2011b). "Health Behavior in Developing Countries." *Annual Review of Economics*

Elder Jr, G. H. (1969). "Appearance and Education in Marriage Mobility." *American Sociological Review*

Field, E., and Ambrus, A. (2008). "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh." *Journal of Political Economy.*

Gale, D., and Shapley, L. S. (1962). "College Admissions and the Stability of Marriage." *The American Mathematical Monthly.*

Hassan, M. M., P. J. Atkins, and C. E. Dunn. (2005). "Social Implications of Arsenic Poisoning in Bangladesh". *Social Science and Medicine*

Kalmijn, M. (1998). "Intermarriage and Homogamy: Causes, Patterns, Trends." *Annual Review of Sociology.*

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.*

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.*

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*

Marini, M. M. (1978). "The Transition to Adulthood: Sex Differences in Educational Attainment and Age at Marriage." *American Sociological Review.*

Milton, A. H., Smith, W., Rahman, B., Hasan, Z., Kulsum, U., Dear, K. and Ali, A. (2005). "Chronic Arsenic Exposure and Adverse Pregnancy Outcomes in Bangladesh." *Epidemiology.*

Mobarak, A. M., P. Dwivedi, R. Bailis, L. Hildemann and G. Miller. (2012) "Low Demand for Nontraditional Cookstove Technologies." *Proceedings of the National Academy of Sciences.*

Rahman, A., Vahter, M., Ekström, E. C., Rahman, M., Golam Mustafa, A. H. M., Wahed, M. A. 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.*

Spivey, C. (2010). "Desperation or Desire? The Role of Risk Aversion in Marriage." *Economic Inquiry.*

Smith AH, Lingas EO, Rahman M. (2000). "Contamination of Drinking Water by Arsenic in Bangladesh: A Public Health Emergency." *Bulletin of World Health Organisation*

UNICEF. (2008). "Arsenic Mitigation in Bangladesh." New York: UNICEF. www.unicef.org/bangladesh/Arsenic.pdf.

van Geen, A., H. Ahsan, A. H. Horneman, R. K. Dhar, Y. Zheng, I. Hussain, K. M. Ahmed, A. Gelman, M. Stute, H. J. Simpson, S. Wallace, C. Small, F. Parvez, V. Slavkovich, N. J. LoIacono, M. Becker, Z. Cheng, H. Momotaj, M. Shahnewaz, A. A. Seddique and J. H. Graziano. (2002). "Promotion of Well-Switching to Mitigate the Current Arsenic Crisis in Bangladesh." *Bulletin of the World Health Organization*

World Bank. (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." Washington, D.C.: World Bank.

Tables and Figures

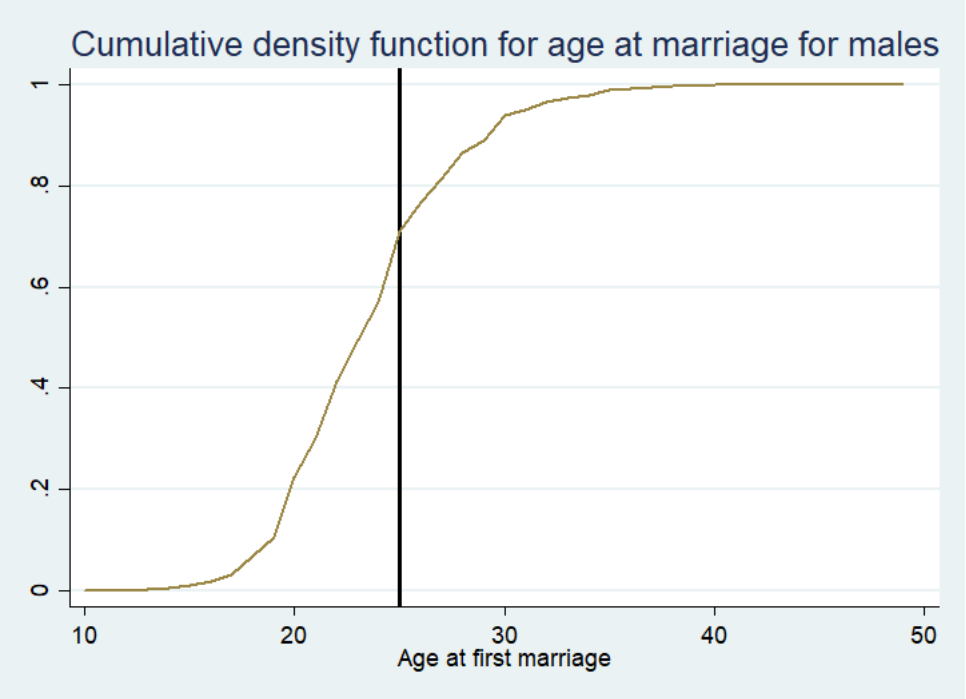


Figure 1: Cumulative Density Function for Age at Marriage for Males: Long Census Data 2011

