Essays in Applied Microeconomics

Inauguraldissertation

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftlichen Fakultät

der Universität zu Köln

2024

vorgelegt von

Khalid Imran, M.Sc. M.S.S

aus

Dinajpur, Bangladesh

Referent: Prof. Dr. Daniel Wiesen

Korreferent: Prof. Dr. Lorenz Goette, Prof. Dr. Ludwig Kuntz

Tag der Promotion: October 11, 2024

To

My late father
and
beloved mother,

who sacrificed so much to give me a better life.

Acknowledgements

It was not easy to come all the long way from a small city in Bangladesh to where I stand now, about to finish my academic journey. This path could not have been undertaken without the support of many individuals to whom I owe my appreciation.

First and foremost, I extend my profound thanks to my primary supervisor, Daniel Wiesen. He has been an invaluable guide, providing exceptional insights that have significantly shaped my research methodology and analytical approach. My deep indebtedness to him extends far beyond excellent academic supervision, from helping with dormitory extension to writing letters for conference visa applications. In the future, when supervising students myself, I will strive to emulate his generosity.

Extreme gratitude goes to my second supervisor, Lorenz Goette, for his continuous encouragement and guidance from Singapore. The distance never felt significant due to his consistent availability, prompt responses, and insightful virtual discussions. Each zoom session left me feeling more motivated and confident than before.

Sincere appreciation is extended to Tom Zimmermann for his invaluable suggestions and feedback on my research, and particular gratitude for his generous help in resolving organizational issues on multiple occasions.

Very special gratitude is reserved for Ingo E. Isphording, my master's thesis supervisor and long-time boss at the Institute of Labor Economics (IZA). Words fail to express how grateful I am for all his favors, instructions, and guidance over the years. His support was crucial for my survival in Germany.

Thanks to all the professors who supported me, especially Rocco Macchiavello from the LSE, Pia Pinger, and Matthias Heinz from the University of Cologne. Heartfelt thanks to Katharina Laske for her unwavering support from the start of my Ph.D. I am deeply grateful to the Cologne Graduate School in Management, Economics, and Social Sciences for the scholarship throughout my doctoral studies. My sincere thanks are extended to Jennifer Mayer for her kindness and spontaneous offers of help. The Center for Social and Economic Behavior (C-SEB) deserves thanks for awarding me the junior start-up grant (Rd14-2022-JSUG-Imran). Many thanks to Nadine Klebe for all her timely support and help over the years.

I am really grateful to Arijit Ghosh for his continuous guidance and encouragement to finish the dissertation. Deep appreciation goes to Anna Stirner for all her kind words and support - her true goodwill was always felt. Gratitude is expressed to Sercan Demir for his encouraging words and sincere concern throughout this process. Thanks to Max R. P. Grossmann for his benevolence.

Sincere thanks to friends from the 2019 economics research masters cohort - Thanh, Marten, Melisa, Niklas, Nobel, and Robin - all of whom helped tremendously during the early years to cope with a new environment. My gratitude for the wonderful time spent with friends and co-workers from IZA, especially Pooja, Daniela, Jessica, and Johannes. I am thankful to Angelika, Cristina, Jingke, and Robin for their wonderful support during my thesis writing phase. I must thank Abu S. Shonchoy for his mentoring role, which has benefited not only me but also the entire Bangladeshi economics student community, both at home and abroad. I am thankful to Aziz bhai for being a true guardian in Germany.

Special thanks to Mohammad Jakaria, a friend and colleague from Hajee Mohammad Danesh Science and Technology University (HSTU). His advice, suggestions, and reviews are invaluable and eased my work in many ways. I would like to thank Prof. Ruhul Amin, former vice-chancellor of HSTU for his encouraging talks during short walks home after Friday prayers. I extend my heartfelt gratitude to all my students at HSTU, hoping that the knowledge I've gained at the University of Cologne will empower them to work towards the betterment of marginalized communities.

Words fall short in expressing my thanks to my friend Akib Khan. For decades, he has been a guiding light, showing me the path forward. His profound understanding of economics and sharp analytical skills have been invaluable to me. Alongside him, Laila apa's kind words have always been a vital source of motivation. Their unwavering support as a couple has been a cornerstone of my journey.

Finally, I would like to express my gratitude to my family members. My sincere thanks go to my uncle and late aunt for their love and support during my stay at their home while studying at the University of Dhaka. I am grateful to my only sister for her constant words of encouragement, and to my brother-in-law for stepping into the role of another son for my mother. Tafheem babu thanks for being a source of joy. My appreciation extends to Mostofa vai and Lucky vabi for their unwavering support. Jacki apa, Minara apa, and Sania apa, I am profoundly grateful for your regular care for my mother. Thanks to Putli for being a faithful companion to her.

Above all, I thank my mother for her remarkable strength, especially during my extended stays in Germany when she was alone at home. Her words and support have been the foundation of my achievements. Alhamdulillah for everything.

Lastly, Urmee, through ups and downs, near and far, she has consistently been my beacon of hope. Her unwavering support and wisdom have guided me through challenges and have profoundly shaped my path. I owe her everything.

Contents

List of Figures

List of Tables

Introduction

Health is a type of human capital, and investing in health has the potential to enhance one's longevity (Becker, 2007). An individual's level of health determines the amount of time available to generate income and produce commodities (Grossman, 1972). To achieve sustainable development and growth, it is essential for any nation to build health capital among its population. It is especially crucial for low and middle income countries (LMICs) where health capital accumulation depends on number of factors, including access to healthcare, behavioral aspects, resource constraints, education level, cultural norms, historical traditions, etc. Investing in the health of women and children in these regions can lead to intergenerational improvements in social welfare and economic well-being.

Given the widespread presence of gender inequalities in healthcare within LMICs, investing in women's health continues to be a matter of utmost importance. Unfortunately, the amount of scientific research unique to women's health in the developing world is inadequate. The studies carried out in this thesis mostly discuss the trajectory of women's health in developing countries using rigorous quasi-experimental methodologies such as difference in difference in differences (DDD), regression discontinuity design (RDD), and experimental methods such as randomized control trial. In general, it aims to explore how cultural norms related to gender, financial barriers, and policy interventions interplay to shape health-related behaviors and decisions among individuals, specifically for women. More specifically, it focuses on societal and paternal behavioral patterns related to the health of women at different stages of their life.

Out of three essays, the first one discusses how child health, household consumption, and the outcome of large intervention programs are impacted in the presence of a second-born female child in the family in Bangladesh. The second essay demonstrates the accumulative effect of nearing the legal age of marriage on womens' health based on an Indian sample. The final chapter examines how people make choices about healthcare providers and take decisions related to medication consumption in rural Bangladesh. All these aspects are underrepresented in the current discussion in the economics literature and yet to be examined from diverse socio-economic dimensions. By exploring these understudied areas, these studies collectively contribute to our understanding of health capital formation and its socio-economic determinants in developing countries, highlighting the complex interplay between gender, age, and healthcare decision-making in building human capital.

The first essay, titled "How Son Preference Affects Child Health and Food Consumption: Causal Evidence from Rural Bangladesh" examines the impact of a second-born female on health outcomes of first-born children, as well as household welfare in rural Bangladesh. With the predominant parental preference for a male child in society, it analyzes how the gender of the second-born child affects the nutritional outcomes of the first-born and the overall food consumption within households. The study also explores how child gender impacts the outcome of a large-scale intervention program called Feed the Future (FTF), designed to enhance household income and food security. To assess how son preference influences the health of the first-born child, the analysis employs a linear regression model under the assumption that the gender of the second-born child is random. To investigate the impact of a second-born female on FTF program outcome, it employs DDD regression framework. The study uses a sample of 1,747 first-born children aged 5-16 with at least one younger sibling within the same age range, representative of rural Bangladeshi households.

The findings reveal that first-born children with a female sibling have a higher height-for-age (HFA) z-score by an average of 0.10 standard deviations compared to those with younger brothers. It also find that when the second-born children is female, first child's calorie share in household increases by 0.05 percentage point, indicating shifted parental attention towards the first-born. The results further indicate a notable decrease in per capita calorie intake when the first-born and second-born children are both female, with a reduction of 121.62 kilocalories on average. This suggests that the presence of a second-born female leads to a redistribution of food resources within the family. The households that are part of the FTF program experience a significant reduction in the Household Dietary Diversity Score (HDDS) by 0.52 points if there is a second-born female in the family. This pattern reflects a broader gender-based discriminatory practice in food allocation and consumption within households. The results underscore the intricate relationship between parental preference for male children and health and consumption patterns at both individual and household levels in rural Bangladeshi communities. These findings reveal how deeply embedded cultural biases toward gender can significantly influence nutritional outcomes and resource distribution within families. Moreover, the study highlights the impact of these gender preferences on the accumulation of child health, suggesting that son preference directly affects the long-term health capital of children by skewing nutritional resources toward offspring of preferred gender.

