



**SERIE  
DOCUMENTOS  
DE TRABAJO**

No. 275

Septiembre de 2021

# Sanitation and Marriage Markets in India: Evidence from the Total Sanitation Campaign

---

**Britta Augsburg**

**Juan P. Baquero**

**Sanghmitra Gautam**

**Paul Rodriguez-Lesmes**

# Sanitation and Marriage Markets in India: Evidence from the Total Sanitation Campaign \*

Britta Augsborg<sup>1</sup>, Juan P. Baquero<sup>2</sup>, Sanghmitra Gautam<sup>† 3</sup>, and Paul Rodriguez-Lesmes<sup>2</sup>

<sup>1</sup>Institute for Fiscal Studies

<sup>2</sup>School of Economics, Universidad del Rosario

<sup>3</sup>Department of Economics, Washington University in St Louis

September 26, 2021

Click [here](#) for an updated version

## Abstract

This paper analyses the marriage decisions of men and women, focusing on the added attractiveness of sanitation within the living arrangement, in rural India. We exploit district and time variation from the Total Sanitation Campaign (TSC) which increased sanitation by 6.6 percent among households with marriage eligible children and generated an exogenous increase in the composition of households with sanitation. Using data from the Indian Human Development household survey (IHDS) and district level census, we show that exposure to TSC increased the probability of marriage for men and women, from poorer households, by 3.8 pp and 6.5 pp respectively. The reduced form estimates incorporate both general equilibrium effects and heterogeneous program effects – two important components of equilibrium marital behavior. To decompose the overall policy impact on marriage market equilibrium we formulate a simple matching model where men and women match on observed and unobserved characteristics. Through model simulations, we show that cohorts within TSC exposed markets experienced a shift in marital gains both across matches but also within a given match. Specifically, the resultant sorting patterns display a marked gender asymmetry with an increase in marital surplus among matches where men are wealthier than their spouse, and a decrease in surplus where the wife is wealthier. Moreover, the increased access to sanitation for TSC exposed women implied a decline in their expected control over resources within the marriage.

**Keywords:** Empirical matching, Marriage markets, Sanitation, Sorting

**JEL Classification:** C78, D13, J12, O18

---

\*We thank Marcus Berliant, George-Levi Gayle, SangMok Lee and Bob Pollak for valuable discussions. Gautam gratefully acknowledges research support grant from the Weidenbaum Center at Washington University in St Louis. Augsborg gratefully acknowledge support of the Economic and Social Research Council (ESRC) Center for the Microeconomic Analysis of Public Policy at the IFS (grant reference ES/M010147/1 and the ESRC's support under the DFID/ESRC Growth Research Programme (grant reference ES/S01571X/1). [Preliminary, please do not cite or redistribute without permission.](#)

<sup>†</sup>Corresponding author e-mail: [sanghmitra.gautam@wustl.edu](mailto:sanghmitra.gautam@wustl.edu)

# 1 Introduction

Many social programs and policy interventions have focused on women in the expectation that women derive larger benefits from them. The female focus is specially evident in the context of Water Sanitation and Hygiene (WASH) interventions where several programs aimed to promote uptake of certain practices and facilities focus both directly and indirectly on women. Notable examples of such policies include the now popular ‘No Toilet, No Bride’ program implemented by the state government of Haryana, India.<sup>1</sup> The advertisement campaign promoted a narrative which encouraged families with marriage-age girls to ‘ask’ prospective suitors to provide, and hence if necessary construct, a latrine in the new home of the bride – thereby defining sanitation as a living arrangement. The emphasis on gender is also true for the Total Sanitation Campaign (TSC), an earlier version of the current Swachh Bharat Mission, the Government of India’s flagship sanitation policy.<sup>2</sup> While there is substantial evidence that the TSC and other sanitation interventions have increased sanitation take-up<sup>3</sup>, little is known with respect to the implications on the marriage market and impact on women.<sup>4</sup> Moreover if sanitation can be considered a type of living arrangement for a couple, policies aimed at increasing sanitation coverage have the potential to modify marriage market outcomes. This would be the case even if policies do not explicitly intertwine sanitation take-up with marriage markets. This paper focuses on quantifying the added marital value of having sanitation within the living arrangement, whether it entails a gain or a loss on the marriage market and, if so, for whom.

Our objective to quantify the added marital gain from sanitation on the marriage market necessarily intersects with two fundamental questions in the study of marriage markets: who marries whom? and how does the market clear? This inter-linkage poses notable challenges to the objective at hand. In particular, the determinants of marital sorting among different living arrangements cannot solely be attributed to taste for sanitation which may correlate with marital preferences and vary across the wealth distribution. In addition, the presence of general equilibrium effects and two-sided unobserved heterogeneity in a bilateral matching process raises non-trivial challenges in our ability to quantify the marital value of sanitation and the impact of sanitation policies that interact with the marriage market. A central contribution of this paper is to explicitly model the marriage market equilibrium in order to decompose the total

---

<sup>1</sup>Since its initial inception the Haryana program has become a model in the use of promotional campaigns aimed at men but with a focus on women (Cavill et al., 2018; Radtke, 2018).

<sup>2</sup> The TSC was implemented between 1999 and 2012. Approx 6%-15% of the program’s overall budget was allocated to gender focused campaigns in 2011 (CBGA, 2011).

<sup>3</sup>Literature focusing on the TSC in particular include Arnold et al. (2010); Barnard et al. (2013); Clasen et al. (2014); Hammer and Spears (2016); Patil et al. (2014); Pattanayak et al. (2009); Stopnitzky (2017). A more recent evaluation of the Government of India’s approach is conducted by Andres et al. (2020). The awareness creation approach used in the program, Community-Led Total Sanitation (CLTS) was shown effective in increasing sanitation investment in for example Tanzania (Briceno et al., 2017), Mali (Pickering et al., 2015), Nigeria (Abramovsky et al., 2019), Ghana (Crocker et al., 2016), Ethiopia (Crocker, 2016). A study by Guiteras et al. (2015) demonstrates effectiveness of CLTS with subsidy provision in the context of Bangladesh.

<sup>4</sup>A notable exception is Stopnitzky (2017), which we discuss below.

effect of programs like the TSC and thereby quantify the marital gains associated with sanitation on the marriage market.

Focusing on the marriage market and the marriage decisions of men and women, we contribute to the literature of marital sorting and the importance of sanitation within the marriage market in several notable ways.<sup>5</sup> First, we exploit quasi-random variation from the TSC that generated exogenous variation in the level of sanitation across India's districts and over time. This is important since sanitation and wealth often co-vary within a marriage type and are otherwise difficult to disentangle. Moreover, the presence of sanitation may be correlated with other unobserved marital preferences that are relevant for matching decisions. Therefore a regression analysis of marriage rates on the sanitation status of the spouse that does not account for endogeneity does not inform us about the added gain from sanitation within the marital living arrangement. Exploiting district and time variation in TSC exposure, coupled with multiple observations of the marriage market, allows us to overcome these identification challenges.

The empirical analysis uses data from the Indian Human Development Survey (IHDS) waves 2004 and 2011 and census data from the 2001 and 2011 at the district level. The IHDS survey is a household-level panel and provides rich micro-data on household composition, including individual demographics and social economic indicators. Additionally, we use TSC performance monitoring data aggregated by WSP (2011) at the district level. This information allows us to exploit district and time variation from the TSC which we show increased sanitation by 6.6 percentage points (ppts) among households with marriage eligible children and generated an exogenous increase in the composition of households with sanitation. Moreover, we show that exposure to TSC increased the probability of marriage for men and women from poorer households by 3.8 ppts and 6.5 ppts respectively.

Our second contribution is to develop and estimate an equilibrium model of the marriage market, where sanitation is modelled as type of living arrangement. We consider a transferable utility matching model where individuals match on both observables and unobservables (Choo and Siow, 2006). The model allows us to decompose the expected gains for men and women into marital sorting i.e., who matches with whom, vis-à-vis changes to the match surplus. In the model individuals on either side of the market, local to district and caste, are defined and matched over a discrete number of wealth types. To allow for composition effects and differential unobserved marital preference, men and women simultaneously choose their spouse and living arrangement. In other words, prospective spouses decide not only whether to marry, and whom to marry, but also the type of living arrangement. We extend the model to multiple markets and exploit variation over time (before and after TSC exposure), and across geographical areas with different populations and marriage market conditions, to achieve model identification. Importantly, in the proposed model marital surplus is specific to a living arrangement and a given couple. In addition,

---

<sup>5</sup>Several other programs are known to have effects on marriage markets. For instance, the introduction of divorce laws (Fernández and Wong, 2017; Greenwood et al., 2016; Rasul, 2006; Reynoso, 2018), incarceration of offenders (Charles and Luoh, 2010), compulsory schooling (Hener and Wilson, 2018), and school construction (André and Dupraz, 2019; Zha, 2019).

we allow the marital surplus that derives from sanitation to vary over the wealth distribution by introducing a match specific random component termed "love-for-hygiene". Doing so has the additional advantage of allowing for sorting on unobserved characteristics on both sides of the market but is match specific.

We use the model to estimate the total effect of the TSC on marital behavior of individuals living in districts with high TSC exposure in rural India. Specifically, the model estimates allow us to decompose and quantify the overall policy impact on: marital sorting, gains to marriage, and division of gains among partners. In addition, the model allows us to separately identify the contribution of marriage and marriage market preferences to the overall importance of sanitation. This is achieved by exploiting both observed matching patterns and associated marital gains over time and across markets under specific assumptions. Since we only use matching patterns, our analysis solely focuses on the marriage market. The focus on marriage markets is in our opinion of first order importance. Intuitively, what makes marriage markets relevant is the magnitude of the unobserved heterogeneity that may be indicative of underlying preferences for marriage and sanitation. Simply put, if there was no such variation, there would be no impact on sorting and thereby no role for the marriage markets. To the best of our knowledge the importance of sanitation on the marriage market has not been explicitly modelled. Stopnitzky (2017), who evaluates the 'No Toilet, No Bride', relies on a triple-difference specification, comparing the sanitation status of households with and without boys of marriageable age, living in the policy implementing state or other, comparable, Northern Indian states, before and after program implementation. Analysing heterogeneity in impacts by scarcity of women, he infers that impacts are driven by markets where women are scarce. However, instead of modelling the marriage market, the paper relies on an empirical proxy i.e., sex ratios, to draw inference on marriage market outcomes from sanitation take-up regressions.

The use of sex ratios as an empirical proxy for marriage market conditions has been widely popular in the literature.<sup>6</sup> Nevertheless, the use of own sex ratios in marriage rate regressions has a key limitation: it ignores the availability of alternatives. By allowing for spousal alternatives the Choo and Siow (2006) marriage matching function encapsulates both the general equilibrium and heterogeneous policy effect thereby capturing the overall equilibrium market response. While our paper, as Stopnitzky (2017), show that changes in marriage market conditions may affect sanitation take-up, the extent to which underlying market clearing transfers adjust is unknown. The estimates from the structural model help address this question and allow us to decompose the overall marriage market response to TSC exposure. Overall, we argue that a careful analysis of the marriage market is an integral component of the overall policy impact on existing marriages, and to infer the determinants of well-being of men and women.<sup>7</sup>

Through model simulations we show that the general equilibrium effects induced by the TSC had unintended consequences both across different sub-markets defined by wealth types but also within a match among spouses. We find that access to sanitation makes it more attractive to be in a marriage for

---

<sup>6</sup>Abramitzky et al. (2011), Angrist (2002), Charles and Luoh (2010), Gupta (2014).