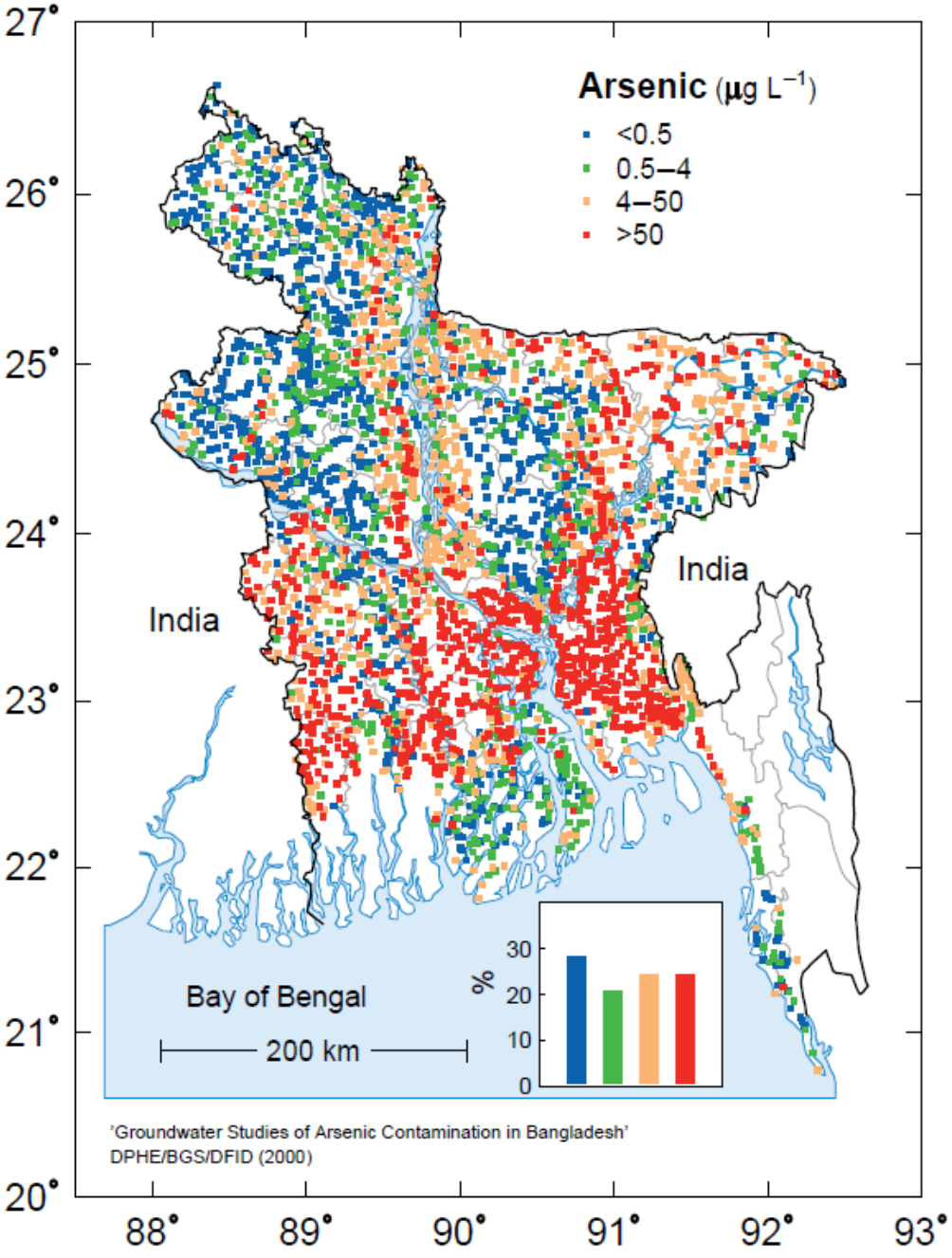


Figure 2: Arsenic Contamination Map for Bangladesh: Source BGS 1999