The second essay is titled as "Marriage Market Induced Health Investments in

Women: Evidence from a Regression Discontinuity Design in India". This study investigates the impact of reaching the legal marriage age of 18 on the health indicators of women in India. More precisely, it investigates if there are any changes in women's health accumulation when they reach the age at which they are legally allowed to marry and enter the marriage market. This study employs a regression discontinuity design (RDD) using a nationally representative sample of 50,745 women aged 17 to 19 years of age from the Indian Demographic and Health Survey. The RDD utilizes the discontinuity in the legal approval status of marriage in India that occurs in the month a woman turns 18 years old. The method used includes local randomization and local polynomial-based RDD estimations to ensure robust findings. This study systematically examines women's health metrics just before and right after they turn 18, evaluating their BMI, BMI-for-age z-scores, and blood parameters such as glucose and hemoglobin levels. The study meticulously tests for covariate balance to ensure that observed effects are attributable to reaching the legal marriage age rather than any other pre-existing differences between the groups.

The result reveals significant improvements in health indicators, which indicates a higher investment in health at the point when women reach the legal marriage age of 18. Body mass index (BMI) showed a notable increase of 0.404 kg/m^2 immediately upon reaching 18, with a stronger effect observed in rural populations. The BMI-for-age z-score, an indicator of relative weight adjusted for age, also increased by 0.129 standard deviations, suggesting enhanced nutritional status. However, while blood glucose levels saw a slight rise, hemoglobin levels experienced a marginally significant decrease, indicating a possible shift in dietary patterns. Additionally, consumption patterns shifted with the increase in the likelihood of daily consumption of meat and fish consumption by 0.08 and 1.3 percentage points, respectively, suggesting that entry into the marriage market influences dietary habits and overall health investment in anticipation of the marriage of a woman. The findings provide evidence that sudden health accumulation occurs among indian women around the legal marriage age when she enters the marriage market. It provides a relevant contribution to the policy discussion regarding change in the legal age for marriage in many developing countries.

The third essay is titled "Addressing Antimicrobial Resistance and Self-Medication Practice in Developing Countries: A Field Experiment in Bangladesh". It is a randomized control trial (RCT) that addresses the improper overuse of antibiotic medicine, which causes antimicrobial resistance (AMR) in the human body. In

Bangladesh, unqualified healthcare providers and self-medication practices are prevalent, particularly in rural areas, exacerbating the AMR issue. Despite laws requiring prescriptions for purchasing antibiotics, poor enforcement allows for widespread non-prescription use. This study investigates whether information interventions can effectively reduce unnecessary antibiotic use, focusing on rural women in Bangladesh. The study aims to determine if providing information about AMR risks, combined with behavioral nudges or transport subsidies, can reduce visits to unqualified healthcare providers and increase visits to qualified doctors.

The research employs a clustered RCT involving 1,129 women from 62 villages in two sub-districts of Dinajpur district. Participants were divided into three groups: a control group receiving no intervention, a group receiving AMR risk information and color-coded medicine storage boxes (visual cue), and a group receiving AMR risk information and a monthly transport subsidy up to 100 bdt (approximately 80 cents) for visiting qualified doctors. Baseline data was collected in October 2023, covering household demographics, medical history, knowledge of medicines, and healthcare provider choices. The main experiment began in January 2024, with follow-up data on healthcare visits collected in March 2024.

The primary findings show a significant relationship between the distance from government hospitals and the frequency of visits to unqualified healthcare providers. Each additional mile to the nearest Upazila Health Complex or Medical College Hospital associates with increased visits to unqualified practitioners. The interventions had significant impact on healthcare seeking behavior of respondents. Participants receiving blue and red color coded boxes as visual cues visited qualified healthcare providers 0.39 times more than the control group, and those receiving transport subsidies visited 0.27 times more. However, neither intervention significantly reduced visits to unqualified providers within the two-month follow-up period. The study makes several contributions: it is the first RCT in a LMIC setting focusing on household women that aim to reduce unnecessary antibiotic consumption; it provides insights into the effects of risk information on health behavior in a developing country context; it explores the impact of behavioral interventions, specifically visual cues, on healthcare choices; and it offers micro-level causal evidence on the determinants of antibiotic consumption. The study demonstrates that information interventions, when combined with behavioral nudges or transport subsidies, can effectively increase visits to qualified healthcare providers, potentially reducing unnecessary antibiotic use. Overall, this study offers a scalable strategy for national public health improvement related to AMR risk.

In these essays, I causally explore the impact of different interventions and existing policies on health outcomes. I also propose innovative intervention strategies and measures that can improve women's health and consequently enhance their socioeconomic status. By offering detailed economic analysis and actionable policy recommendations, these research aims to provide comprehensive solutions to stakeholders and policymakers committed to closing the gender gap in the accumulation of health capital in developing economies.

Contribution

Chapter 1, "How Son Preference Affects Child Health and Food Consumption: Causal Evidence from Rural Bangladesh", is single authored based on my idea. This paper is an expanded version of my master's thesis.

Chapter 2, "Marriage Market Induced Health Investments in Women: Evidence from a Regression Discontinuity Design in India", is a collaborative effort with Daniel Wiesen and Johannes Schmieden. I conceived the idea, and then we worked collaboratively to develop and refine the paper's content. Daniel Wiesen and I wrote the initial draft.

Chapter 3, "Addressing Antimicrobial Resistance and Self-Medication Practice in Developing Countries: A Field Experiment in Bangladesh", is a joint work with Daniel Wiesen and Lorenz Goette. The idea emerged from our collective discussions. All of three contributed to the design of the experiment. I conducted fieldwork in Bangladesh. All authors contributed to developing the content of the paper.

Chapter 1

How Son Preference Affects Child Health and Food Consumption: Causal Evidence from Rural Bangladesh*

Abstract

We analyze the impact of son preferences on individual and household-level outcomes. First, we show how the presence of a second-born female in a household affects the health and nutritional outcomes of first-born children. Using nationally representative survey data from rural Bangladesh and utilizing the randomness of the second child's gender, we find that first-born children's height-for-age z-score increases significantly by 0.10 standard deviation if the second-born is a female. This is due to the higher proportion of household calories allocated to the eldest child in this sibling gender combination. Second, using two-period panel data on the same households and employing a triple difference regression framework, we show that removed financial constraints through the Feed the Future program, a US government's initiative addressing global hunger and food security, do not encourage families with second-born female children to diversify food consumption. Our findings thus imply that despite the efforts of the Feed the Future program being a daughter is still a distress in rural areas of Bangladesh.

*We are grateful for valuable comments and suggestions from Sibbir Ahmed, Jishnu Das, Lorenz Götte, Arijit Ghosh, Shahidul Islam, Mohammad Jakaria, Ingo Isphording, Akib Khan, Robin Möllerherm, Mohammad Mashiur Rahman, and seminar and conference participants at dggö annual conference Hannover, Essen Health Conference 2023, EuHEA PhD and supervisor conference Bologna, iHEA Cape Town, Global Health Economics Doctoral Student Mentorship Workshop. For insightful and valuable discussions of our paper, we thank Alessandro Toppeta, Sandy Tubeuf, and Zachary Wagner. We specially thank Wahid Quabili (International Food Policy Research Institute) for the support in providing the geospatial and additional datasets.

1.1 Introduction

Women worldwide face discrimination, which is often associated with individual malnutrition and mortality (Sen and Sengupta, 1983; Sethuraman and Duvvury, 2007; Marphatia et al., 2016). Malnutrition not only affects women's reproductive health and causes maternal mortality in the future, but it is also transmitted intergenerationally to the women's offspring (Emanuel, 1986; Rush, 2000; Sen, 2002; Delisle, 2008; Subramanian et al., 2009; Özaltin et al., 2010; Khatun et al., 2018; Wells et al., 2020).[1] Therefore, to have a healthy future generation, paying additional attention to the health and welfare of women in developing countries of key importance. Instead of extra care, however, it is well documented that women in the developing world face extreme inequalities in every sphere of life, from healthcare to education (Okojie, 1994; Jayachandran, 2015; Kennedy et al., 2020).