<sup>7</sup>The importance of marriage markets and changes in sorting was first emphasised by Lundberg and Pollak (1996).

both men and women, inducing some of them to change their status from being single to being married. We find in particular that the final match is rearranged, so that men of low and median wealth status become more likely to marry women of low wealth status. The new sorting pattern results in a higher match surplus for couples where men are wealthier than women, and in a lower match surplus for most remaining cases. Moreover, we show through our simulations that these new sorting patterns would have increased low wealth expected surplus in areas with no shortage of women, whereas it would have unambiguously decreases the expected surplus in areas with shortage of women. These latter findings effectively imply a decline in women’s control over resources in the household. Instead of preventing the unintended consequences, the presence of less women per men ended up exacerbating the loss of control of resources. Therefore, scarcity per se does not imply that female-targeted programs necessarily result in better outcomes for women.

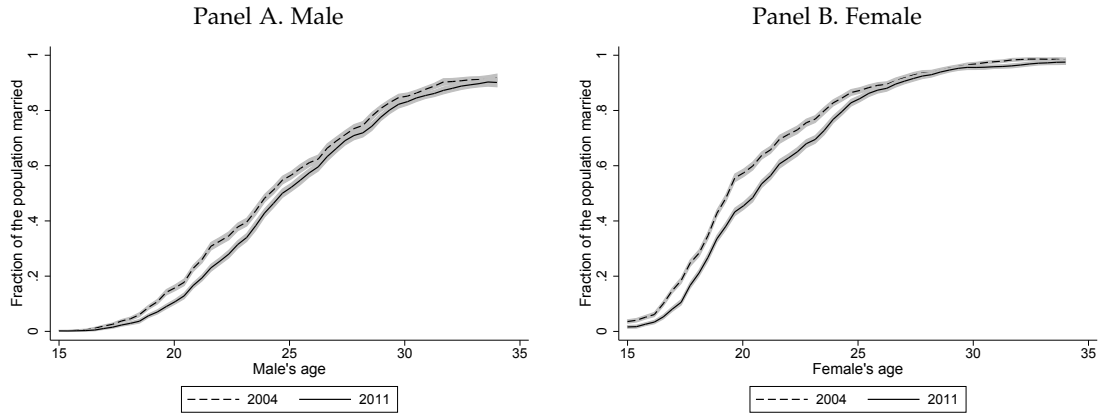
The rest of the article is organized as follows. Section (2) describes the data and characteristics of the Indian marriage market, which make it a particularly apt context for our analysis. Section (3) describes the TSC and presents the effect of the policy on sorting patterns and sanitation take-up among households with marriage eligible children. The theoretical framework is developed and estimated in Section (4). Our main empirical findings are presented and discussed in Section (5). Section (6) concludes.

## **2 Context**

### **2.1 Marriage in India**

Marriages in India are fundamentally shaped by social norms and by caste (and geographical) endogamy. Social norms include that girls must marry, marriages tend to be arranged, and newlyweds are expected to live with the groom’s parents (patrilocality). For women, remaining single carries a heavy social stigma in India with the consequence that marriage is near universal for women, occurs early in life and is rarely reversed. The median age of brides was 18 years in 2011 (one year higher than in 2004, Table 1) and by the time women reach their late twenties, their likelihood of being married approaches 1 (Figure 1). Men on the contrary get married at a later age and their likelihood of marriage levels off at just over 90% when they reach their mid-thirties. Gupta (2014) discusses how male marriage rates show greater variation by region and over time than female marriage rates and depend on the availability of brides.

Figure 1: Fraction of respondents who are married by age, gender, and year of survey



Note. Own calculations using data from the IHDS waves 2004 and 2011. Sample includes all married men and women 15 to 34 years.

Decision-makers about getting married, and whom to marry, are typically neither the bride nor the groom, but their guardians. Especially in view of the stigma of having an older, unmarried daughter, it becomes a parents' primary duty to marry off their daughter. According to the IHDS questionnaire for wives, only 35% of women reported to have participated in the matching decision in 2011 (37% in 2004, Table 1). The matching process is dominated by a concern for status and economic outcomes, in that parents seek a partner that is of higher or equal caste, educational qualification, or wealth. Especially caste continues to play a crucial role on determining marriage outcomes: over the period 2004-2011, less than 6% of marriages took place across caste (Table 1). This share has remained stable despite strong economic growth. Banerjee et al. (2013) show that underlying this statistic is a strong preference and low cost to marry someone of the same caste, and recent genetic analyses have established that these patterns of endogamous marriage have been in place for over 2,000 years (Moorjani et al., 2013). Also, education is seen as a particularly desirable quality for grooms (Banerjee et al., 2013) and perceived relevant by parents for their daughter to find a better marriage match<sup>8</sup> However, while such individual characteristics like education and looks (Banerjee et al. (2013)), do play a role, it is household characteristics that dominate the matching choice: in addition to caste especially wealth (such as land and income), have been shown to matter more for matching than individual characteristics in the Indian context (Borker et al., 2018; Rao, 1993).<sup>9</sup>

Once married, due to patrilocal norms, more than 90% of newly-weds begin their married life co-residing with the groom's parents (Desai and Andrist (2010)). In line, just 12% of women captured in the

<sup>8</sup> The quality of the match though deteriorates quickly with her age on leaving school (Adams-Prassl and Andrew, 2019), further pushing the existence of unofficial age limits by which women are expected to be married (Jaggi, 2001).

<sup>9</sup> Assortative matching based on characteristics such as income, education, physical characteristics, and age have been similarly established in other contexts (see for example Becker and Becker (2009); Chiappori et al. (2018,1,1); Choo and Siow (2006); Grossbard (1993); Hitsch et al. (2010); Pencavel (1998)).

IHDS grew up in the same village as their husband (Table 1). At the same time, spouses are chosen from geographically close areas - the bride's village of origin is on average a distance of 3 hours away with locally common transportation. Only 4% of total female migration (including for education and other non-marriage purposes) is inter-state, and 9.8% inter-district (but intra-state) (Kone et al., 2018), implying that the vast majority of brides' destinations are intra-district.

Table 1: Descriptive statistics IHDS sample

	2004	2011
<b>Demographics:</b>		
<i>Age at marriage (women):</i>		
p25	15.0	16.0
p50	17.0	18.0
p75	19.0	20.0
mean	17.1	18.0
<i>Age gap (male-female):</i>		
p25	0.0	0.0
p50	1.0	1.0
p75	4.0	3.0
mean	2.1	1.8
<b>Marriage market characteristics:</b>		
Bride participated in matching decision (%)	0.374	0.351
Newly-weds living with groom's parents (%)	0.924	0.974
Wife from same village as groom (%)	0.119	0.105
Hours to natal home (wife)	3.287	3.065
Inter-caste marriage (%)	0.049	0.060
<b>Marriage associated costs:</b>		
Gifts occurrence (0: never 1: usually)	0.448	0.364
Share of household income spent by groom	1.7	1.9
Share of household income spent by bride	2.6	2.9

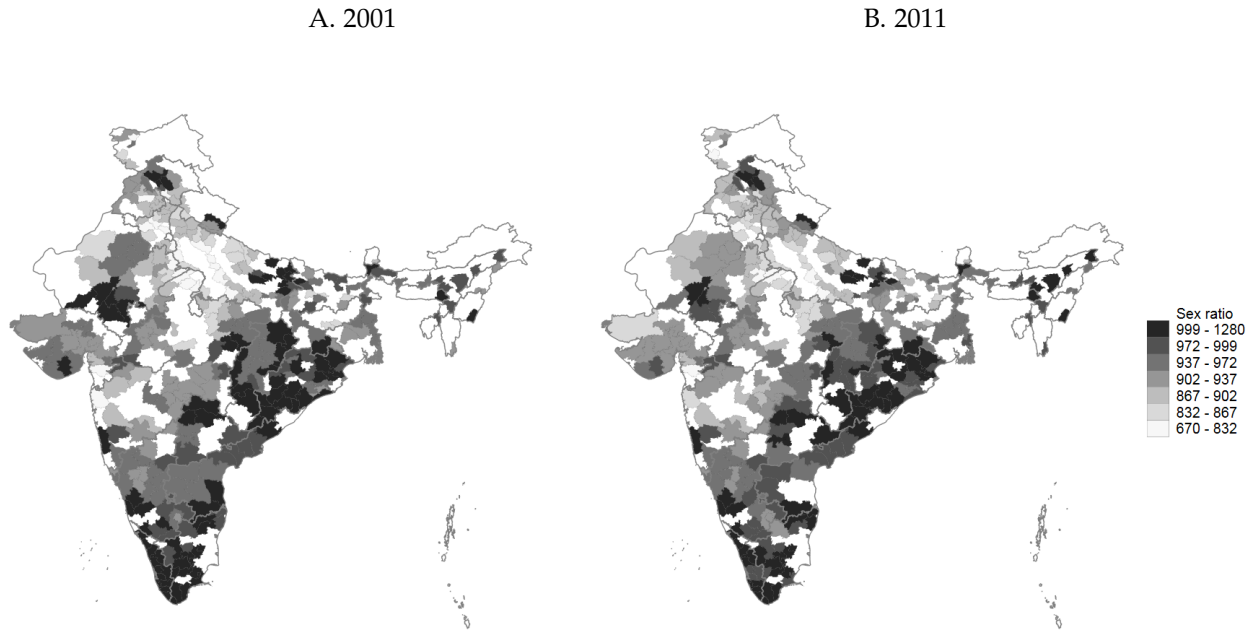
*Notes:* Own calculations using data from the IHDS waves 2004 and 2011. Sample includes men and women aged 15 to 34 years and married at the time of the survey.

## 2.2 Indian marriage markets and implications for analysis

In many ways these characteristics of marriage in India, makes the Indian marriage market look more like a market than it does in other parts of the world. Importantly, the longstanding practise of intra-caste marriage, as well as the strong tendency to marry within ones birth district imply not just one, but many distinct marriage markets, as most recently highlighted by Beauchamp et al. (2017). This feature provides

us with multiple observations of markets, which we will make use of in our analysis.

Figure 2: District-level sex ratio



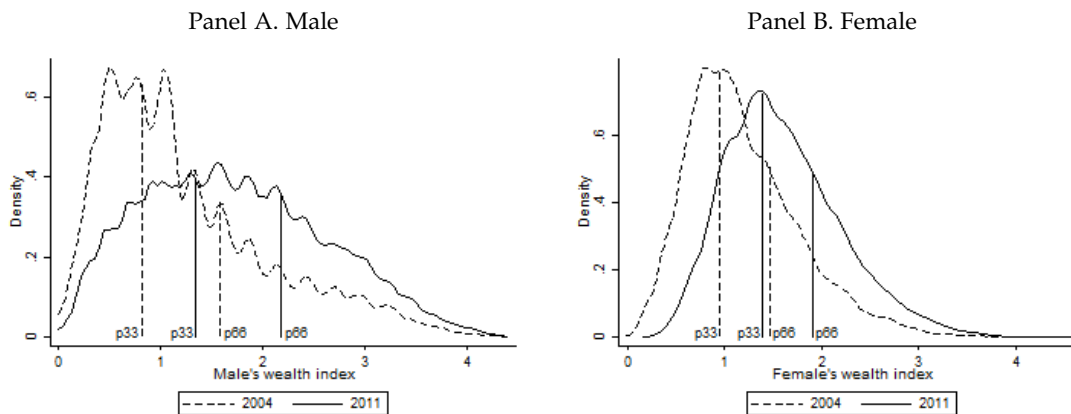
Notes: Source: Indian census data, 2001 and 2011. Sample includes 15-34 year-olds.