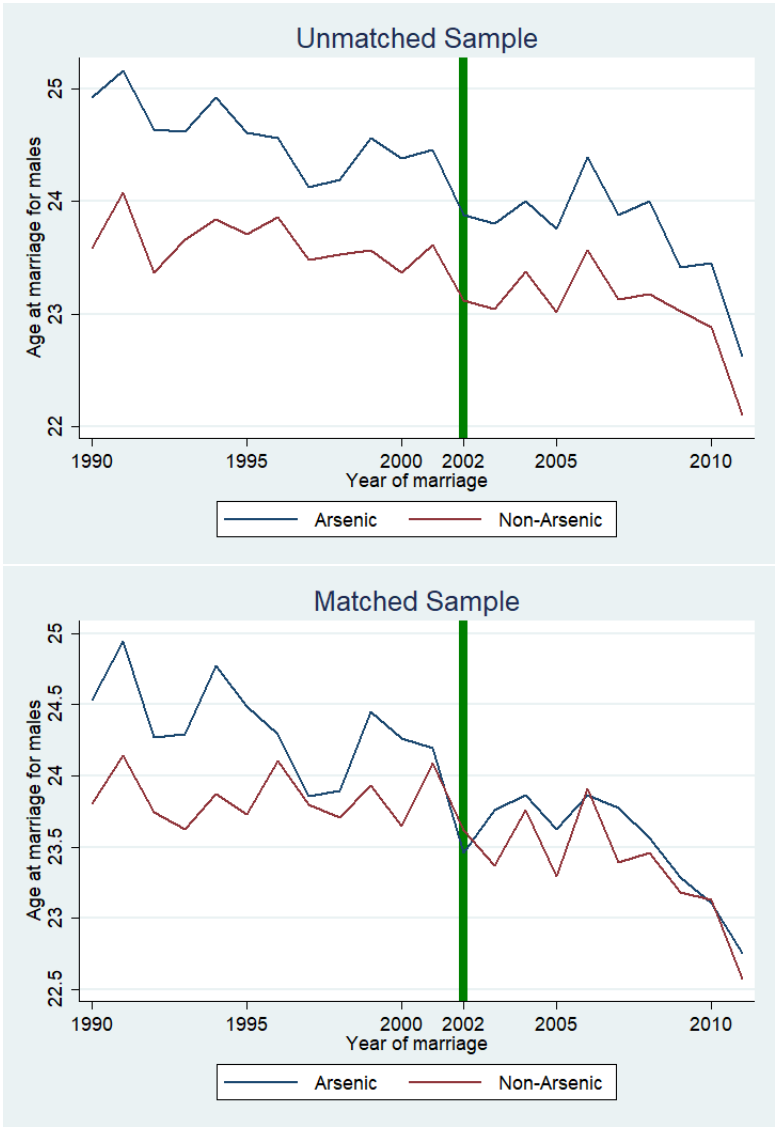


Figure 3: Mean age at marriage for males over time using long census data - Unmatched and Matched sample