Many of the inequalities women are facing originate from the parental preference for a male child, the so-called son preference. In Bangladesh, as in many other developing countries, a son preferences is still persistent. Discrimination starts at the time of birth and continues over a lifetime, which is reflected through their vulnerable health status. Thirty-nine percent of Bangladeshi women in the age group of 10-18 years are undernourished according to the standard set by the World Health Organization (Helen Keller International, 2016). The way food and healthcare are distributed within households clearly favors males in Bangladesh (Chen et al., 1981; Abdullah and Wheeler, 1985; Koenig and D'Souza, 1986; Mitra et al., 2000; D'Souza and Tandon, 2019; Hossain et al., 2021).

In developing countries, poverty, food insecurity, malnutrition, and gender inequality are widely prevalent. Governments and international organizations frequently implement initiatives to address these issues. Son preference is a special kind of intra-household gender discrimination which is typically not considered in the initiatives targeting the household level. Addressing the issue is necessary because the presence of son preference is more pronounced in South Asian countries in comparison to societies in other parts of the world (Williamson, 1976; Filmer et al., 2008). For example, the widespread presence of son preference in India does not seem to be reduced by economic development and wealth (Pande et al., 2006). In Pakistan, the preference for the male child is so strong that mothers stop breastfeeding significantly earlier if

[1] According to World Health Organization (WHO), maternal mortality is the number of female deaths that occur from complications during or after 42 days of pregnancy.

the child's gender is female, in the expectation of quickly conceiving another child that is a son (Hafeez and Quintana-Domeque, 2018). In the context of Bangladesh, Kabeer et al. (2014) found that 40% women expressed preference for a son, while the preference for a daughter was only 7%. Jayachandran and Pande (2017) show how parents' decisions of further childbearing and allocation of resources among children are influenced by the preference for the eldest son in India, which also affects the height of the next children, who stay behind in the birth order.

In this paper, we focus the impact of son preference on individual and household-level welfare in rural Bangladesh. We thereby contribute to the discussion of gender-discriminatory household preferences based on the child's gender. More specifically, we investigate how the health outcomes of first-born children vary depending on the gender of the second-born, if and if not, the second one is a parent's preferred male child. Second, we analyze how children's gender composition in the family affects the outcome of the Feed the Future (FTF) program, a US government's initiative aiming to increase household income and food security.

We provide causal evidence on how the gender of a second-born child affects both individual and household welfare related to consumption and health indicators. The first part of our analysis explains the variations in the anthropometric outcomes of first-born children, depending on the gender of their immediate younger sibling. Operating under the assumption of a predominant preference for sons in rural Bangladeshi society, we explore the hypothesis that parents exercise greater diligence in nutritional care towards first-born when a second child is female, which is less preferred over male. At the same time, in the presence of a female child, there is a reduced emphasis on overall food quality within the household, reflecting a potential bias in the distribution of resources based on the child's gender. This discriminatory approach in parental action results in differential health outcomes for first-born children. Also, we infer that the second-born female child has a heterogeneous influence on the health of first-born male and female children. A first-born male with a second-born sister potentially receives more parental attention, benefiting more in terms of food share than a first-born female with a second-born sister, who receives relatively less attention. Our analysis further explores that the presence of a second-born female child does not cause parents to increase overall food consumption within the household. This restrained behavior eventually influences overall per capita consumption within the family.

To analyze the impact of the second-born childs's gender on the anthropometric outcomes of first children, we use a regression framework where we assume that the gender of second-born children in a family is random. This approach aligns with Brenøe (2021), who used a similar identification strategy (randomness of the second child's gender) to show how Danish women with a brother enjoy better gender conformity. A similar argument has been used in several other recent pieces of literature, such as Cools and Patacchini (2019) and Peter et al. (2018). Here, we have ruled out the incidence of sex-selective abortion in the sample. Selective abortion is not only legally banned in Bangladesh; it is also socially and religiously condemned. Additionally, in Bangladesh, the sex ratio at birth (SRB) is around 104 males per 100 female birth, which is in line with the natural rate of 105 males per 100 female birth (Talukder et al., 2014). To our knowledge, no study or statistics support the incidence of significant sex-selective abortion in Bangladesh. Lastly, the cost of sex-selective ultrasound and abortion seems beyond the capacity of rural households. Based on this, we can rule out the case of selective abortion and consider the gender of the second-born child as random. The first part of the analysis includes a sample of 1,747 first-born children aged 5-16 with at least one younger sibling who is also in that age range from the Bangladesh Integrated Household Survey (BIHS)-2011 (Ahmed et al., 2012), which is representative of rural Bangladesh. In addition to ordinary least squares, we also utilize quantile regressions in this part of the analysis. It allows us to investigate the effect of the gender of the second-born sibling (our treatment variable) on different outcomes of the first-born child at various points of the distribution. Using this, we explore whether the effect of having a second-born sister varies among first-born children who belong to distinct health outcome spectrums.

In the second part of our analysis, we focus on examining the household-level dynamics. Here, we investigate how households shape food consumption decisions in response to interventions designed to enhance food security and household income. Specifically, we investigate whether improved access to resources leads to increased diversity in household food consumption with separate child gender composition. This part is focused on determining whether the presence of a second-born female child influences intervention program outcomes. Our goal is to demonstrate whether better resource accessibility contributes to more varied dietary choices within households in a symmetric way and if this effect is mediate by the gender composition of the children, especially in the context of a family with a second-born female child.

We employed the difference-in-difference-in difference (DDD) framework to examine the influence of child gender on the outcome of an external intervention program on household consumption. We used two waves of Bangladesh Integrated Household Survey (BIHS) data from 2011/12 and 2018/19 (Ahmed et al., 2012; International Food Policy Research Institute (IFPRI), 2020). The BIHS-2011/12 is a baseline and BIHS-2018/19 is an endline survey. The study sample comprises households with a minimum of two children.[2] Between the two survey periods, a subset of households were beneficiaries of the Feed the Future Bangladesh (FTF) program a large-scale intervention initiative by the United States Agency for International Development (USAID); see also section 1.A.2 in Appendix). The primary objectives of the FTF were to increase household income, improve agricultural production, and enhance nutrition among people. To evaluate the impact of this intervention on households with different second-born child genders, we employ a DDD regression design. In this approach, the third difference comes from the presence of second-born female children in these households, operating under the assumption that the gender of the second-born is random, while the other two differences are household's participation status in the FTF program and the time period delineating observations before and after the program's commencement. This approach allows us to understand the distinctive effects of the FTF program, both in general and in the context of households with second-born female children.

The results of the first part of our study show an apparent change in the nutritional condition of first-born children aged 5 to 16 years when they are accompanied by a second-born sister. We discovered that the height-for-age (HFA) z-score of first-born children with a younger sister is, on average, 0.10 standard deviations higher than that of first-born children with younger brothers.[3] Our results also revealed the mechanism operating within the family food allocation dynamics, which seems to be influencing the observed increase in the HFA z-scores. Our investigation indicates that the proportion of daily calories consumed by the first-born child, relative to the total household consumption, increases by 0.5 percentage points when the second-born child is female. The calorie intake increases by 4 percent if the first-born is a male child and the second born is a female. These findings indicate that if the second child is a female, the first child is more likely to receive a greater share of

[2]To keep our household analysis connected with individual impact of son preference, we exclusively selected households that had a first-born child under the age of 16 during the baseline survey.

[3]HFA z-score is our key anthropometric variable which shows observed height of a child is how many standard deviations ahead or behind from the median height of reference group children of that age. Besides age group 0-5, this indicator is frequently used in literature for analyzing the nutritional status of the school aged children (Zhang et al., 2019; Sachdev et al., 2021; Eilat-Adar et al., 2021).

the family's food resources, reflecting better care by parents. This reallocation of resources within the family is a crucial component that contributes to the observed improvement in the HFA z-score of the eldest children, driven by preference for the son or, in other words, possible ignorance of the daughters. This ignorance is reflected in our result as we found that on average, per capita household calorie consumption is significantly reduced by 121.62 kilo calories if both first-born and second-born child are female, while there is no impact on household per capita calorie intake if first-born is a male. In our household-level analysis, we found that large-scale intervention programs to address food security and income do not seem to diversify household food consumption, specifically in cases where there is a second-born daughter within the family. Following the Food and Agricultural Organization guideline, we constructed the Household Dietary Diversity Score (HDDS) using all the foods consumed by households in the reference period (Swindale and Bilinksy, 2006; Kennedy et al., 2011).[4] The difference-in-difference-in-difference estimate shows that families with second-born females significantly lose 0.52 in HDDS score.