A distinguishing characteristic of Indian marriage markets is substantial variation in the degree of gender balance. Considerable attention has been paid to understanding the impact of a change in the sex ratio on marriage outcomes. In contrast to early contributions (Cox (1940); Easterlin (1961)), more recent studies in contexts other than India are able to take account of reverse causality between sex ratios and marriage market outcomes (Abramitzky et al. (2011); Angrist (2002); Charles and Luoh (2010)), demonstrating that gender imbalances tend to lead to improved marriage outcomes for the scarcer sex. While we are not aware of any study making use of exogenous variation in sex ratios in India, studies analysing the co-existence of sex selection and dowries in India (Basu (1999); Borker et al. (2019); Das Gupta et al. (2003); Murthi et al. (1995)) provide evidence in support of a link between sex selection and marriage markets. We consider sex ratios as a given market characteristic in our analysis that varies across districts. Figure 2 highlight the great disparity in sex ratios across the country ranging from equally balanced to only two women for every three men in the North of the country. We will consider differences between districts of high and low sex ratio in our analysis, where we define high sex ratio as a district where there are at least 999 women per 1000 men aged 15-34; nearly 20% of the sample, both in 2004 and 2011.

Other aspect of consideration is the presence of dowries. Table 1 shows that more than a third of households of the sample report that 'gifts' at marriage are usual in their context. Yet, the illegality of

dowry in the Indian context makes accurate measurement a challenge. Moreover, Botticini and Siow (2003) warns against interpreting dowries as the sole form of transfer that clears the marriage market. As such for our analysis, the dowry response can be thought to be encapsulated within the market clearing transfer.

Figure 3: Wealth index distribution by gender and over time



*Note.* Own calculations using data from the IHDS waves 2004 and 2011. The wealth index, defined at household level, is the sum of self-reported assets at the household, which include: bicycle, sewing machine, generator set, grinder, motorcycle, TV, air cooler, clock, electric fan, chair/table, cot, telephone, cell phone, refrigerator, pressure cooker, car, Air conditioner, washing machine, computer, credit card, two clothes, footwear. For males it corresponds to the household information, while for female its is the prediction based on observed characteristics: age, education, literacy, English knowledge, caste and state. The prediction comes from a model for single women where the dependent variable is the wealth index of their family.

## 2.3 Wealth and sorting

We use a measure of wealth as our matching variable of interest between the bride and the groom. The choice of to use wealth as the key attribute is reinforced by the strong positive correlation between wealth other marriage market relevant characteristics.<sup>10</sup>

Our proxy for wealth is based on self-reported asset ownership using principal component analysis (PCA).<sup>11</sup>, but will take different approaches to construct pre-marital wealth for the wife and the husband, driven by the practice of patrilocal exogamy, i.e. the bride to move into the groom’s family’s house. This almost universal practice implies that we can proxy the groom’s pre-marital wealth by using information on asset ownership as reported for the household the couple resides in. For the bride’s pre-marital wealth we instead use a predicted asset score which is based on how observable characteristics of single women (age, education, literacy, English knowledge, caste and state) predict the asset index of their families. Figure 3 shows the distribution of the asset index by year for both genders. In line with significant

<sup>10</sup>See (Borker et al., 2019) for a discussion of the importance of wealth in a context with marriage market frictions.

<sup>11</sup>The assets included in the component are: bicycle, sewing machine, generator set, grinder, motorcycle, TV, air cooler, clock, electric fan, chair/table, cot, telephone, cell phone, refrigerator, pressure cooker, car, Air conditioner, washing machine, computer, credit card, two clothes, footwear.

economic growth over this period, we observe a rightward shift in the distribution between 2004 and 2011. In our model, we consider a discrete version of this variables based on the terciles defines over gender and year. The vertical lines of the line indicates the limits of these categories.

We calculate the Pearson's correlation coefficients to characterise the degree of marital matching along wealth. We find a correlation coefficient of 0.679 between the groom's and bride's wealth index in 2005 and of 0.676 in 2011. A value of 1 would indicate perfect correlation, implying that these statistics confirm a high degree of matching along wealth in the Indian context. If we consider instead the discrete variable (low, medium, high wealth), Kruskal's Gamma correlation coefficients present similar results: a rank correlation of 0.736 in 2005 and 0.724 in 2011.

### 3 Sanitation policy and marriage markets

#### 3.1 The Total Sanitation Campaign

The Indian government has devoted large amounts of resources to make India "open defecation free" through its flagship Swachh Bharat Mission (Clean India) policy introduced in 2014, following similar previous national policies. First efforts go back to the Central Rural Sanitation Program (CRSP) in 1986, replaced by the Total Sanitation Campaign (TSC) in 1999 and the Nirmal Bharat Abhiyan (NBA) in 2012, which in 2014 evolved into the current Swachh Bharat Mission.

The TSC introduced continued a focus on constructing new private household latrines with subsidies for vulnerable households, but added a component with Information, Education and Communication (IEC) activities around sanitation and introduced in 2003 the provision of a financial award to communities that achiev Open Defecation Free (ODF) status, the Nirmal Gram Puraskar (NGP), or Clean Village Prize.<sup>12</sup> The state and central governments have a facilitating role in the implementation of the policy, that takes the form of framing, enabling, providing financial and capacity-building support, and monitoring progress (WSP, 2011).

A strong upward trend in rural sanitation coverage has been documented over the course of program implementation, however, with significant disparities across states and districtsWSP (2011). The study illustrates the unequal performance of TSC among Indian districts through a global monitoring index ("The Grand Score"), which weights performance based on eight key performance indicators.<sup>13</sup> Figure 4 shows the geographic variation of this indicator. 54 districts (16% of the IHDS sample) receive a score of

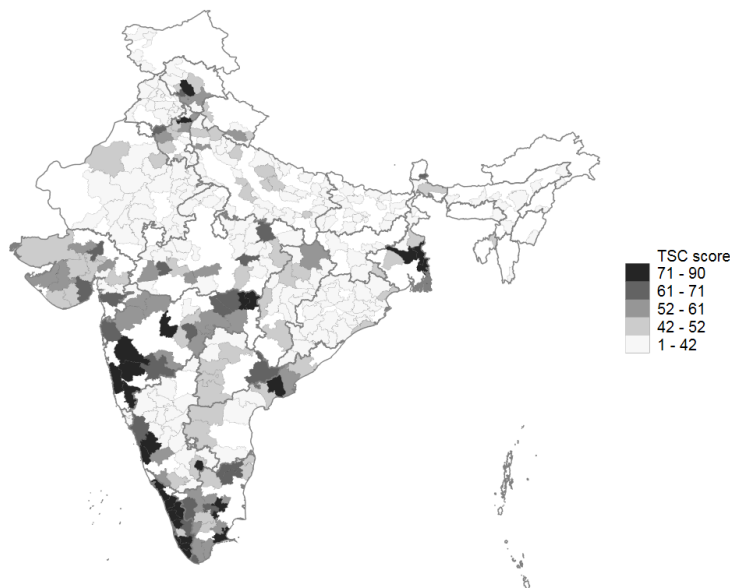
---

<sup>12</sup> The NBA program further introduced the concept of community-led total sanitation (CLTS), as part of the education and communication activities on sanitation. The SBM expanded to urban areas, among other changes.

<sup>13</sup>The eight indicators cover inputs, outputs, process and outcomes and each indicator is allocated a maximum score. The first indicator is an input (% of TSC budget spent) for which the maximum score is 5; the next two are outputs (% of household toilet targets achieved, 15 scores, and % of school sanitation target achieved, 10 scores); there are three process indicators with maximal 10 scores for each (financial efficiency (cost per NGP community), average population per community (gram panchayat), success rate of NGP applications; and two outcomes indicators (number of NGP panchayats, 30 scores, and % of NGP panchayats, 10 scores).

61 or above (the two darkest colors in the map), which we use as a cut-off for high-performing districts. These high-performing districts tend to have a better educated population, both women and men marry on average at later ages (one and two years respectively), and bride and spouse are more often from the same village (Appendix Table 10).

Figure 4: Grand Score TSC implementation performance by district



*Note.* Data source: WSP (2011). The “Grand Score” is an index that weights performance based on eight key TSC performance indicators, covering inputs, outputs, process and outcomes and each indicator is allocated a maximum score. Scores add up to a maximum of 100 and are available at the district level.

### 3.2 Toilet take-up in marriage markets

Several studies have analysed the effectiveness of this policy (Arnold et al. (2010), Clasen et al. (2014), Hammer and Spears (2016), Patil et al. (2014), Pattanayak et al. (2009), Stopnitzky (2017)), providing evidence of significant increases in toilet ownership, which can be considered high compared to many other evaluated sanitation policies Garn et al. (2017).<sup>14</sup> We analyze whether we can replicate the effectiveness of the TSC policy in increasing household toilet ownership when focusing on cross-Indian households that are likely to enter/be in the marriage market between 2004/05 and 2011, and consider differential impacts by underlying market conditions, as indicated by the district-level sex ratio. We use a repeated

<sup>14</sup>While its impact on child health could not be established (Clasen et al. (2014)), increases in cognitive scores due to the policy have been identified (Orgill-Meyer and Pattanayak (2020); Spears and Lamba (2016)).

cross-section of all households where the household head has at least one son or daughter aged 15-34.<sup>15</sup> We estimate the following difference-in-difference specification:

$$Toilet_{it} = \pi_0 dT_t + \pi_1 TSC_i + \pi_2 TSC_i \cdot dT_t + \mathbf{Z}_{it}\Lambda + e_{it}, \quad (1)$$

where  $Toilet_{it}$  is toilet adoption in the household  $i$  in survey wave  $t$ . We define  $dT = 1$  for survey wave 2011, and  $dT = 0$  for 2004/05;  $TSC = 1$  for households located in districts with a high TSC performance, which corresponds to a high grand score above or equal 61, and 0 otherwise. The matrix  $\mathbf{Z}$  represent the control variables: age and marital status of the oldest marriageable female (male); the wealth index of the female's (male's); household size; education level of the household head (no education [base], primary, incomplete secondary, secondary, and above secondary); caste (Brahmin [base], High caste, Other backward caste, Dalit, Adivasi, Muslim, Sikh - Jain); an indicator variable for the rural zone; and fixed effects at the state level.  $e_{it}$  corresponds to the unobserved component. Standard errors are clustered at the district level. The coefficient of interest is  $\pi_2$ .

Table 2 shows the results, considering separately household where the oldest member within the marriage age range is male (columns 1–3) and where it is female (columns 4–6). The impact of high TSC-performance ( $dT \times TSC$ ). We find that for households where the eldest individual in the age range 15-34 is a male, toilet ownership increases by 6.57 percentage points (pppts) due to living in a high-performing TSC district (column 1). The result is significant at the 5% level. For females, the estimated coefficient is smaller at 3.85, but statistically insignificant (column 4). In Appendix 12, we show that these results are robust to alternative definitions of shortage and TSC exposure. These impacts hide heterogeneity with larger impacts where women are scarce (low sex ratio, columns 3 and 6), and even negative (significant at the 10% level) where women are abundant. In low sex ration districts, households where the eldest marriageable child of the household head is male, we find a 13.8pppts increase in toilet ownership, and where it is female, one of 7.7 pppts. These estimates are significantly different from each other, suggesting that households with a marriageable boy tend to react more to the TSC campaign. These results are in line with Stopnitzky (2017) in the context of the 'No Toilet No Bride' campaign in Haryana. The author found that households with men of the marriageable age were more likely to take-up sanitation when exposed to the campaign than households with women of similar ages.

---

<sup>15</sup>Appendix Table 11 demonstrates that results are in line when all households with a 15-34 year-old member are included in the sample.