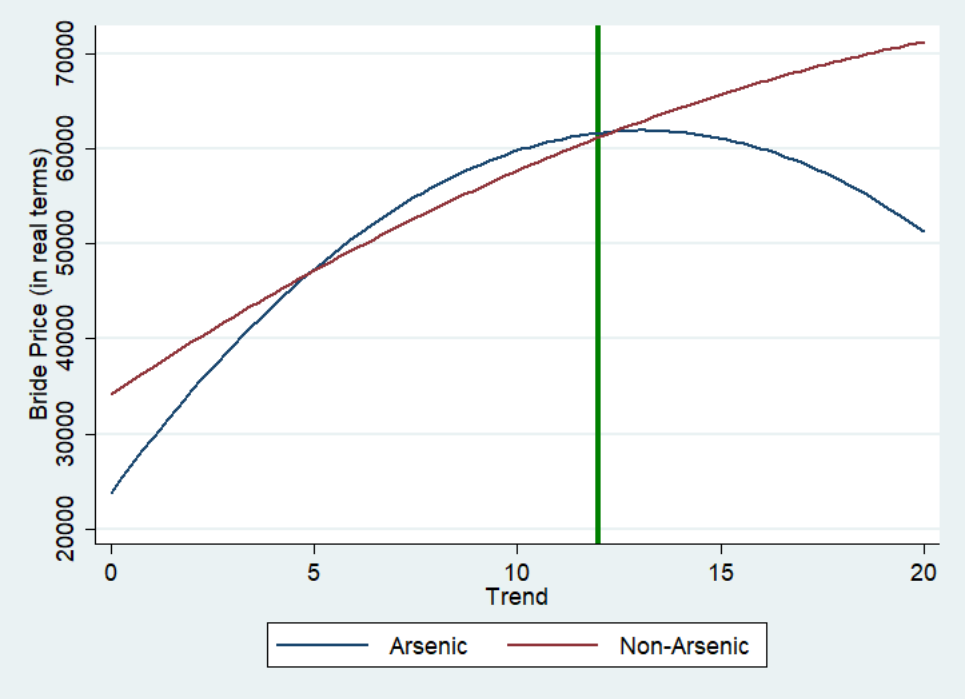
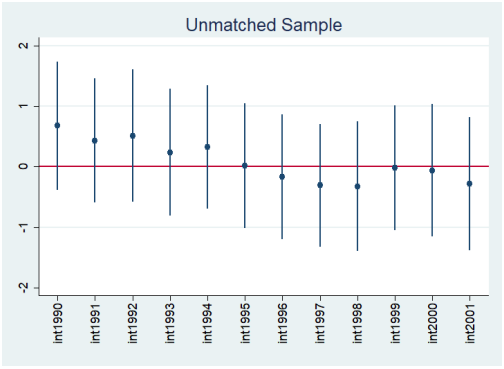
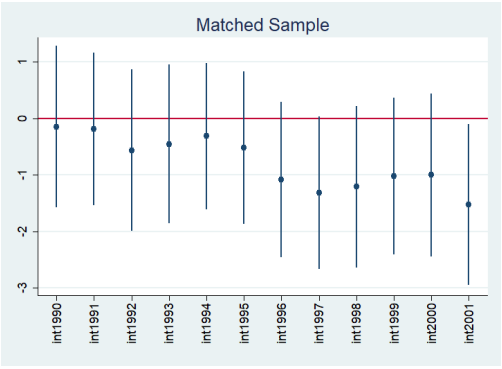


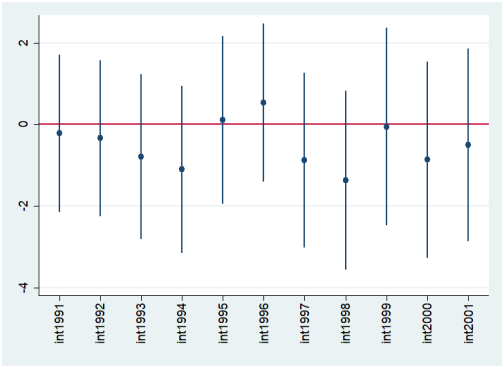
Figure 4: Real Bride Price over time: this figure plots fitted values from a regression of real bride price on a trend variable and its square, where the trend variable takes value 1 for year 1990, 2 for 1991 and so on till year 2010.



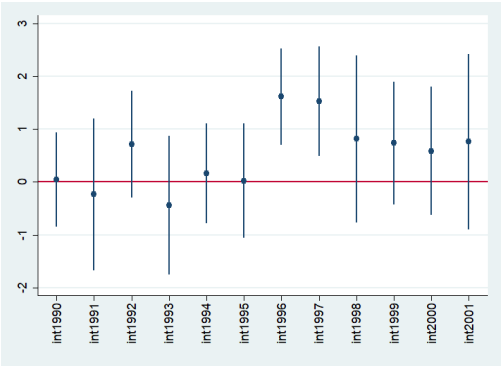
(a) Age at marriage for males: Long census data - Unmatched Sample



(b) Age at marriage for males: Long census data - Matched Sample



(c) Age at marriage for males: BDHS data



(d) Bride Price: Primary survey data

Figure 5: Testing Parallel Assumption using specification 4: 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. Joint significance of all coefficients has also been tested in all regressions and it was rejected in each of the cases at 10 percent level. The error bars are 95% confidence intervals, clustering is at sub-district or cluster level depending upon data being used