With these results, our article discusses several strands of literature related to the economics of gender. Firstly, our study contributes to the literature discussing parental attitudes towards child gender and its outcomes (Dahl and Moretti, 2008; Filmer et al., 2008; Jayachandran and Pande, 2017). We make specific contribution to the existing literature on son preference and its implication on child health by unveiling gender-discriminatory household decisions based on the second child's gender. While Jayachandran and Pande (2017) provides robust evidence that the health outcomes of the later-born child can depend on the preference for the older son in India, we provide evidence that the nutritional status of older children can also depend on the gender of the later-born child in rural Bangladesh. Here, our contribution is two-fold. Not only we unveil the impact of son preference or daughter neglect on an individual health level, but we also document the impact of this gender bias on consumption at the household level with different sibling sex compositions. Secondly, our research adds significant value to the current academic discourse on the various determinants that impact the accumulation of human capital in children, residing in low or middle-income countries (LMIC) Jayachandran (2015); Jakiela et al. (2020); Andrabi et al. (2023). Specifically, our research focuses on the aspect of the child health accumulation process with different sibling gender compositions.

In Bangladesh and other LMICs, there is a notable gap in the literature regarding

[4]HDDS is a 12 points score based on 12 food groups consumed in the last 24-hour period.

the health outcomes of school-aged children (5-15 years), as most studies concentrate on children up to 5 years old. For the case of Bangladesh, to the best of our knowledge, this is the first quasi-experimental study on the impact of son preference, where we provided empirical evidence of the heterogeneous effect of child gender on household food allocation and child health. This research causally unveils how the preference of the son affects the health of school-aged children through a lower overall nutrient intake and informs us of the direct implication of the neglect of the daughter, which many families are unaware of. The policy implications of these findings are very large as developing countries seek to improve the health and nutritional condition of all children. Specifically, the results will help combat the incidence of malnutrition among children in rural communities where preferences for the male child are noticeable. It will help reduce gender discrimination within the household by promoting extra care for Bangladeshi female children who are lower in birth order. Result also advocates the need for diversified food consumption in families with female children. Our result suggest that, while designing large-scale community level development interventions in developing countries, policymakers should consider household gender composition and gender norms to benefit all families similarly in society.

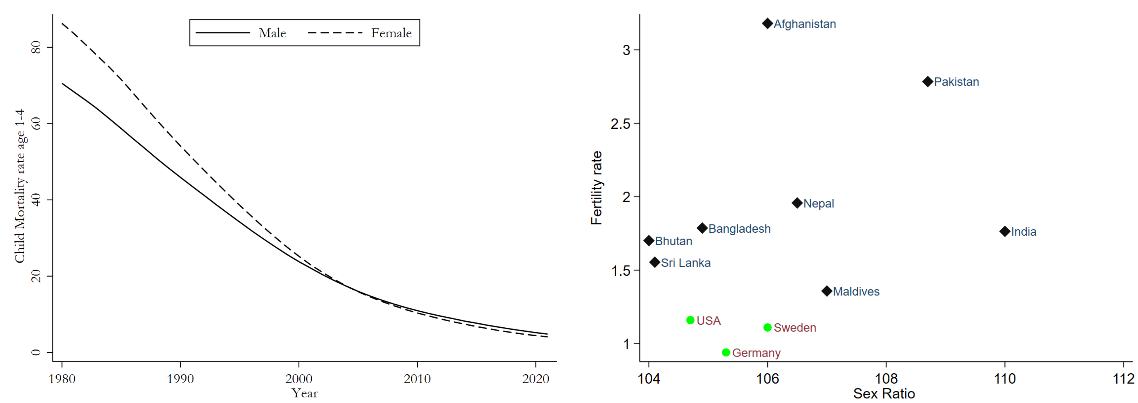
The rest of the document is structured in the following manner: Section 1.2 gives a detailed background of son preference, malnutrition, and the Feed the Future Program in Bangladesh. Section 1.3 describes the identification strategy for the son preference effect. Section 1.4 outlines the data source, sample collection process, variable description, and descriptive statistics. Section 1.5 presents detailed results. Section 1.6 is the limitation, policy implication, and conclusion of the study.

1.2 Background

1.2.1 Son preference in Bangladesh

Due to various socio-cultural and historical reasons, South Asian countries are the hub where a son is preferred to a daughter (Sen, 2002; Das Gupta et al., 2003; Attané, 2006; Jayachandran, 2015; Rai et al., 2014). Several studies investigating the cause and effect of son preference in South Asia showed that preferring sons over daughters imposes various derogatory effects ranging from maternal mortality to child mortality (Arnold et al., 1998; Self and Grabowski, 2012; Milazzo, 2018). Das Gupta et al. (2003) finds that the existing social security system in some Asian countries makes a woman feel vulnerable at old age, and she wants her son to protect her at that age.

In addition, the prevalence of the dowry system in society during the marriage of a female child imposes the financial burden of having daughters that makes parents want a son instead of a daughter (Diamond-Smith et al., 2008). For all these reasons, the preference for the son is also prominent in Bangladesh. One of the indicators of son preference in society is the differential rate of child mortality among male and female children. A higher mortality rate for female child is an indicator of unequal healthcare, nutrition, and other resources. We demonstrated this historical trend in the child mortality rate for children aged 1-4 in Bangladesh using data from Chao et al. (2023) in Panel A of Figure 1.1. It shows long-term elevated mortality rates among female children within this age range in Bangladesh. However, current patterns indicate a subtle decrease in the death rate among female children.



(a) Child mortality trends in Bangladesh (b) Fertility rates and sex ratios across countries

Figure 1.1: Child mortality, fertility rate, and sex ratio trends in South Asia

Notes: Panel A presents the child mortality rates per 1,000 live births for male and female children aged 1-14 in Bangladesh from 1980 to 2021. Panel B depicts the birth rate trends and sex ratio at birth in South Asian countries, as per the World Development Indicators-2019. The sex ratio is defined as the number of male births per 100 female births.

Simultaneously, an imbalanced sex ratio at birth acts as additional noteworthy evidence of a preference for male children. This imbalance in sex ratio is often a consequence of sex-selective abortion practices, which symbolize the use of sonographic screening during pregnancy and birth control measures depending on children's gender. Although the preference for the son is present in Bangladesh, it does not lead to selective abortion due to social and religious (Islamic) norms. Kabeer et al. (2014) examines how the context of son preference in Bangladesh changed differently from India, especially since the 1980s. The study stated that selective abortion contributed to an imbalanced child sex ratio in India, while the decrease in fertility reduced the gender gap in Bangladesh in the last few years. Panel B of Figure 1.1 provides an overview of South Asian countries' fertility rates and sex ratio at birth in 2019 relative to some of The Organisation for Economic Co-operation and Development (OECD) countries.

Traditionally, the presence of son preference is higher in patriarchal societies, which is a general practice in Bangladesh. Bangladeshi rural society is an agricultural society in which almost all households are engaged in agriculture in rural communities. A male child can help parents on their farm fields. Due to cultural and religious reasons, a female child is generally not allowed to help in farming fields. In addition to farming, male offspring support their parents in their businesses, which is also not common exercise for a female child. Another major reason for preferring a son over a daughter is the widespread prevalence of dowry in South Asian society. The dowry system makes parents concerned when a girl is born, as the parents know that they have to pay a significant amount of money during the marriage of that daughter. According to Diamond-Smith et al. (2008), the potential financial burden resulting from the dowry in a marriage is one of the main drivers of the preference of the son. Like any patriarchal society, boys enjoy a higher social value than the girl in Bangladesh. Previous studies confirm that the presence of son preference contributed to a high fertility rate with a negative relation to contraceptive use among women (Amin and Mariam, 1987). However, in recent years, Bangladesh has experienced positive changes in different socioeconomic indicators. Girls receive more school education, more women are involved in the workforce, and mortality of children under five years of age decreased to 41% of live births in 2014, which was 82.5 per 1,000 in 1994 (Khan et al., 2021). Asadullah et al. (2021b) demonstrated that the stated preference for a son in Bangladesh has declined, but in practice, fertility decisions are still determined by son preference; for more, see section 1.A.1 in Appendix.