Table 2: TSC impact on Sanitation for households with an oldest marriageable son (daughter)

Variables	Oldest marriageable son household			Oldest marriageable daughter household		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	High sex ratio	Low sex ratio	All	High sex ratio	Low sex ratio
$dT$	-0.0345*** (0.0112)	-0.0326 (0.0334)	-0.0363*** (0.0119)	-0.0220* (0.0129)	-0.0118 (0.0287)	-0.0253* (0.0147)
$TSC$	0.0547 (0.0411)	0.150** (0.0625)	0.0143 (0.0480)	0.0871** (0.0391)	0.110 (0.0735)	0.0601 (0.0459)
$dT \times TSC$	0.0657** (0.0333)	-0.0849* (0.0439)	0.138*** (0.0372)	0.0385 (0.0296)	-0.0344 (0.0527)	0.0769** (0.0368)
Observations	27,993	5,784	22,209	10,437	2,619	7,818
$R^2$	0.416	0.482	0.407	0.404	0.460	0.396

*Notes:* Own calculations using data from the IHDS waves 2004 ( $dT = 0$ ) and 2011 ( $dT = 1$ ). The sample consists of households of all single and married females (males) aged 15 to 34 at the survey time. Households are classified according to the gender of the eldest son (columns 1 to 3) or daughter (columns 4 to 6) or the household head within such age range.  $TSC = 1$  correspond to a grand score for implementation of 61 or above, and  $TSC = 0$  to those with a score of 61 or above. The grand score per district is taken from WSP (2011). The high sex ratio corresponds to districts with at least 999 women per 1000 men in the age range 15-34 (columns 2 and 5), while the low sex ratio is districts with less than 999 (columns 3 and 6). District level sex ratio information was computed using data from the population census 2001 and 2011. Apart from the coefficients presented in the table, as controls, we consider the age and marital status of the individual for whom the household is in the sample; the wealth index of the household; household size; education level of the household head (no education [base], primary, incomplete secondary, secondary, and above secondary); caste (Brahmin [base], High caste, Other backward caste, Dalit, Adivasi, Muslim, Sikh - Jain); an indicator variable for the rural zone; and fixed effects at the state level. Clustered at the district level standard errors in parentheses. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

In Table 3 we repeat the analysis by whether a household falls within a high- or low-wealth category, based on the median of the wealth index distribution for each wave. We find that the impact on toilet ownership is higher for households at the top of the wealth distribution, living in areas with a scarcity of women. For high-wealth households with a boy in line to get married the impact is 20.0 ppts compared to 12.1% for low-wealth households with a groom-to-be; for high-wealth households with a girl in marriageable age the impact is 11.6% compared to an insignificant 2.85% for low-wealth households with a bride-to-be.

Table 3: TSC impact on Sanitation for Households with an oldest marriageable son (daughter) by wealth and sex ratio

Variables	Oldest marriageable son household				Oldest marriageable daughter household			
	High sex ratio		Low sex ratio		High sex ratio		Low sex ratio	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	High wealth	Low wealth	High wealth	Low wealth	High wealth	Low wealth	High wealth	Low wealth
$dT$	-0.0155 (0.0335)	-0.0522 (0.0436)	-0.0557*** (0.0183)	-0.0300** (0.0128)	-0.0416 (0.0351)	-0.0651 (0.0416)	-0.0134 (0.0186)	-0.0637** (0.0247)
$TSC$	0.159** (0.0659)	0.102 (0.0764)	-0.0592 (0.0515)	0.0711 (0.0631)	0.102 (0.0750)	0.0982 (0.0764)	-0.0119 (0.0499)	0.183** (0.0711)
$dT \times TSC$	-0.0713 (0.0437)	-0.0431 (0.0669)	0.200*** (0.0451)	0.121* (0.0627)	-0.00346 (0.0511)	-0.0601 (0.0940)	0.116** (0.0495)	0.0285 (0.0457)
Observations	3,294	2,490	12,349	9,860	1,655	964	5,173	2,645
$R^2$	0.373	0.337	0.273	0.295	0.386	0.344	0.327	0.332

Notes: Own calculations using data from the IHDS waves 2004 ( $dT = 0$ ) and 2011 ( $dT = 1$ ). The sample consists of households of all single and married females (males) aged 15 to 34 at the survey time. Households are classified according to the gender of the eldest son (columns 1 to 4) or daughter (columns 5 to 8) or the household head within such age range.  $TSC = 1$  correspond to a grand score for implementation of 61 or above, and  $TSC = 0$  to those with a score of 61 or above. The grand score per district is taken from WSP (2011). The high sex ratio corresponds to districts with at least 999 women per 1000 men in the age range 15-34 (columns 1, 2, 5 and 6), while the low sex ratio is districts with less than 999 (columns 3, 4, 7 and 8). District level sex ratio information was computed using data from the population census 2001 and 2011. High wealth correspond to individuals whose asset index is above the 50 percentile of the entire country distribution per wave (columns 1, 3, 5 and 7). Those households below such cutoff are classified as low wealth (columns 2, 4, 6 and 8). Apart from the coefficients presented in the table, as controls, we consider the age and marital status of the individual for whom the household is in the sample; the wealth index of the household; household size; education level of the household head (no education [base], primary, incomplete secondary, secondary, and above secondary); caste (Brahmin [base], High caste, Other backward caste, Dalit, Adivasi, Muslim, Sikh - Jain); an indicator variable for the rural zone; and fixed effects at the state level. Clustered at the district level standard errors in parentheses. Significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

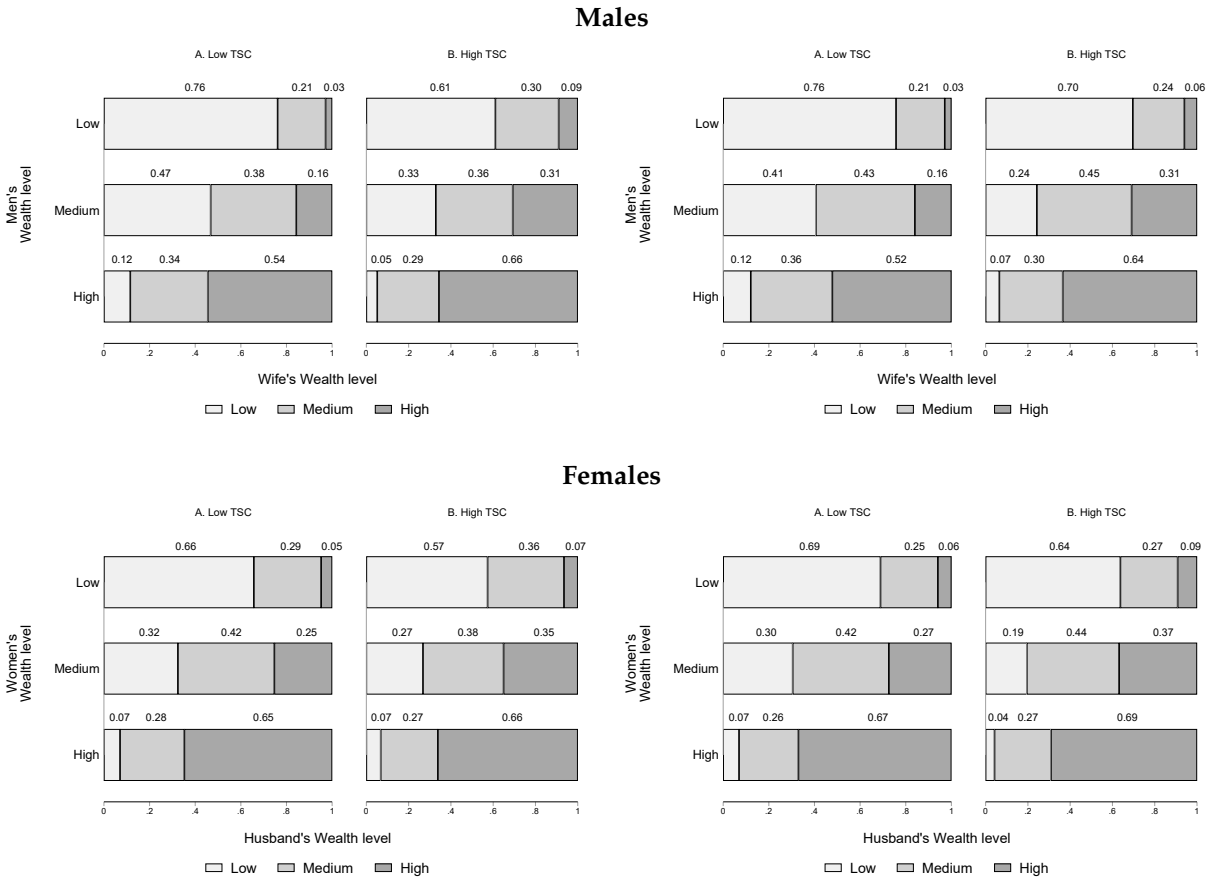
An important take-away from the analysis presented in this section is that the TSC program likely impacted the marriage market equilibrium: the extent of the uptake is a function of marriage market characteristics and attributes such as the scarcity of women and wealth. In addition Tables (2) and (3) provide indirect evidence on the importance of incorporating heterogeneous policy effects of the TSC in our analysis of the marriage market equilibrium. We unpack this link in Section (5), based on the structural model presented in Section (4). Before doing so, we once more turn back to marital sorting.

### 3.3 Marital Sorting

We discussed in Section 2.3 the high degree of assortative matching along wealth in the Indian marriage market, which we present here graphically and by high and low TSC performance (Figure 5). The figure clearly shows that spouses tend to marry partners of the same wealth category. For example, in 2011, 76% of men with low pre-marital wealth living in low-performing TSC areas were matched with a woman

whose pre-marital wealth also fell into the low-wealth category; in high-performing TSC areas the percentage was 61. In 2011, the percentage of low-wealth matches in low-performing TSC areas remains at 76% and increases from 61% to 70% in high-performing TSC areas. Similar types of patterns, with same-wealth levels dominating matches are found for women, both in high- and low-performing TSC areas.

Figure 5: Marriage patterns of women and men over time across TSC exposure groups



Notes: Own calculations using data from the IHDS 2005 and 2011.

While these marriage rates provide indirect evidence of the importance of marital matching along wealth in both high- and low-performing TSC areas, they do not take into account general equilibrium effects such as changes in the number of available matching opportunities in the market – notably, the large increase in the number of households with sanitation, particularly in high-performing TSC areas.

Once taking general equilibrium effects into account, as we will do in the next part of the paper, we are not only able to quantify how individual marital outcomes are related to individual and couple characteristics, e.g. wealth and sanitation, but also subsequent marriage market general equilibrium response which are non-trivial. In such cases any reduced form analysis in the context of evaluating marriage market outcomes would conflate individual and market effects as individual of different types interact within

a marriage market.<sup>16</sup> This issue is independent of the non-overlapping nature of the marriage markets and would persist within each geographic and caste isolated market cluster.

## 4 Theoretical Framework

We use a structural model to rationalise what we see in the data in terms of sorting, and to quantify the resulting impacts on marital surplus and division of surplus, both of which are determined from the marriage market equilibrium and are thereby endogenous. For example, if TSC increases the gains from participation in the marriage market for both men and women of a given match type this would imply an increase in the total match specific surplus derived from that match. The impact on the surplus division, however, is less straightforward and would depend on the differential impacts by gender. In what follows we employ the Choo and Siow (2006) marriage matching function which in addition to being nonparametric allows for general equilibrium effects.