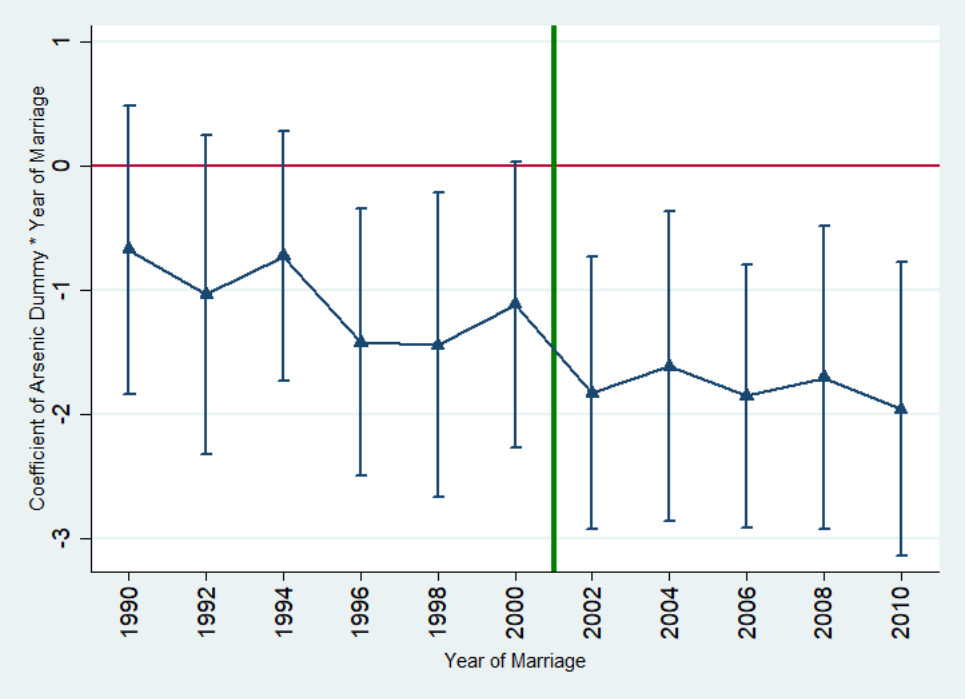


Figure 6: This figure plots the coefficients of interaction between arsenic dummy and year of marriage dummies for the matched sample of long census data, in a regression which is similar to specification 4 but which covers the entire study period. The error bars are 95% confidence intervals, clustering at sub-district level

Table 1: sub-district (upzila) characteristics, full sample and matched sample of sub-districts; means and t-stat for differences

SUB-DISTRICT LEVEL MEANS	1 Control	2 Treated	3 t-stat ((1) vs (2))	4 Matched Control	5 Matched Treated	6 t-stat ((4) vs (5))
employment category						
employed in agriculture, perc	0.65	0.57	-5.31	0.60	0.59	-0.51
employed in formal sector, perc	0.14	0.15	2.91	0.15	0.15	0.58
employed in business, perc	0.11	0.14	6.16	0.13	0.13	0.36
employed in others, perc	0.10	0.14	4.47	0.12	0.13	0.31
education details						
literacy level, perc	0.40	0.44	3.94	0.41	0.42	0.44
number of years of education	2.96	3.30	4.20	3.08	3.14	0.49
completed primary education, perc	0.91	0.91	-1.34	0.91	0.91	-0.56
employment status details						
employed, perc	0.43	0.42	-4.54	0.43	0.43	-0.45
unemployed, perc	0.02	0.02	3.16	0.02	0.02	0.48
inactive, perc	0.14	0.15	3.15	0.15	0.15	0.68
involved in housework, perc	0.40	0.40	0.29	0.40	0.40	-0.34
household characteristics						
number of children	1.61	1.67	3.05	1.65	1.64	-0.06
number of families	1.39	1.45	5.20	1.44	1.44	0.11
number of couples	0.89	0.86	-4.32	0.89	0.88	-0.77
electricity connection, perc	0.18	0.24	4.45	0.22	0.22	0.13
age	25.21	25.26	0.34	25.27	25.29	0.09
ownership of house, perc	0.94	0.95	2.55	0.95	0.95	-0.71
number of mothers	0.94	0.98	5.66	0.97	0.97	-0.30
number of fathers	0.81	0.80	-2.39	0.82	0.81	-0.90
religion is muslim, perc	0.90	0.88	-1.97	0.90	0.90	0.04
polygamous, perc	0.00	0.00	-5.03	0.00	0.00	-0.50
demographic characteristics of unmarried adults above age 16						
female age	24.41	24.58	1.30	24.75	24.54	-1.02
male age	22.57	22.96	4.79	22.91	22.88	-0.27
ratio of males	0.71	0.69	-4.60	0.70	0.70	-0.14
ratio of males in children with age < 1 year	0.52	0.52	-0.87	0.52	0.52	0.85
Number of sub-districts	277	132		83	83	