1.2.2 Malnutrition in Bangladesh

The association between poverty and malnutrition has been well proven, and the case of Bangladesh is no exception. 56. 7% of Bangladesh's population were below the poverty line in 1991-1992, which fell to 20. 5% according to Asian Development Bank (2021). The incidence of severe acute malnutrition is also prevalent in Bangladesh. The Multiple Indicator Cluster Survey 2019, conducted by the Bangladesh Bureau of Statistics and UNICEF, found that 28% of Bangladeshi children were suffering from chronic malnutrition in 2019, which was 42% in 2013. Hosegood and Campbell (2003) studied the case of Bangladeshi women, which indicated that lower BMI of women at reproductive age is associated with high maternal mortality. They found that in the lowest 10% of the decile distribution of BMI, the risk of death is highest among women. Helen Keller International (2016) summarizes the results of the Bangladesh Food Security Nutritional Surveillance Project and reported that around 70% girls in the age group 10-18 are mild to severely short. It also revealed that around 50% of girls are mild to severely undernourished in terms of body mass index. Bangladeshi women are not only underweight and short; they also suffer from various related diseases as they are severely malnourished. For example, a large part of Bangladeshi women suffer from anemia. Rahman et al. (2021) reported 41.8% women of reproductive age was anemic in 2011 in Bangladesh.

1.2.3 Feed The Future Bangladesh

As a key component of the worldwide campaign against hunger, poverty, and malnutrition, Feed the Future (FTF) is a global initiative from the government of the United States of America (Lawson et al., 2016). The United States Agency for International Development (USAID) led the FTF initiative supports root-level farmers, women, and children in getting out of poverty and hunger. USAID partnered with the Bangladesh government to explore the potential of agriculture-led growth and improve the lives of Bangladeshi people. The whole program in Bangladesh was finalized in March 2011. One of the key objectives of the FTF Bangladesh program was to improve the nutritional status of people in rural areas. Bangladesh's south and southwestern regions were selected as the operational areas of the FTF. Our study exploited this intervention as a quasi-experimental variation, as only households in this zone encountered the FTF program; see section 1.A.2 in Appendix for more details on the FTF program.

1.3 Method

1.3.1 Identifying son preference effect on first-born child

The aim of this research is to systematically estimate the causal effect of son preference, specifically by analyzing the presence of second-born female children in the family. Our assumption is that first-born children, get extra parental care, which possibly result in their better health status if the second-born is not a preferred male. In other words, the parent's attention shifts towards first-born when the second-born child is female. To identify this son preference effect, we admit the potential selection in the child sex composition in the family. Children's sex composition in a family is not always exogenous, as parents' decision on having further children may depend on the sex composition of existing children. So, it would not give a causal effect if we estimate the next children's outcome on the gender of older children based on which parents might took the decision of another child. One potential solution is to think the other way and focus on the randomness of the gender of the second child. The randomness of the gender of the second child allows us to estimate its causal impact on the health outcome of the first-born, because the birth of the first child was not based on the gender of the second-born. By ruling out the case of selective abortion, the assumption implies that the gender of second-born children is random.[5] This randomness allows us to perform a causal analysis of son preference on child health and family-level outcomes. Our identification argument is similar to Brenøe (2021), who utilized Danish administrative data under the assumption that the gender of a second-born child is a random phenomenon and assessed the gender conformity of first-born women with a second-born brother compared with first-born women with a second-born sister. Using the random variation in the gender of different sex twins in Sweden, Peter et al. (2018) shows how the earning of the men with brothers is higher compared to men with sisters. To study the impact of a younger brother on female earning, Cools and Patacchini (2019) also uses the argument that the next youngest child's gender is also random. This study was conducted on U.S. women and found that in adulthood, they earn seven percent less if they have a younger brother. In line with these literature, we also assume that the gender of second-born children is random and there is no sex section among our sample. Our balance table confirms our assumption that there are no significant differences in the predetermined covariates among families with second-born male and second-born female children, except for their gender.

The specification for analyzing son preference effect on the health of first born child's health is

$$Y_i^{\text{First-Born}} = \alpha_0 + \alpha_1 \text{ D}_i + X_i'\delta + \epsilon_i, \tag{1.1}$$

where $Y_i^{\text{First-Born}}$ is health-related outcomes of first-born child; Dummy D_i=1 if second-

[5] Abortion is not legal in Bangladesh and selective abortion seems not to occur in Bangladeshi society (Kabeer et al., 2014; Talukder et al., 2014).

born is a female, and 0 for male; X_i = is vector of regressors such as mother's age and school year, total number of children in a family, household income, first child's age and gender, second child's age, household head age and school year; and ϵ_i is the error term.

For analyzing the incidence of stunting among children, we used a linear probability model. In the linear probability model, our dependent variable $P_i^{\text{First Born's stunting}}$ takes the value 1 if the first-born child is stunted and 0 otherwise:

$$P_i^{\text{First Born's stunting}} = \alpha_0 + \alpha_1 \text{ S }_i^{\text{Second-Born}} + X_i'\delta + \epsilon_i. \qquad (1.2)$$

1.3.2 Identifying son preference effect on household

For our study of the impact of son preference on household level, we used the FTF program as a case study. We explore changes in household consumption patterns among households with FTF interactions, specifically in cases where there are second-born female children in families. To measure these impacts, we first employed the difference-in-difference (DiD) regression methodology. Our data allows the use of the DiD framework because households outside of southern Bangladesh were completely unexposed to the program before and after the FTF started in 2011. Only households in the southern region of Bangladesh (panel B of Figure 1.A.2 in Appendix) were exposed to the FTF program. With two periods and two groups, the best strategy to capture the impact of FTF is the use of the difference-in-difference (DiD) approach (Card and Krueger, 1994; Heckman et al., 1998; Athey and Imbens, 2006; Abadie, 2005).

Using the DiD method, we estimated the average treatment effect of the FTF program on the household food consumption indicators of FTF households (treated) after the treatment period in 2018. In short, our DiD estimator gives the following change in the outcome of household food consumption Y, which is caused by the intervention of the FTF program. The Difference-in-Difference (DiD) estimate is $[(Y_{ftf,2018} - Y_{ftf,2011}) - (Y_{con,2018} - Y_{con,2011})]$. here $Y_{ftf,2018}$ and $Y_{ftf,2011}$ are the household food consumption outcomes for the FTF program regions for the year 2018 and 2011 respectively. $Y_{con,2018}$ and $Y_{con,2011}$ are the outputs for the control group household from non-FTF zones who did not receive treatment between the year 2018 and 2011. The DiD model specification is as follows:

$$Y_{it} = \beta_1 + \beta_2 (\text{ treat }_i) + \beta_3 (\text{ post }_t) + \rho (\text{ treat }_i \times \text{ post }_t) + X_i' + \epsilon_{it} \qquad (1.3)$$

where $Y_{it}$is the outcome variable of household- (Household Dietary Diversity Index) in period t; treat is the treatment dummy. which takes value 1 if the household is a Feed the Future household, 0 otherwise; post is the time dummy. It takes the value 0 if the period is before(2011) and 1 if the period is after(2018) the program; treat x post is the interaction term of treatment and time dummy from which we estimate DiD. X_i' is the control variables including the size of the household, the age and education of the household head, the regional fixed effect, the year fixed effect.

By specifying this framework, we estimate the average treatment effect of the FTF program on the households food consumption who were part of the program. We also capture the FTF treatment effects on households with different sibling gender combinations. That is, we intend to see what happens to a household's food consumption decision in the presence of son preference and external intervention programs. To capture these two effects, we use the Difference in Difference in Difference (DDD) or triple difference method, which was first used by Gruber (1994), later DDD framework has become popular for capturing the heterogeneous effect of an intervention. For example, Ravi and Engler (2015) used the DDD framework to evaluate the impact of an Indian anti-poverty program on savings, food security and health. Bandiera et al. (2017) used DDD specification to analyze the difference in program outcome in ultra-poor and non-poor households in a large-scale randomized control trial (RCT), where livestock asset and skill intervention was provided to Bangladeshi village women.