### 4.1 Marriage Matching and Gains from Marriage

To decompose the impact of TSC on marriage market outcomes and rationalize the sorting patterns observed in the data we employ a simple matching model.<sup>17</sup> We begin by describing the Choo and Siow (2006) transferable utility model of a marriage market, where the unique Indian context allows for multiple observations of the relevant marriage market  $g$  defined over caste and region.<sup>18</sup> In particular we consider two large finite populations of male and female populations to be matched on the basis of their wealth type. Men and women within each group  $g$  differ only in their wealth defined by a wealth index  $I(W_\kappa)$  and  $J(W_\kappa)$  respectively. The distribution of wealth has a finite support  $K$  and each individual belongs to a finite set of wealth classes  $W = \{W_1, \dots, W_K\}$ . Let  $\mathbf{M}$  ( $\mathbf{F}$ ) be the vector of available male (and female) types, the  $I^{th}$  and  $J^{th}$  element of which denoted the measure of different types of men  $m_I$  and women  $f_j$  respectively. Each man  $i \in I$  and women  $j \in J$  must chose a type of partner. The choice of remaining unmatched is captured by augmenting the type space for men and women, e.g.  $I = \{0, 1, \dots, I, \dots, \mathcal{I}\}$ , where choosing 0 by woman  $J$  corresponds to remaining single and vice versa for males.<sup>19</sup>

The total marital gains generated by a match  $\pi_{ij}$  between  $i \in I$  and  $j \in J$  is given by

$$\pi_{ij} = \pi_{IJ} + \varepsilon_i^I(J) + \eta_j^J(I) \quad (2)$$

---

<sup>16</sup>The limitations of using a regression framework on individual level data to account for general equilibrium effects are well documented. For example, Angrist (2002) finds the causal effect of changes in the sex ratio on the male marriage rate to be inconsistent with that found for the female marriage rate.

<sup>17</sup>This characterization is similar to that undertaken by Chiappori et al. (2009,1); Choo and Siow (2006); Iyigun and Walsh (2007).

<sup>18</sup>We subsume all  $g$  subscript notation and return to this feature for the estimation of the model.

<sup>19</sup>Female type space  $J = \{0, 1, \dots, J, \dots, \mathcal{J}\}$

where in addition to the systematic match surplus  $\pi_{IJ}$ , a match between a type  $I$  male and type  $J$  is also subject to a vector of gender specific preference shocks. The preference heterogeneity components  $\varepsilon$  and  $\eta$  are additive and separable in the gains function and depend only on the partner's type, not his or her exact identity. The matrix  $\Pi = (\pi_{IJ})$  denotes aggregate surplus. The systematic gains to marriage  $\pi_{IJ}$  in equation (2) also referred to as the marital surplus will be the main object of interest in our empirical analysis.

## 4.2 Adding Sanitation to Gains from Marriage

Since we only observe the sanitation status together with the matched status of individuals we model sanitation as a type of living arrangement. In other words a living arrangement establishes types of marriages which individuals can choose to sort into and thereby allows us to incorporate sanitation without additional assumptions on the underlying preferences.<sup>20</sup> To account for how the presence of sanitation in the match affects the gains to marriage we follow Choo and Siow (2006), and measure the numerator of Equation (2) alternatively by the number of marriage living arrangements  $l$  with and without sanitation  $l \in \{T, NT\}$ .

In particular, as discussed, the marriage rates depicted in Figure (5) do not take into account changes in the number of available matching opportunities in the market – notably, the large increase in the number of households with sanitation. An increase in marriages with sanitation could occur either because those marriages are desirable relative to marriages without sanitation or, mechanically, because there are more available households with sanitation. To take this change in the pool of available marriage types and available men and women into account, we first follow the methodology introduced by Choo and Siow (2006) to generate a marriage matching function for each marriage types  $l$ . Specifically, we incorporate marriage types or living arrangements which delivers an additional set of equilibrium relationships which define the marriage matching functions to calculate the gains to marriage. By allowing for different marriage types each individual can decide whether to marry in an arrangement with or without sanitation or to remain unmatched. For each type  $I$  man to match with a type  $J$  woman in living arrangement  $l$ , he must transfer to her a part of his utility that he values at  $\tau_{IJ}^l$ . The woman values the transfer as  $\tau_{IJ}^l$  which maybe positive or negative.

## 4.3 Marital Preferences over Types

Let the pay-off of a man  $i$  of wealth type  $I$  who matches with a woman of wealth type  $J$  in a living arrangement  $l$  with own idiosyncratic preferences  $\varepsilon_{iIJ}$  be given by

$$U_{iIJ}^l = \tilde{u}_{IJ}^l - \tau_{IJ}^l + \varepsilon_{iIJ} \quad l \in \{T, NT\} \quad (3)$$

---

<sup>20</sup>Incorporating unobserved heterogeneity in a two-sided matching model is non trivial.

where like in Choo and Siow (2006)  $\tilde{u}_{IJ}^l$  denotes the systematic gross returns common to all males of type  $I$  matching to a female of type  $J$  in living arrangement  $l$ . The idiosyncratic component of male marital preference measures the departure of each individual male  $i$ 's pay-off  $U_{iIJ}^l$  from the systematic component which is common to all male type  $I$  who marry females of type  $J$ . Similarly the pay-off for a male  $i$  of type  $I$  who remains unmatched ( $J = 0$ ) is given by:

$$U_{iI0} = \tilde{u}_{I0} + \varepsilon_{iI0} \quad (4)$$

where  $\tilde{u}_{I0}$  denotes the systematic gain that is common to all type  $I$  males from remaining unmatched. In the empirical specification we relax the assumption that men and women face similar gains from remaining unmatched. Moreover, the inclusion of living arrangements, allows us to specify the joint choice of partner and living arrangement with or without sanitation. Specifically, the underlying decision that would govern each male type  $I$  behaviour maximizes the objective function given by Equation (5).

$$U_{iI} = \max_{j,l} \{U_{iI0}, U_{iI1}^T, \dots, U_{iIJ}^T, \dots, U_{iI\mathcal{J}}^T, U_{iI1}^{NT}, \dots, U_{iIJ}^{NT}, \dots, U_{iI\mathcal{J}}^{NT}\} \quad (5)$$

The random utility pay-offs for women have a similar form. Specifically, the choice specific pay-off of a woman  $j$  of wealth type  $J$  who marries a man of wealth type  $I$  with own idiosyncratic preferences  $\eta_{jIJ}$ , and  $\tilde{v}_{0J}$  systematic gain for type  $j$  women, be given by:

$$V_{jIJ}^l = \tilde{v}_{IJ}^l + \tau_{IJ}^l + \eta_{jIJ} \quad l \in \{T, NT\} \quad (6)$$

$$V_{j0J} = \tilde{v}_{0J} + \eta_{j0J} \quad (7)$$

The pay-off specifications above are analytically convenient allow us to incorporate sanitation without additional ad-hoc assumptions. The idiosyncratic marital preference for individual men  $i$  and women  $j$  are assumed to be independently and identically distributed with a type I extreme value distribution. In addition the idiosyncratic marital preference only depends on their spousal type  $I$  or  $J$  and not their individual identity  $i$  or  $j$ . It should be stressed that in our framework, these idiosyncratic, additively separable shocks are the only source of unobserved heterogeneity.

#### 4.4 Marriage Market Equilibrium

The matching market clears when, given equilibrium transfers  $\tau_{IJ}^l$ , the demand by type  $I$  males is equal to the supply of females of type  $J$  for all  $(l, I, J)$

$$(\mu_{IJ}^l)^d = (\mu_{IJ}^l)^s = \mu_{IJ}^l; \quad l = (T, NT) \quad (8)$$

substituting (8) into the quasi demand and supply equations above yields the market equilibrium moment

condition where the measure of type  $I$  men who chose to marry type  $J$  women equals the measure of type  $J$  women who chose to marry type  $I$  men. The equilibrium distribution of marriages is a function of population vectors and exogenous parameters that determine the systematic and idiosyncratic payoffs. Thus, this yields the equilibrium distribution of marriages as a Marriage Matching Function (MMF) in Equation (9) under different living arrangements.

$$\pi_{IJ}^l = \ln \frac{\mu_{IJ}^{l,2}}{\mu_{I0} \cdot \mu_{0J}} \quad \text{for } l \in \{T, NT\} \quad (9)$$

Equation (9) has an intuitive interpretation where the left hand side reflects the expected total gain per spouse from an  $(I, J)$  match in living arrangement  $l$  relative to the common gains from remaining unmatched is a function of the number of such matches  $\mu_{IJ}^l$  scaled by the geometric average of the number of unmarried men and women. The marriage matching function under each living arrangement  $\mu(M, F; l,)$  returns a  $\mathcal{I} \times \mathcal{J}$  matrix, where the element  $\mu_{IJ}^l$  denotes the measure of marriages between type  $I$  men and type  $J$  women in living arrangement  $l$ . The marriage matching function must satisfy the following accounting conditions:

$$\mu_{0J} + \sum_l \sum_{I=1}^K \mu_{IJ}^l = f_J \quad \forall J \quad l = (T, NT) \quad (10)$$

$$\mu_{I0} + \sum_l \sum_{J=1}^K \mu_{IJ}^l = m_I \quad \forall I \quad l = (T, NT) \quad (11)$$

$$\mu_{0J}, \mu_{I0}, \mu_{IJ}^l \geq 0 \quad \forall l, I, J \quad (12)$$

where the first equation denotes that the total number of men who marry type  $J$  women and the number of unmarried type  $J$  women must be equal to the number of available type  $J$  women, for all  $J$ . While the second equation provides a set of accounting conditions that must hold for all male types. The last accounting constraint holds because the number of unmarried persons of any type and gender and the number of marriages between type  $I$  men and type  $J$  women must be non-negative. These accounting constraints are crucial towards ensuring that the predicted marriage rates are not above 1.

## 4.5 Empirical specification

To bring the model to data we implement the Type (1) extreme value assumption for the idiosyncratic marital tastes.<sup>21</sup> The crucial feature of our identification approach, like others in the literature, is that we observe and thereby are able to analyse multiple markets. While several papers have made use of

---

<sup>21</sup>This is a standard assumption and widely employed within the marriage market literature for parametrizing the distribution of the shocks see for e.g. Choo and Siow (2006), Chiappori, Salanie, Weiss (2017)

time variation in male and female characteristics in the population, this requires assumptions on the independence across different age cohorts. In this paper we combine quasi-random variation from the TSC policy with non-overlapping marriage markets along caste and region (geography) to generate over-identifying restrictions from cross-sectional variation from the IHDS sample.

The model is estimated in two parts using a two-step conditional choice probability (CCP) estimator. First stage estimates the sub-marriage market level demand and supply using male and female quasi-demand equations for each spouse type conditional on their own type with additional identifying restrictions using our caste and region demographic groups. In the second stage we use first stage estimates for marital preference and the distribution of wealth type for both men and women, to estimate the total gains from marriage for each type of match relative to single-hood. Specifically, to identify the surplus matrix  $\Pi$  and generate testable restrictions we include gender specific drifts  $\nu^I$  and  $\zeta^J$  that vary by caste and region categories denoted by  $g$ , such that the  $\Pi^{IJ}$  terms vary according to:

$$\Pi_g^{IJ} = \Pi^{IJ} + \nu_g^I + \zeta_g^J \quad (13)$$

The drifts absorb possible changes across castes and regions in the surplus generated by marriage capturing taste heterogeneity across different cultural and socio-economic groups. Marital preference estimates are used to solve a system of equalities derived from the key identified moment of the matching model under the assumption of transferable utility.<sup>22</sup>

$$\log(\hat{P}(J|I, g) \cdot \hat{P}(I|J, g)) - G_{IJ} - E_g^I - F_g^J = 0 \quad (14)$$

Using our sample of multiple non-overlapping groups and within group exposure to the TSC adds to the total number of equations and thereby generates additional identifying variation. The parameters of interest here are  $(G_{IJ}, E_g^I, F_g^J)$ , that together capture and provide an estimate of the marital gains. The parameters are estimated using a Minimum Distance Estimator (MDE) as outlined in Chiappori, Salanie and Weiss (2017). For this model, Choo and Siow (2006) show that the solution is locally unique.