Source: IPUM Survey 2001

Table 2: Dependent Var: Age at Marriage - Males (Census Data)

	Unmatched Sample		**Matched Sample**	
	(1)	(2)	(3)	(4)
arsenic * post	-0.296***	-0.289**	-0.359**	-0.405***
	(0.105)	(0.114)	(0.150)	(0.154)
sub-district-year of marriage level means				
literate	0.305**	0.275**	0.481***	0.466***
	(0.125)	(0.125)	(0.165)	(0.167)
muslim	-0.906***	-0.945***	-1.282***	-1.364***
	(0.236)	(0.236)	(0.398)	(0.401)
ownland	0.454	0.499	0.329	0.425
	(0.370)	(0.371)	(0.455)	(0.459)
ownhouse	0.247	0.238	-0.207	-0.229
	(0.258)	(0.262)	(0.414)	(0.421)
cement_wall	0.540**	0.534**	0.642**	0.672**
	(0.209)	(0.208)	(0.302)	(0.297)
Observations	8823	8823	3580	3580
Number of sub-districts	409	409	166	166
R-square	0.36	0.37	0.35	0.37
Year of Marriage FE	Yes	Yes	Yes	Yes
Sub-district FE	Yes	Yes	Yes	Yes
district trends	No	Yes	No	Yes
mean age at marriage (in Non-Arsenic areas)	23.36		23.63	

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

Table 3: Dependent Var: Age at Marriage - Males (DHS Data)

	(1)	**(2)**
arsenic * post	-0.795**	-0.876*
	(0.400)	(0.466)
dummy for no education	-0.349*	-0.340*
	(0.193)	(0.198)
muslim	-1.783***	-1.816***
	(0.350)	(0.347)
electricity	0.614**	0.608**
	(0.240)	(0.242)
cement_wall	1.450***	1.429***
	(0.270)	(0.270)
telephone	-0.696	-0.594
	(0.999)	(0.996)
tv	-0.178	-0.161
	(0.242)	(0.243)
bicycle	0.374*	0.384*
	(0.204)	(0.208)
Observations	4869	4869
Number of Clusters	1067	1067
Year of Marriage FE	Yes	Yes
Cluster FE	Yes	Yes
district trends	No	Yes
mean age at marriage (in Non-arsenic areas)	23.60	

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

Table 4: Dependent Var: Log of Bride Price

	PKSF		**PKSF + BRULS**	
	(1)	(2)	(3)	(4)
meanarsenic * post	-0.00741**		-0.00664**	
	(0.00273)		(0.00262)	
arsenic * post		-0.644*		-0.605*
		(0.368)		(0.312)
edu_wife	0.109***	0.108***	0.0983***	0.0980***
	(0.0241)	(0.0242)	(0.0123)	(0.0123)
edu_dif	-0.0354	-0.0361	-0.0502***	-0.0505***
	(0.0216)	(0.0215)	(0.0134)	(0.0134)
age_bride	-0.0541	-0.0532	-0.0458**	-0.0448*
	(0.0406)	(0.0407)	(0.0226)	(0.0227)
age_dif	-0.0134	-0.0139	0.00769	0.00722
	(0.0275)	(0.0276)	(0.0155)	(0.0156)
who_chose	0.0928	0.106	0.0952	0.101
	(0.309)	(0.310)	(0.217)	(0.218)
rich1	0.141	0.144	0.0344	0.0363
	(0.157)	(0.158)	(0.0928)	(0.0930)
rich2	0.271*	0.266	0.112	0.110
	(0.153)	(0.158)	(0.0890)	(0.0907)
Observations	875	875	1699	1699
Year of Marriage FE	Yes	Yes	Yes	Yes
sub-district FE	Yes	Yes	Yes	Yes
district trends	Yes	Yes	Yes	Yes
mean bride price in real terms (in Non Arsenic areas)	53980		51862	