Our assumption states program effect might vary depending on the gender composition of the participant household. When there is a second-born girl in the household, parents might be less interested in spending on better food and education even after their enhanced purchasing capacity through external interventions like feed the future. In this context, we add another treatment, the gender of the second child, which is random. The DDD framework yields: $[(Y_{ftf,f,\ 2018} - Y_{ftf,f,\ 2011}) - (Y_{con,f,\ 2018} - Y_{con,f,2011})]$ $-[(Y_{ftf,m,\ 2018} - Y_{ftf,m,2011}) - (Y_{con,m,2018} - Y_{con,m,2011})]$. Here, $Y_{ftf,f,\ 2018}$ and $Y_{ftf,f,\ 2011}$ is the outcome for FTF households with a second-born daughter in 2018 and 2011 respectively. While $Y_{con,f,\ 2018}$ and $Y_{con,f,\ 2011}$ is the outcome for control households (non-FTF) with the second born daughter in 2018 and 2011. $Y_{ftf,m,\ 2018}$ and $Y_{ftf,m,\ 2011}$ is the outcome for FTF households with second born male child for year 2018 and 2011. $Y_{con,m,\ 2018}$ and $Y_{con,m,\ 2011}$ is the outcome control households with second born male child for year 2018 and 2011. The DDD model takes the following form:

$$Y_{it} = \beta_0 + \beta_1 treat + \beta_2 female + \beta_3 \text{ post } + \beta_4 treat \times female +$$
$$\beta_5 treat \times \text{ post } + \beta_6 female \times \text{ post } + \beta_7 treat \times female \times \text{ post } + X_i' + \epsilon_{it} \quad (1.4)$$

Here Y_{it} is the outcome variable of households- (Household Dietary Diversity Score) in period t; treat is FTF treatment dummy, it takes value 1 if the household is in FTF regions, 0 otherwise; post is Time dummy which takes value 0 if the period is before (2011) and 1 if the period is after(2018) the program; female is the gender treatment dummy- 1 if the gender of the second child of the household is female (which is random) and 0 otherwise; treat x post is the interaction term of ftf treatment and time dummy; treat x female is the interaction term of ftf treatment and gender treatment; treat x post x female is the interaction term of ftf, gender treatment, and time dummy- from which we estimate DDD; X_i' is control variables including household size, Household head's age, and education, regional fixed effect in village level, year fixed effect.

1.4 Data sources and sample construction

1.4.1 Data sources and sample construction

For our analysis, we use data from two rounds of the Bangladesh Integrated Household Survey (BIHS). The BIHS is a national representative dataset of rural Bangladesh. Experienced researchers from IFPRI and Bangladesh Bureau of Statistics (BBS) constructed and designed this survey, which is one of the very few nationally representative and comprehensive surveys conducted in Bangladesh. This dataset is claimed to be the only dataset in Bangladesh that collects dietary intake and anthropometric measurements for all households, while other data sets contain this information only for a particular age group or gender. The BIHS data was collected on same households over three periods; the first wave was conducted in 2011, before FTF, and considered baseline. The midline data was collected in 2015, and the final wave of data was collected in 2018. The dataset has a total sample size of 6,500 households, of which 2,040 are for the FTF zone.

We constructed several dependent and independent variables from the first and third waves of the BIHS survey to investigate the effects of son preference at the children and household levels. For individual analysis, we used data from BIHS-

2011 wave, which was collected before the FTF intervention. Our respondents are first-born children aged 5 to 16 years with at least one younger sibling with the same age group. In the sample, 2,556 first-born children under 16 years of age had at least one younger sibling, and 1,747 first-born children had a younger sibling of the same age group (5-16). For household level analysis, we constructed a balanced panel of households from first (2011, before the FTF intervention) and third round (2018, after the FTF intervention) of BIHS survey. Several households in the baseline survey experienced split as a result of children growing up, getting married or starting as a new household on their own; these households were removed from the panel dataset in order to maintain the integrity of our analysis. To maintain methodological rigor, we employed a differentiated approach in sample construction for household and individual-level dynamics of son preference. At the household level, we included all first-born children below the age of 16 instead of 5-16. This age range was crucial for an accurate intention-to-treat (ITT) analysis. Restricting the age range to 5-16 years at the household level could exclude younger second-born children, whose presence and gender are key variables in our analysis of the outcome of the FTF program. Finally, our 2 x 2 (two period , two group) panel includes data of 1,960 households.

Height-for-age (HFA) z-score is our key anthropometric variable which we use as indicator of child health outcome. We used the WHO 2007 growth reference to construct our anthropometric variables.[6] Additionally, based on the HFA z-score, we calculated the incidence of stunting among children. The use of z-scores helps us to understand the overall situation as the distribution is from the global reference population. For household-level analysis, we have generated food consumption indicator variables, such as household dietary diversity scores (HDDS) from dietary data following the guideline Gina Kennedy (2011). Table 1.B.1 in Appendix describes the main outcome variables used in our study.

1.4.2 Descriptive statistics

To utilize the gender of the second-born child as an accurate random variation, it is crucial to ensure that other factors within the family setting are completely balanced. Our balace Table 1.1 supports the above-mentioned assumption, as the predetermined

[6]The WHO 2007 growth reference is a comprehensive set of growth standards and charts developed by the World Health Organization, which provides guidelines for evaluation and monitoring of the physical growth, development, and nutritional status of children and adolescents globally. see more on - `https://www.who.int/tools/growth-reference-data-for-5to19-years/application-tools`

Table 1.1: Balance table

	All	Child ($\leq 16)$*sample*			Experiment(5-16) sample		
	All	2nd-born male(C)	2nd-born female(T)	diff (T - C)	2nd-born male(C)	2nd-born female(T)	diff(T - C)
Panel A: Pre-determined covariates							
Household head age	38.28	38.39	38.16	0.23	40.92	40.44	0.48
	(8.27)	(8.55)	(7.99)	[0.71]	(8.22)	(7.59)	[1.28]
Household head school year	3.28	3.30	3.26	0.04	3.10	2.98	0.12
	(3.88)	(3.88)	(3.88)	[0.26]	(3.78)	(3.82)	[0.65]
Currently married hh head(=1)	0.97	0.98	0.97	0.01	0.97	0.96	0.01
	(0.16)	(0.15)	(0.17)	[1.14]	(0.17)	(0.19)	[0.83]
Mother's School Year	3.63	3.71	3.56	0.15	3.30	3.14	0.16
	(3.52)	(3.53)	(3.51)	[1.10]	(3.51)	(3.43)	[0.95]
Mother Age	31.90	31.77	32.03	-0.26	33.96	34.31	-0.34
	(6.43)	(6.41)	(6.46)	[-1.02]	(6.06)	(5.96)	[-1.19]
Religion(Muslim=1)	0.89	0.89	0.90	-0.01	0.89	0.90	-0.01
	(0.31)	(0.31)	(0.30)	[-0.42]	(0.31)	(0.30)	[-0.44]
Ethnicity(Bengali=1)	1.00	1.00	1.00	0.00	1.00	1.00	0.00
	(0.05)	(0.05)	(0.06)	[0.34]	(0.05)	(0.06)	[0.42]
Panel B: Outcome and mediator variables							
Household size	4.67	4.61	4.72	-0.11**	4.77	4.91	-0.13**
	(1.16)	(1.10)	(1.21)	[-2.42]	(1.15)	(1.25)	[-2.29]
Total Children in HH	2.57	2.51	2.63	-0.13***	2.70	2.87	-0.18***
	(0.82)	(0.76)	(0.86)	[-3.96]	(0.83)	(0.92)	[-4.18]
Household head is male	0.82	0.83	0.80	0.03*	0.83	0.79	0.04**
	(0.38)	(0.37)	(0.40)	[1.95]	(0.38)	(0.41)	[2.20]
hdds	6.83	6.82	6.85	-0.03	6.81	6.80	0.00
	(1.35)	(1.34)	(1.36)	[-0.51]	(1.33)	(1.39)	[0.05]
Per capita calorie intake	2156.95	2179.65	2134.85	44.80*	2258.71	2195.15	63.56**
	(587.75)	(610.51)	(564.06)	[1.93]	(613.52)	(562.79)	[2.25]
Monthly food expense/pc	1390.74	1418.15	1364.05	54.10**	1430.62	1353.94	76.68**
	(692.54)	(735.08)	(647.60)	[1.97]	(738.50)	(627.12)	[2.34]
Monthly nonfood expense/pc	931.13	929.64	932.58	-2.94	967.65	949.74	17.92
	(666.24)	(691.26)	(641.21)	[-0.11]	(706.27)	(652.04)	[0.55]
Panel C: First child's health outcomes							
Height-for-age z-score	-1.23	-1.24	-1.22	-0.02	-1.24	-1.18	-0.06
	(1.19)	(1.24)	(1.15)	[-0.39]	(1.19)	(1.14)	[-1.07]
BMI-for-age z-score	-1.07	-1.02	-1.11	0.09*	-1.04	-1.05	0.01
	(1.12)	(1.14)	(1.10)	[1.85]	(1.11)	(1.06)	[0.22]
Weight-for-age z-score	-1.55	-1.50	-1.60	0.11	-1.51	-1.54	0.02
	(1.12)	(1.14)	(1.11)	[1.43]	(1.18)	(1.14)	[0.20]
Observations	2,556	1,261	1,295	2,556	861	886	1,747