## 4.6 Model Estimates

### Systematic/Total Gains

To show the difference in systematic gains between marriages with and without sanitation - we show in Figure (6) the systematic total gains in marriages for different living arrangements i.e., with  $\pi_{IJ}^T$  and without  $\pi_{IJ}^{NT}$  sanitation for couples of the same wealth type in 2005 prior to TSC exposure. Recall that the systematic (or total) gain measures the expected marital surplus to a random  $(I, J)$  pair marrying relative to them not marrying.

---

<sup>22</sup>In practice we exploit variation across a total of 20 groups ( $g = 1, \dots, 20$ ) using caste affiliation and region.

Figure 6: Systematic Gains under different living arrangements in 2005

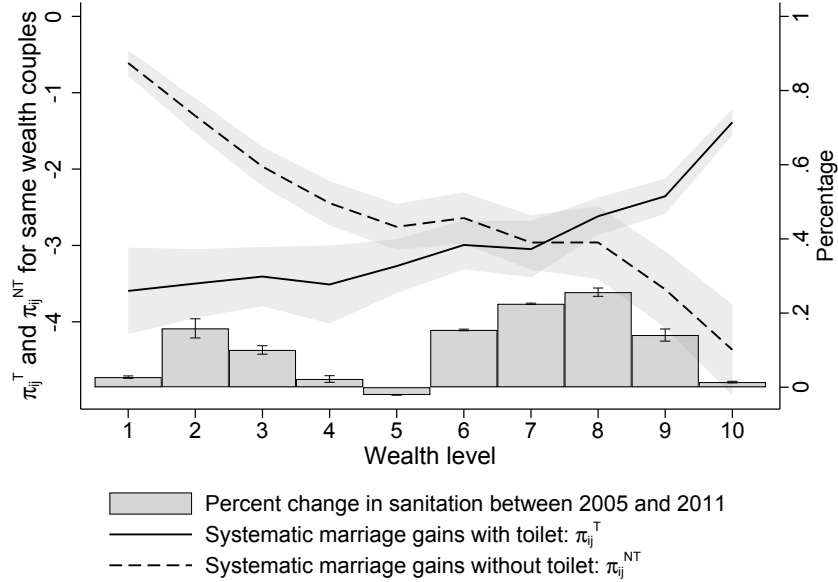


Figure (6) highlights two important points. First, the systematic gains under both living arrangements lie below zero indicating the large fraction of younger women and men in our sample. An average annual population growth rate of 1.44% between 2005 and 2011 (World Bank) would necessarily imply a further decrease in the systematic gain from marriage. Moreover, these composition changes due to population shifts are distinct from any changes in marital preferences but are observationally equivalent.

Second, the systematic gains to marriage with sanitation are not unilaterally larger but instead vary with the wealth distribution across the sub-marriage markets. Highlighting the scope for large heterogeneous policy effects. Specifically the systematic gains to marriage with sanitation are higher than that of marriages without only among the wealthiest of the same wealth matches. In our framework the equilibrium effects operate on both the type of living arrangement and type of spouse. In the first case the increase in living arrangements which come with sanitation, with TSC exposure, will imply a necessary decrease in the systematic gains  $\pi_{IJ}^T$  vs  $\pi_{IJ}^{NT}$ . In the second case, heterogeneous prevalence of living arrangements across the wealth distribution may generate composition changes in the choice of spouses. Therefore to quantify the total effect from TSC exposure would necessitate fixing the level of systematic gains.

Table 4: Probability of Marriage

	Men				Women			
	High TSC	Low TSC	Change	p-val	High TSC	Low TSC	Change	p-val
Wealth Type L	0.473	0.434	0.038	<b>0.042</b>	0.680	0.615	0.065	<b>0.006</b>
Wealth Type M	0.383	0.374	0.009	0.503	0.490	0.503	-0.012	0.478
Wealth Type H	0.361	0.360	0.001	0.635	0.401	0.398	0.003	0.838

Notes: Probability of marriage with caste, region and time group effects. Wealth Types L,M,H refer to low, medium and high wealth respectively Household level controls include age, assets, Education of Household Head. Standard errors clustered.

Table 5: Partners of Men

<b>Wife →</b>	Wealth Type L	Wealth Type M	Wealth Type H
<b>Husband ↓</b>	High TSC		
Wealth Type L	0.703	0.256	0.041
Wealth Type M	0.324	0.369	0.307
Wealth Type H	0.073	0.247	0.679
<b>Husband ↓</b>	Low TSC		
Wealth Type L	0.640	0.299	0.061
Wealth Type M	0.245	0.453	0.301
Wealth Type H	0.046	0.254	0.700
<b>Husband ↓</b>	Change		
Wealth Type L	0.063**	-0.043*	-0.020
Wealth Type M	0.078***	-0.084***	0.006
Wealth Type H	0.027	-0.006	-0.021

Notes: Conditional Probability of matching with Woman by wealth type with caste, region and time group effects. Wealth Types L,M,H refer to low, medium and high wealth respectively. Household level controls include age, assets, Education of Household Head. Standard errors clustered.

### Sorting Outcomes

In this section we use the estimated equilibrium model to analyse how the total sanitation campaign affected outcomes in the marriage markets. Table 4 presents how the probability of marriage between types is affected by the program intensity. The marriage probabilities increase among the poorest men

and women with TSC exposure.<sup>23</sup>

Table 6: Partners of Women

<b>Husband</b> →	Wealth Type L	Wealth Type M	Wealth Type H
<b>Wife</b> ↓	High TSC		
Wealth Type L	0.647	0.273	0.079
Wealth Type M	0.196	0.391	0.413
Wealth Type H	0.023	0.208	0.769
<b>Wife</b> ↓	Low TSC		
Wealth Type L	0.625	0.303	0.072
Wealth Type M	0.252	0.422	0.326
Wealth Type H	0.043	0.228	0.729
<b>Wife</b> ↓	Change		
Wealth Type L	0.023	-0.030	0.007
Wealth Type M	-0.056***	-0.031**	0.087
Wealth Type H	-0.020*	-0.020**	0.040

*Notes:* Conditional Probability of matching with Woman by wealth type with caste, region and time group effects. Wealth Types L,M,H refer to low, medium and high wealth respectively. Household level controls include age, assets, Education of Household Head. Standard errors clustered.

The partner choice for men and women in Table (5) and Table (6) rationalize patterns that we observe previously. By definition, the sum of the shares across different types of spouses for a given type of male or female is one. Table (5) shows a clear pattern for men where with TSC exposure men were more likely to marry "down" and less likely to marry "up".<sup>24</sup> However, the composition changes induced by TSC for women are less clear. Table (6) shows that while non-poor women were less-likely to marry "down" this change was not accompanied by a corresponding shift to marry "up" among poor and non-poor women alike. As such the fall in the marriage shares for women where the husband was as wealthy or less would be the equilibrium response to the increase in men who are now marrying "down". It is important to note that while Table (5) and Table (6) indicate changes in spousal choice i.e., whom to marry they do not incorporate the decision of whether or not to marry. In contrast the marriage matching function in equation (9) encapsulates both these effects into a single summary statistic. In the subsequent section, we

<sup>23</sup>Appendix tables 14 and 13 presents the same results for high and low sex ratio districts separately. Results are similar to those presented here.

<sup>24</sup>Sorting at the top of the wealth distribution showed no discernible changes.

use this summary statistic to analyse the total effect of the TSC looking at changes to both the marital surplus and the expected surplus share. Overall, the changes in marital choices in Table (5) and Table (6) induced by the Total Sanitation Campaign illustrate the importance of heterogeneous changes in marital surplus.

Table 7: Match Surplus

<b>Wife →</b>	Wealth Type L	Wealth Type M	Wealth Type H
<b>Husband ↓</b>	High TSC		
Wealth Type L	-0.788	-2.991	-6.969
Wealth Type M	-2.425	-1.936	-2.751
Wealth Type H	-5.146	-2.281	-0.649
<b>Husband ↓</b>	Low TSC		
Wealth Type L	-0.918	-2.586	-5.931
Wealth Type M	-2.598	-1.653	-2.678
Wealth Type H	-5.712	-2.493	-0.673
<b>Husband ↓</b>	Change		
Wealth Type L	<b>0.130</b>	-0.405	-1.038
Wealth Type M	<b>0.173</b>	-0.282	-0.074
Wealth Type H	<b>0.565</b>	<b>0.211</b>	<b>0.024</b>

*Notes:* Men in rows and women in columns. Wealth Types L,M,H refer to low, medium and high wealth respectively. Bold indicates increase in total match surplus.

## 5 How does the TSC affect Marriage Markets

The tables below are specific to the structural approach and illustrate the impact of the policy on marital surplus and expected female surplus share using model simulations.

### 5.1 Effect of TSC on Marital Surplus

In order to empirically quantify the impact of the TSC policy on the gains to marriage, we consider variation to a type of match characterised by wealth and whether the matched couple belongs to a high or low TSC district in each year. The change in the total match surplus for an (I,J) match type in the High

TSC exposed districts can be estimated by a difference in difference type estimator <sup>25</sup>

Table (7) shows the change in the gains to marriage across district exposure for spouses of different wealth types. We find that households with marriageable men and women exposed to the TSC policy experience a heterogeneous impact in match surplus for a given match. The total surplus from a match of a given type for each sub marriage market is given in Table (7) where a small negative number denotes a larger match specific surplus. Overall the match surplus increases in 5 out of total of 9 match types concern couples formed where the man is at least as wealthy as their spouse if not more i.e., lower triangle of the match surplus matrix. Table (7) highlights the importance of the general equilibrium effects generated by the Total Sanitation Campaign on the redistribution of gains across different sub-markets or matched types.

Table 8: Wife’s Expected Surplus Share (Female Shortage)

<b>Wife →</b>	Wealth Type L	Wealth Type M	Wealth Type H
<b>Husband ↓</b>	High TSC		
Wealth Type L	0.623	0.537	0.455
Wealth Type M	0.668	0.586	0.504
Wealth Type H	0.684	0.603	0.523
<b>Husband ↓</b>	Low TSC		
Wealth Type L	0.635	0.563	0.480
Wealth Type M	0.673	0.603	0.521
Wealth Type H	0.681	0.612	0.530
<b>Husband ↓</b>	Change		
Wealth Type L	-0.012	-0.026	-0.025
Wealth Type M	-0.004	-0.018	-0.017
Wealth Type H	0.004	-0.009	-0.007

*Notes:* Men in rows and women in columns. Wealth Types L,M,H refer to low, medium and high wealth respectively.

## 5.2 Effect of TSC on Surplus Share

While our context does not allow for any exogenous variation in the population supplies we are able to use model simulations to analyse the impact of the TSC program on different underlying marriage market

<sup>25</sup>This is similar to the approach implemented by Choo and Siow (2006) to analyse the impact of the partial legalization of abortion in the US on the total gains to marriage.

conditions.<sup>26</sup> Using the estimated parameters for the systematic gains from marriage under different living arrangements Table (8) illustrates the overall impact of the TSC on the wife's expected surplus share. Even though the relative scarcity of women is thought to imply a larger sharing weight the net equilibrium impact of the TSC seems to have eroded away these female gains. In other words, while for some type of couples there are benefits for getting married as seen in Table (7), wives will have less control of the resources of the household. The counterfactual estimates from the model in Table (9) show us that some women, in particular the poorest, would have experienced an increase in their expected surplus share if the gender differences were removed. This result is particularly striking in light of previous literature but nevertheless straightforward once equilibrium effects are taken into account.