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

Appendix

Table 1: Dependent Var: Age at Marriage - Females (Census Data)

	Unmatched Sample (1)	**Matched Sample** (2)
arsenic * post	-0.170**	-0.172*
	(0.0738)	(0.0981)
sub-district level means		
literate	0.0600	0.228
	(0.0990)	(0.160)
muslim	-0.727***	-0.707***
	(0.167)	(0.227)
Observations	8895	3615
Number of sub-districts	409	166
Year of Marriage FE	Yes	Yes
Sub-district FE	Yes	Yes
District trends	Yes	Yes
mean age at marriage (in Non-Arsenic areas)	17.79	17.91

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. Regressions include a constant term and all other controls are mentioned in the table.

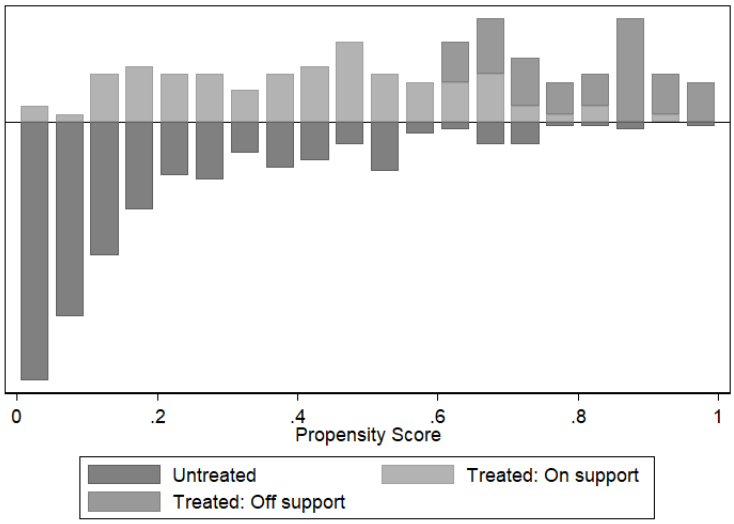


Figure 1: Density of sub-districts over propensity score
The matching is done using psmatch2 command in stata, with caliper value 0.05 and unique matching strategy with no replacement

Table 2: Probit Estimation for Propensity Score Matching

sub-district level means	**Dependent variable: Dummy for Arsenic Contaminated Sub-district**
employment category	
employed in agriculture	0.360
	(1.959)
employed in formal sector	1.387
	(2.786)
employed in business	8.964**
	(3.777)
employed in others	0
	(.)
education details	
literacy level	-6.019*
	(3.498)
number of years of education	1.422**
	(0.610)
completed primary education	25.59***
	(6.941)
employment status details	
employed	3.323
	(6.390)
unemployed	-34.36
	(25.22)
inactive	-1.776
	(3.669)
involved in housework	0
	(.)
household characteristics	
number of children	2.148**
	(1.039)
number of families	-1.536
	(1.499)
number of couples	2.440
	(7.969)
have electricity connection	2.782***
	(0.904)
age	0.194
	(0.134)
ownership of house	9.396***
	(3.025)
number of mothers	10.95***
	(3.478)
number of fathers	-11.28
	(9.526)
religion is Islam	-3.557***
	(1.064)
polygamous household	-116.4
	(71.46)
demographic characteristics of unmarried adults above age 16	
female age	0.0542
	(0.0844)
male age	0.303*
	(0.183)
ratio of males to females	2.163
	(3.832)
ratio of males in children with age less than 1 year	-3.887
	(3.870)
Psuedo R-square	0.3052
Observations	409

Notation for p-values *** is $p < 0.01$, ** is $p < 0.05$ & * is $p < 0.1$. Regressions include a constant term and all other controls are mentioned in the table.

Comments

Sonia Bhalotra's comments:

1. Beauty angle - The effects we see can be explained by fear/strategic move on the part of active people in marriage market about developing skin lesions which possibly reduces their chances of finding a mate.

2. Labor Market angle - Cognitive skills are possibly lower for people who have been exposed to arsenic which might lead to worse off labor market achievements. These are the people who tend to get married earlier.