Notes: The balance table presents variables at the household and individual level for all first-born children between the ages of 5 and 16 who have at least one younger sibling. It compares the baseline characteristics and outcomes between households with a second-born male ("2nd-born male") and those with a second-born female ("2nd-born female"), offering perspectives on potential differences depending on the sex of the second children. The "Experiment(5-16)" sample is our experiment sample of first-born children in the specified age range whose younger sibling, the second-born, is also between 5 to 16 years old. Means and standard deviations (in parentheses) are presented for each subgroup; t-statistics for differences between groups are provided in square brackets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

covariates of households such as household head's age, schooling year, marital status, mother's age, schooling year, and religion do not significantly vary among households with different sibling sex composition. We compare essential family characteristics based on different gender compositions of siblings. For this purpose, we first compared children in the following two groups: (i) first-born any gender and second-born male; (ii) first-born any gender and second-born female. Then we also compare both children male and both female. The comparisons provide some important rationale for further analysis. Most importantly, it validates our primary assumption: the randomness of the sex of a second-born child in the household.

We confirm that predetermined covariates do not significantly vary among families with second-born children's different gender compositions, thus ensuring the validity of randomness assumption. Table 1.1 shows the descriptive statistics and the balance in the predetermined covariates in our experiment sample, as well as all children sample. The average age of the household head is around 40 years in our sample which does not vary by gender of second child. At the same time, the average year of education is only around three years among the heads of the households. Almost all household heads are currently married. The mothers also have very low education, on average less than four years. The religion of the majority of households is Muslim, and the ratio is in line with the national rate of approximately 90 percent Muslim. The ethnicity of almost all the respondents are Bengali. None of these covariates significantly differs bases on second-born's gender.

Panel B of Table 1.1 shows the average values of household-level outcomes and mediator variables by sibling gender composition. The household size is significantly larger when the second-born children is female. In our experimental sample, the average number of children per household is 2.78 when the second child is female, which is the highest value among all four groups. This also supports the inference that there is a widespread preference for sons in Bangladeshi society. When the first two children are girls, parents seek further children in expectation of a boy, contributing to the increase in family size. The more female children there are, the larger the size of the family, but with the limited resources rural families then struggle to meet basic needs. Panel B further shows that per capita calorie consumption is the lowest among the households where the first two children are female. In addition, monthly household food expenses are the lowest in this group, which implies that households with second-born female children spend less on food. The Panel C shows the mean value of the height-for-age, weight-for-age, and BMI-for-age z-scores for the first-born children in the age group of 5-16 years old.

1.5 Results

1.5.1 Son preference and child health

In this section, we discuss the causal implication of parental preference for male children on the health outcomes of first-born child. We begin by presenting the assumptions, followed by our results. Our first assumption is that parental attention shifts towards the first-born when the second child is not of the preferred male gender. This shift in attention and care poetically results in a better health outcome for the first-born child. To empirically evaluate this supposition, we examine two primary indicators of first child's health well-being: the height-for-age (hfa) Z-score and incidence of stunting. These measures serve as an indicator of the extent of parental care and attention provided to the eldest child. Our treatment indicator is the presence of a second-born female child in the family, with the assumption of randomness of the child's gender. Our experimental sample comprises first-born children aged 5-16 years who have at least one younger sibling within the same age range of 5-16 years. We exclude first-born with second-born sibling under 5 years of age, as their dietary share and household consumption patterns are not comparable with those of the first-born child who has a sibling of above age 5. The mean HFA z-score of our sample with a second-born sister is -1.18, which is higher than the mean HFA z-score score (-1.24) of the first-born with a second-born brother. These variations are visually illustrated in Figure 1.2, which displays a set of box plots showing the HFA z-scores of first children for each the ages from 7 to 16 years.[7]

The box plots visually compare the HFA z-scores of first-born children, categorised by the gender of their subsequent sibling. We detect some interesting patterns in this indicator throughout different ages. At age 7, the median HFA z-scores of first-born children with younger brothers (control group, marked as "0") and those with younger sisters (treatment group, marked as "1") shows slight deviation, with a small upsurge for control group. At age 8, the treatment group shows an increase in median z-scores. At age 9 to 11, the scores for both groups are nearly similar between the control and treatment groups, suggesting that there is little noticeable difference in these ages based on the gender of the younger sibling. For age 12-15, the graph indicates that first-born children with younger sisters tend to have higher HFA z-scores. At 16, the medians are aligned closely again, indicating that any differences detected in previous

[7]In the graph we omit 5 and 6 years old child as there is very few children with younger sibling with similar age range.

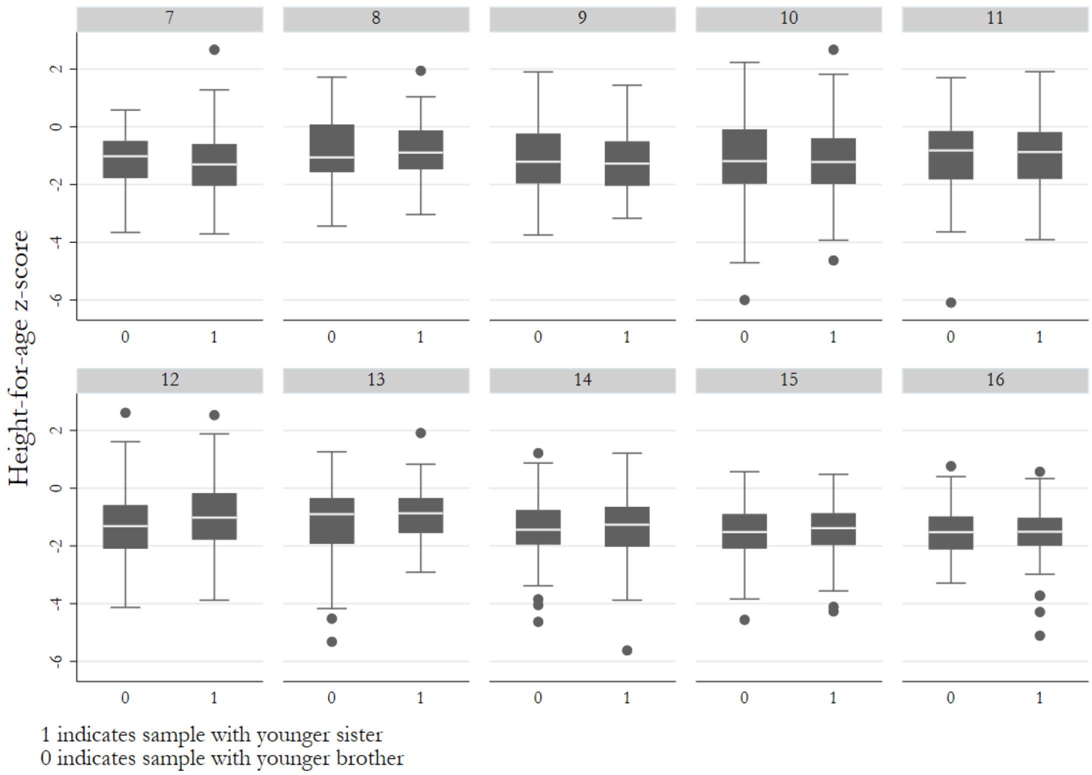


Figure 1.2: Height-for-age z-score of first-born child at different ages

Notes: This box plot illustrates the height-for-age (HFA) z-scores of first-born children ranging from 7 to 16 years old, categorized by the gender of their younger sibling. The graph omits first-born children younger than 7 due to the minimum age criterion of 5 for the younger siblings, due to limited sample. In horizontal axis, 1 indicates first-born children with second-born sister and 0 indicates first-born children with second-born brother. The vertical axis shows HFA z-scores.

years tend to fade away as children reach late adolescence.