Table 9: Wife's Expected Surplus Share (No Female Shortage)

<b>Wife</b> →	Wealth Type L	Wealth Type M	Wealth Type H
<b>Husband</b> ↓	High TSC		
Wealth Type L	0.720	0.434	0.388
Wealth Type M	0.797	0.540	0.492
Wealth Type H	0.826	0.586	0.539
<b>Husband</b> ↓	Low TSC		
Wealth Type L	0.597	0.490	0.435
Wealth Type M	0.675	0.573	0.519
Wealth Type H	0.698	0.600	0.546
<b>Husband</b> ↓	Change		
Wealth Type L	0.122	-0.055	-0.048
Wealth Type M	0.122	-0.033	-0.027
Wealth Type H	0.127	-0.013	-0.007

*Notes:* Men in rows and women in columns. Wealth Types L,M,H refer to low, medium and high wealth respectively.

Table (8) and (9) also illustrate the importance of equilibrium effects where the underlying sharing weight<sup>27</sup> is not a fixed parameter solely dependent on the scarcity of women relative to men. Both the weight and the surplus share of husbands and wives are also determined by their willingness to enter the marriage market. In fact from the table we see that the average surplus share for women in markets

<sup>26</sup>Note that there is very little regional variation in sex ratio with over 80% of high sex ratio districts located in the South and East of the country, according to 2011 census data. As such implied "shortage" of women and unobserved differences in marital preferences between the South and the rest of the country are observationally equivalent.

<sup>27</sup>sometimes referred to as the 'bargaining power of missing women'.

with female shortage is 0.582 while the average surplus share in markets with an excess of women is 0.581 illustrating that in fact that there is no difference in the surplus share associated with the supposed scarcity of women.

## 6 Conclusion

While the case for investment in sanitation has generally been convincingly made, there remains an incomplete understanding of impacts of sanitation interventions (Augsburg and Sainati (2020); Gautam (2017)). This maybe due to the complexity and heterogeneity of sanitation, which implies direct monetary and non-monetary costs and benefits, many of which are hard to measure, as well as positive and negative externalities. Gautam (2020) highlights the importance of inter-household interactions and quantifies the welfare effects in the presence of externalities from sanitation through an ex-ante approach. Nevertheless, there may be indirect impacts on individuals via alternative markets, such as marriage markets, which to date, have been largely ignored in the literature. In this paper we show that sanitation matters in marriage markets: the Indian Total Sanitation Campaign changed marriage market outcomes. We exploit in this paper quasi-random variation from the TSC, that shifted the distribution of households with sanitation, and the incentives of men and women to sort into marriage, or not. We show that exposure to a high-performing TSC district had a significant impact on both the composition of marriage types, and the sorting into marriages by men and women. The analysis relies on reduced-form techniques and a structural approach. Both approaches exploit information on marriages - including the chosen living arrangement of couples - across geographical areas and time. Identification is achieved through a difference-in-differences approach in a multi-market framework. For the structural analysis we develop and estimate a simple friction-less matching model à la Choo and Siow (2006), allowing for different types of living arrangements through a couple-specific random component affecting marriage surplus and other socio-economic spousal characteristics, such as wealth.

To analyse the importance of marriage market conditions and to illustrate the overall sorting effect we use the estimated equilibrium model to conduct two model simulations. First, we show that cohorts within TSC exposed markets experienced a shift in marital gains both across matches but also within a given match. Specifically, the resultant sorting patterns display a marked gender asymmetry with an increase in marital surplus among matches, where men are wealthier than their spouse, and a decrease in surplus where the wife is wealthier. Second, we consider the role of relative scarcity of women within the marriage market and find that the increase in female gains were negligible. Instead, the TSC had an unintended countervailing impact on female empowerment as measured by the effective control over household resources and reflected in the division of surplus. Our results indicate that this increased sorting of men and women into households with sanitation on the marriage market came at the expense of women losing control of household resources and a decline in female empowerment within marriage.

While evaluations of the TSC generally highlighted the policy's positive impacts on sanitation uptake, our paper draws attention to an important indirect effect, that might well be present in other such policies, including the 'No toilet, no bride' campaign. We argue that – to obtain a comprehensive picture of such female-focused policies including costs and benefits – their interaction with marriage markets should not be ignored.

The insights derived from our empirical analysis further highlight interesting avenues for future research that are beyond the scope of this paper. In particular, the importance of sanitation in the marriage market determined by the magnitude of unobserved heterogeneity raises interesting questions on the presence and nature of frictions specifically on the marriage market. Even though the transferable utility assumption limits our ability to explore such questions, our results emphasize a promising avenue to be explored in future research. In addition, while the marriage matching function in our analysis is static, it may be reasonable to expect additional aspects of marital behaviour, in response to policy exposure, to be dynamic e.g. marriage delay. Moreover, this response may differ across men and women. In such a case, we would need a dynamic marriage matching function. While a dynamic matching model lies beyond the primary focus of this paper, an important contribution of this paper is to provide a rigorous structural foundation that can be extended to explore resulting dynamics from the TSC. In summary, our analysis emphasises the importance of accounting for general equilibrium effects which necessitates going beyond reduced form methods and yet has been largely missing from the policy discourse within the literature at the intersection of sanitation and marriage markets.

## References

- Abramitzky, R., Delavande, A., and Vasconcelos, L. (2011). Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, 3(3):124–57.
- Abramovsky, L., Augsburg, B., Luuhrmann, M., Oteiza, F., and Rud, J. P. (2019). Community matters: heterogeneous impacts of a sanitation intervention.
- Adams-Prassl, A. and Andrew, A. (2019). Preferences and beliefs in the marriage market for young brides.
- André, P. and Dupraz, Y. (2019). Education and polygamy: Evidence from Cameroon. Technical report, Warwick Economics Research Papers No 1219.
- Andres, L. A., Deb, S., Joseph, G., Larenas, M. I., and Grabinsky Zabludovsky, J. (2020). A multiple-arm, cluster-randomized impact evaluation of the Clean India (Swachh Bharat) Mission program in rural Punjab, India. Working Paper 9249, World Bank".
- Angrist, J. (2002). How do sex ratios affect marriage and labor markets? Evidence from America's second generation. *The Quarterly Journal of Economics*, 117(3):997–1038.
- Arnold, B. F., Khush, R. S., Ramaswamy, P., London, A. G., Rajkumar, P., Ramaprabha, P., Durairaj, N., Hubbard, A. E., Balakrishnan, K., and Colford, J. M. (2010). Causal inference methods to study non-randomized, preexisting development interventions. *Proceedings of the National Academy of Sciences*, 107(52):22605–22610.
- Augsburg, B. and Sainati, T. (2020). Editorial: WASH economics and financing: towards a better understanding of costs and benefits. *Journal of WASH for Development*, 10:615—617.
- Banerjee, A., Duflo, E., Ghatak, M., and Lafortune, J. (2013). Marry for what? Caste and mate selection in modern India. *American Economic Journal: Microeconomics*, 5(2):33–72.
- Barnard, S., Routray, P., Majorin, F., Peletz, R., Boisson, S., Sinha, A., and Clasen, T. (2013). Impact of Indian Total Sanitation Campaign on latrine coverage and use: A cross-sectional study in Orissa three years following programme implementation. *PLOS One*, 8.
- Basu, A. (1999). Fertility decline and increasing gender imbalance in India, including a possible South Indian turnaround. *Development and Change*, 30:237–263.
- Beauchamp, A., Calvi, R., and Fulford, S. (2017). Terms of engagement: Marriage and migration in India. In *Structural Models in Development: Migration, Marriage and the Family*. Econometric Society.
- Becker, G. S. and Becker, G. S. (2009). *A Treatise on the Family*. Harvard university press.
- Borker, G., Eeckout, J., Luke, N., Minz, S., Munshi, K., and Swaminathan, S. (2018). Wealth, marriage, and sex selection. PAA 2018 Annual Meeting.
- Borker, G., Eeckout, J., Luke, N., Minz, S., Munshi, K., and Swaminathan, S. (2019). Wealth, marriage, and sex selection. Working paper.
- Botticini, M. and Siow, A. (2003). Why dowries? *American Economic Review*, 93(4):1385–1398.
- Briceno, B., Coville, A., Gertler, P., and Martinez, S. (2017). Are there synergies from combining hygiene

- and sanitation promotion campaigns: Evidence from a large-scale cluster-randomized trial in rural Tanzania. *PLoS One*, 12(11):e0186228.
- Cavill, S., Mott, J., Tyndale-Biscoe, P., Bond, M., Edström, J., Huggett, C., and Wamera, E. (2018). Men and boys in sanitation and hygiene: A desk-based review. CLTS Knowledge Hub and Learning Paper, Institute of Development Studies.
- CBGA, U. (2011). Total sanitation campaign (TSC). budgeting for change series, 2011. Centre for Budget and Governance Accountability; Social Policy, Planning, Monitoring and evaluation (SPPME), UNICEF.
- Charles, K. K. and Luoh, M. C. (2010). Male incarceration, the marriage market, and female outcomes. *The Review of Economics and Statistics*, 92(3):614–627.
- Chiappori, P.-A., Dias, M. C., and Meghir, C. (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy*, 126(S1):S26–S72.
- Chiappori, P.-A., Iyigun, M., and Weiss, Y. (2009). Investment in schooling and the marriage market. *American Economic Review*, 99(5):1689–1713.
- 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.
- Chiappori, P.-A., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–67.
- Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of political Economy*, 114(1):175–201.
- Clasen, T., Boisson, S., Routray, P., Torondel, B., Bell, M., Cumming, O., Ensink, J., Freeman, M., Jenkins, M., Odagiri, M., et al. (2014). Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: A cluster-randomised trial. *The Lancet Global Health*, 2(11):e645–e653.
- Cox, O. C. (1940). Sex ratio and marital status among negroes. *American Sociological Review*, 6(5):937–947.
- Crocker, J. (2016). Teachers and sanitation promotion: An assessment of community-led total sanitation in Ethiopia. *Environmental Science and Technology*, 50(12).
- Crocker, J., Abodoo, E., Asamani, D., Domapielle, W., Gyapong, B., and Bartram, J. (2016). Impact evaluation of training Natural Leaders during a Community-Led Total sanitation intervention: A cluster-randomized field trial in Ghana. *Environ. Sci. Technol.*, 50(16):8867–8875.
- Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., and Hwa-Ok, B. (2003). Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *The Journal of Development Studies*, 40:153–187.
- Desai, S. and Andrist, L. (2010). Erratum: Fatter attraction: Anthropometric and socioeconomic matching on the marriage market. *Demography*, 47:667–687.
- Easterlin, R. A. (1961). The American baby boom in historical perspective. *American Economic Review*, 51(5):869–911.

- Fernández, R. and Wong, J. C. (2017). Free to leave? A welfare analysis of divorce regimes. *American Economic Journal: Macroeconomics*, 9(3):72–115.
- Garn, J. V., Sclar, G. D., Freeman, M. C., Penakalapati, G., Alexander, K. T., Brooks, P., Rehfuess, E. A., Boisson, S., Medlicott, K. O., and Clasen, T. F. (2017). The impact of sanitation interventions on latrine coverage and latrine use: A systematic review and meta-analysis. *International journal of hygiene and environmental health*, 220(2):329–340.
- Gautam, S. (2017). *Household Demand in the Presence of Externalities: Model and Applications*. PhD thesis, University College London, United Kingdom.
- Gautam, S. (2020). Quantifying welfare effects in the presence of externalities: An ex-ante evaluation of sanitation interventions. Working Paper, Washington University in St. Louis.
- Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2016). Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation. *American Economic Journal: Macroeconomics*, 8(1):1–41.
- Grossbard, S. (1993). On the economics of marriage—a theory of marriage, labor and divorce. *MPRA Paper*, 1832.
- Guiteras, R., Levinsohn, J., and Mobarak, A. M. (2015). Sanitation subsidies. encouraging sanitation investment in the developing world: A cluster-randomized trial. *Science*, 348(6237):903–906.
- Gupta, B. (2014). Where have all the brides gone? Son preference and marriage in India over the twentieth century. *The Economic History Review*, 67(1):1–24.
- Hammer, J. S. and Spears, D. (2016). Village sanitation and children’s human capital: evidence from a randomized experiment by the Maharashtra government. World Bank Policy Research Working Paper No. 6580.
- Hener, T. and Wilson, T. (2018). Marital age gaps and educational homogamy—evidence from a compulsory schooling reform in the UK. Technical report, ifo Working Paper.
- Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010). Matching and sorting in online dating. *American Economic Review*, 100(1):130–63.
- Iyigun, M. and Walsh, R. P. (2007). Endogenous gender power, household labor supply and the demographic transition. *Journal of Development Economics*, 82(1):138–155.
- Jaggi, T. (2001). The economics of dowry: Causes and effects of an Indian tradition. *University Avenue Undergraduate Journal of Economics*, 5(1):2.
- Kone, Z. L., Liu, M. Y., Mattoo, A., Ozden, C., and Sharma, S. (2018). Internal borders and migration in India. *Journal of Economic Geography*, 18(4):729–759.
- Lundberg, S. and Pollak, R. A. (1996). Bargaining and distribution in marriage. *Journal of economic perspectives*, 10(4):139–158.
- Moorjani, P., Thangaraj, K., Patterson, N., Lipson, M., Loh, P.-R., Govindaraj, P., Berger, B., Reich, D., and Singh, L. (2013). Genetic evidence for recent population mixture in India. *The American Journal of*

- Human Genetics*, 93(3):422–438.
- Murthi, M., Guio, A.-C., and Dréze, J. (1995). Mortality, fertility, and gender bias in India: A district level analysis. *Population and Development Review*, pages 745–782.
- Orgill-Meyer, J. and Pattanayak, S. K. (2020). Improved sanitation increases long-term cognitive test scores. *World Development*, 132(104975).
- Patil, S. R., Arnold, B. F., Salvatore, A. L., Briceno, B., Ganguly, S., Colford Jr, J. M., and Gertler, P. J. (2014). The effect of India’s Total Sanitation Campaign on defecation behaviors and child health in rural Madhya Pradesh: A cluster randomized controlled trial. *PLoS Med*, 11(8):e1001709.
- Pattanayak, S. K., Yang, J.-C., Dickinson, K. L., Poulos, C., Patil, S. R., Mallick, R. K., Blitstein, J. L., and Praharaj, P. (2009). Shame or subsidy revisited: social mobilization for sanitation in orissa, india. *Bulletin of the World Health Organization*, 87:580–587.
- Pencavel, J. (1998). Assortative mating by schooling and the work behavior of wives and husbands. *The American Economic Review*, 88(2):326–329.
- Pickering, A. J., Djebbari, H., Lopez, C., Coulibaly, M., and Alzua, M. L. (2015). Effect of a Community-Led Sanitation intervention on child diarrhoea and child growth in rural Mali: A cluster-randomised controlled trial. *The Lancet Global Health*, 3(11):e701–e711.
- Radtke, I. (2018). WASH guidelines for field practitioners. Part 2: Sanitation. Malteser International.
- Rao, V. (1993). The rising price of husbands: A hedonic analysis of dowry increases in rural India. *Journal of political Economy*, 101(4):666–677.
- Rasul, I. (2006). Marriage markets and divorce laws. *Journal of Law, Economics, and organization*, 22(1):30–69.
- Reynoso, A. (2018). The impact of divorce laws on the equilibrium in the marriage market. *University of Michigan, Job Market Paper*.
- Spears, D. and Lamba, S. (2016). Effects of early-life exposure to sanitation on childhood cognitive skills: Evidence from India’s Total Sanitation Campaign. *Journal of Human Resources*, 51(2):298–327.
- Stopnitzky, Y. (2017). No toilet no bride? Intrahousehold bargaining in male-skewed marriage markets in India. *Journal of Development Economics*, 127:269–282.
- Water and Sanitation Program (WSP) (2011). A decade of the Total Sanitation Campaign: Rapid assessment of processes and outcomes. Volume 1: Main Report.
- Zha, D. (2019). Schooling expansion and the female marriage age: Evidence from Indonesia. Columbia University, Job Market Paper.

## A Descriptives and Additional Take-up exercises

Table 10: Descriptive statistics in 2004/05 by TSC groups

	All	Low TSC	High TSC	p-val Diff
<b>Age gap:</b>				
p25	0.0	0.0	0.0	
p50	1.0	1.0	0.0	
p75	4.0	3.0	4.0	
mean	2.1	2.0	2.3	0.0
<b>Age at marriage (women):</b>				
p25	15.0	15.0	16.0	
p50	17.0	17.0	18.0	
p75	19.0	19.0	20.0	
mean	17.1	17.0	18.0	0.000
<b>Education level (men):</b>				
None	0.148	0.163	0.086	0.000
Primary	0.149	0.154	0.127	0.000
Incomplete secondary	0.197	0.202	0.172	0.000
Secondary	0.408	0.391	0.478	0.000
More than secondary	0.098	0.090	0.137	0.000
<b>Education level (Women):</b>				
None	0.299	0.333	0.147	0.000
Primary	0.147	0.154	0.114	0.000
Incomplete secondary	0.178	0.177	0.180	0.191
Secondary	0.308	0.280	0.435	0.000
More than secondary	0.068	0.056	0.124	0.000
<b>Marriage market definition</b>				
Percentage inter-caste marriage	0.049	0.047	0.056	0.296
Grow up in husband's village	0.119	0.105	0.183	0.000
Hours to natal home (wife):	3.287	3.282	3.306	0.652

*Notes:* Own calculations using data from the IHDS waves 20041. Low TSC districts corresponds to a grand score for implementation below 61, and high TSC to those with a number of 61 or above. The grand score per district is taken from WSP (2011).

Table 11: TSC impact on sanitation for households with marriageable male and females by household type

Variables	Only son marriageable households			Only daughter marriageable households		
	All	High sex ratio	Low sex ratio	All	High sex ratio	Low sex ratio
$dT$	-0.0450*** (0.0124)	-0.0378 (0.0340)	-0.0483*** (0.0134)	-0.0192 (0.0172)	-0.00332 (0.0450)	-0.0225 (0.0179)
$TSC$	0.0435 (0.0414)	0.132* (0.0675)	0.00612 (0.0478)	0.110*** (0.0386)	0.145** (0.0695)	0.0847* (0.0460)
$dT \times TSC$	0.0779** (0.0333)	-0.0653 (0.0492)	0.148*** (0.0374)	0.0364 (0.0352)	-0.0359 (0.0605)	0.0738* (0.0446)
Age oldest marriageable individual	0.00654*** (0.00109)	0.00358 (0.00251)	0.00678*** (0.00121)	0.00846*** (0.00228)	0.00943*** (0.00295)	0.00769*** (0.00288)
Marital status oldest marriageable individual	-0.0439*** (0.0132)	-0.00254 (0.0227)	-0.0472*** (0.0150)	-0.0825*** (0.0295)	-0.0740** (0.0336)	-0.0849** (0.0389)
Constant	0.267 (0.179)	0.339*** (0.0865)	0.227 (0.178)	0.132 (0.163)	0.176 (0.166)	0.133 (0.169)
Observations	21,245	4,437	16,808	6,594	1,747	4,847
$R^2$	0.424	0.491	0.414	0.401	0.462	0.393

Notes: Own calculations using data from the IHDS waves 2004 and 2011. The sample consists of all single and married females (males) aged 15 to 34 at the survey time. The high sex ratio corresponds to districts with at least 999 women per 1000 men in the age range 15-34, while the low sex ratio is districts with less than 999 sex-ratio. District level sex ratio information was computed using information from the population census 2001 and 2011. High TSC districts correspond to a grand score for implementation of 61 or above, and high TSC to those with a score of 61 or above. The grand score per district is taken from WSP (2011). Oldest marriageable female (male) is an indicator variable that takes the value of 1 if the respondent is the oldest marriageable (age 15-34) daughter (son) in the household. Apart from the coefficients presented in the table, as controls, we consider an index of the female's (male's) wealth; age of the female (male); household size; education level of the household head (no education [base], primary, incomplete secondary, secondary, and above secondary); caste (Brahmin [base], High caste, Other backward caste, Dalit, Adivasi, Muslim, Sikh - Jain); and indicator variable for the rural zone; and fixed effects at the state level. Clustered at district level standard errors in parentheses. Significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 12: Sensitivity to the cut-offs definitions

		High Sex Ratio cut-off *				
		990	995	999	1005	1015
Oldest marriageable son	High Sex Ratio	-0.065*	-0.093**	-0.083*	-0.096**	-0.106**
	Low Sex ratio	0.144***	0.144***	0.136***	0.138***	0.131***
Oldest marriageable daughter	High Sex Ratio	-0.007	-0.023	-0.038	-0.057	-0.063
	Low Sex ratio	0.074*	0.074*	0.077**	0.086**	0.085**
		TSC exposure cut-off**				
		50	60	61	65	67
Oldest marriageable son	High Sex Ratio	-0.045	-0.083*	-0.083*	-0.090*	-0.098**
	Low Sex ratio	0.088***	0.132***	0.136***	0.147***	0.151***
Oldest marriageable daughter	High Sex Ratio	-0.011	-0.038	-0.038	-0.054	-0.077
	Low Sex ratio	0.061*	0.077**	0.077**	0.082**	0.078*

\* TSC=61 \*\* Sex ratio = 999

Table 13: Probability of Marriage (LSR)

	Men				Women			
	High TSC	Low TSC	Change	p-val	High TSC	Low TSC	Change	p-val
Wealth Type L	0.451	0.438	0.014	0.504	0.629	0.633	-0.004	0.881
Wealth Type M	0.389	0.386	0.003	0.865	0.502	0.524	-0.023	0.256
Wealth Type H	0.367	0.376	-0.008	0.569	0.394	0.412	-0.018	0.302

Notes: Probability of marriage with caste, region and time group effects. Wealth Types L,M,H refer to low, medium and high wealth respectively. Household level controls age, assets, Education of Household Head. Standard errors clustered.

Table 14: Probability of Marriage (HSR)

	Men				Women			
	High TSC	Low TSC	Change	p-val	High TSC	Low TSC	Change	p-val
Wealth Type L	0.493	0.423	0.070	0.169	0.826	0.558	0.268	<b>0.000</b>
Wealth Type M	0.359	0.325	0.034	0.141	0.407	0.410	-0.003	0.924
Wealth Type H	0.308	0.297	0.011	0.600	0.350	0.345	0.004	0.881

Notes: Probability of marriage with caste, region and time group effects. Wealth Types L,M,H refer to low, medium and high wealth respectively. Household level controls include age, assets, Education of Household Head. Standard errors clustered.