In addition to the median-based illustration shown in Figure 1.2, we look at the kernel densities of the HFA z-scores to get a better idea of how they are distributed. The K-density analysis presented Figure 1.B.3 in Appendix captures the HFA z-score distributions highlighting the possible differences linked to the younger sibling's gender. The shape of the kernel density graph implies that the distribution is not entirely identical and varies based on the gender of the second child. The distribution shows a marginal, yet notable shift to the right for first-born children with a second-born sister.[8] This shift is particularly evident when the HFA z-score is negative. A move-

[8]Two-sample Kolmogorov–Smirnov (KS) test on our experimental sample ($p<0.1$, N=1,605) confirms that the distribution is significantly different for both groups.

ment on the right in the negative range of the distribution implies an improvement in nutritional status.

In Table 1.2, we analyze the effect of the second born's gender on the HFA z-score and incidence of stunting of the first-born. According to our identification strategy, we consider the gender of the second child as random (validated by balance Table 1.1), which allows us to causally explain the relationship of health indicators of the first-born child and the second child's gender. The result of least squares regressions supports our hypothesis and shows that having a second-born female sibling, on average, causes an increase in the HFA z-score of the first-born children by 0.10 standard deviations ($p<0.10$). When we divide the sample and perform the same regression for first-born males and first-born females separately, the z-score trends in the same direction but lacks statistical significance. As part of robustness check, we run similar regression with cluster around the PSUs (villages) and upazilas (sub-districts) and the result almost remain similar as shown Table 1.C.7 and 1.C.8 in Appendix. The linear probability model does not show any significant effect on the occurrence of stunting, but the negative coefficient suggests a potential inclination towards improved nutritional outcomes for first-born children when the second-born child is a female. In the regression models, we controlled for the mother's age and education, the first child's age and gender, the second child's age, household savings, land ownership, and household size. In addition, we include the districts which are the second highest level administrative regions in Bangladesh.

In order to examine the distinct impacts that younger female siblings may have on different quantiles of the first child's HFA z-score, we performed a quantile regression analysis. The quantile regression plot in Figure 1.3 and the accompanying Table 1.B.4 in the Appendix demonstrate that in lower quantiles, there are significant positive associations between the HFA z-score of the first child and the gender of the second born, but these associations decrease and become statistically insignificant at higher quantiles. The results implies that families on the lower nutritional status spectrum show more care for the first child when the second born is a female. Based on the above results, it can be inferred that when the second-born is female, there is a noticeable increase in parental attention towards the first-born child, which is reflects through their increased HFA z-score.

Table 1.2: First child's HFA z-score and incidence of stunting

	HFA z-score						Stunting					
	All		Male		Female		All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Second child is female	0.062	0.105*	0.061	0.068	0.061	0.118	-0.026	-0.030	-0.026	-0.024	-0.026	-0.026
	(0.058)	(0.058)	(0.088)	(0.088)	(0.077)	(0.079)	(0.021)	(0.021)	(0.031)	(0.031)	(0.029)	(0.031)
Mother's School Year		0.008		0.011		0.013		-0.002		-0.005		-0.002
		(0.011)		(0.017)		(0.015)		(0.004)		(0.006)		(0.006)
Mother Age		-0.002		-0.006		0.002		0.000		0.003		-0.002
		(0.009)		(0.014)		(0.011)		(0.003)		(0.005)		(0.004)
Household size		-0.083***		-0.030		-0.124***		0.015		0.002		0.026*
		(0.025)		(0.040)		(0.030)		(0.010)		(0.014)		(0.014)
Child Age		-0.116***		-0.099***		-0.144***		0.009		0.015		0.009
		(0.019)		(0.029)		(0.024)		(0.007)		(0.011)		(0.009)
Second child age		0.055***		0.082***		0.036*		-0.006		-0.019*		0.002
		(0.017)		(0.030)		(0.021)		(0.007)		(0.010)		(0.009)
Household head age		0.005		0.009		0.004		0.000		-0.002		0.001
		(0.006)		(0.010)		(0.008)		(0.002)		(0.004)		(0.003)
Household head school year		0.006		0.006		0.003		0.000		0.001		0.000
		(0.011)		(0.016)		(0.015)		(0.004)		(0.005)		(0.006)
Mean depvar	-1.21	-1.21	-1.13	-1.13	-1.28	-1.28	0.23	0.23	0.22	0.22	0.24	0.24
Division FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	1,605	1,599	734	730	871	866	1,605	1,599	734	730	871	866
R-Squared	0.001	0.120	0.001	0.127	0.001	0.192	0.001	0.073	0.001	0.112	0.001	0.106

Notes: Columns 1-6 of this table report the impact of second-born's gender on the height-for-age z-score of the first-born child while columns 7-12 show the effect on incidence of stunting. Stunting is a binary variable that indicates if HFA z-score falls below minus two standard deviations from the median of the World Health Organization (WHO) child growth standards. "All" denotes sample for all first-born children, irrespective of gender, while "Male" and "Female" segregate the sample into first-born male and first-born female children, respectively. All child is between 5-16 years old. Second child age range is also 5-16. Robust standard errors are in parentheses. Controlled for seven administrative division fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

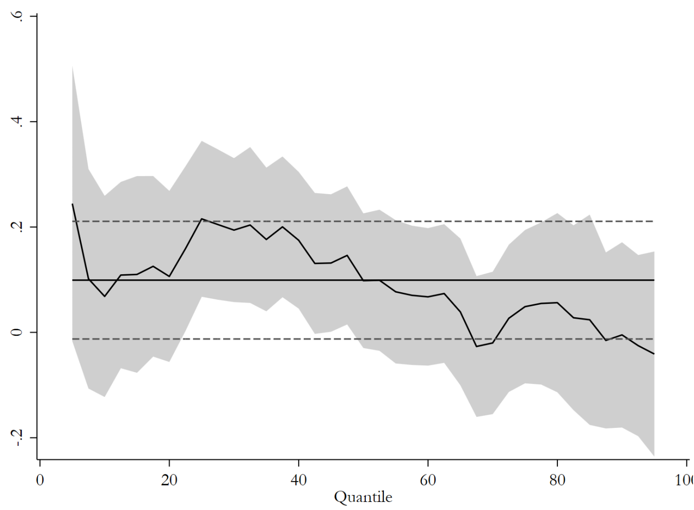


Figure 1.3: Quantile regression: HFA z-score

Notes: This figure displays quantile regression estimates of the impact of having a second-born female child on the HFA z-score of the first-born child, controlling for other covariates. The y-axis of the graph is depicting the coefficients of quantile regression. The dependent variable is HFA z-score of the first-born. The solid line represents OLS estimates, while the shaded area depicts the confidence interval from quantile regressions estimated at every 2.5 quantile between the 5th and 95th percentiles.

Change in first-born's calorie intake

Our second assumption is that the improved anthropometric outcomes observed in first-born children with a second-born sister are attributable to the mechanism of enhanced nutritional care provided by parents, based on shifted attention from the second-born female, which is assumed to be not a preferred gender. To assess our second assumption, we first investigate how the daily calorie share (share in total daily calorie consumption by the household) of the first-born children varies in presence of a immediate younger sister. We get evidence favor of our supposition that due to son preference, the firstborn receives an additional share of food when the second child is not a preferred male. The result of Table 1.3 supports our assumption, since the daily calorie share in total household calorie of the first child increases by a fraction of 0.005 or 0.5 percentage point ($p<0.05$) if the second child is female in the household.

The calorie share of male first-born individuals exhibits a higher overall increase compared to the female samples, although both are marginally insignificant. The analysis shows that the daily calorie intake of the first-born child is 3.1 percent ($p<0.10$, without controls) higher when the second-born child is female. Furthermore, this rise

Table 1.3: First child's calorie share and calorie intake

	Calorie share in household						Calorie Intake					
	All		Male		Female		All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Second child is female	0.004	0.005**	0.007	0.007	0.001	0.005	0.031*	0.022	0.049**	0.040*	0.015	0.015
	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.004)	(0.016)	(0.015)	(0.023)	(0.021)	(0.023)	(0.023)
Mother's School Year		-0.000		-0.001		-0.000		0.001		-0.003		0.005
		(0.001)		(0.001)		(0.001)		(0.003)		(0.004)		(0.005)
Mother Age		0.001**		0.000		0.002***		0.002		-0.002		0.007*
		(0.000)		(0.001)		(0.001)		(0.002)		(0.003)		(0.003)
Household size		-0.025***		-0.026***		-0.024***		0.003		0.019*		-0.007
		(0.001)		(0.002)		(0.002)		(0.007)		(0.011)		(0.010)
Child Age		0.006***		0.007***		0.005***		0.041***		0.042***		0.039***
		(0.001)		(0.001)		(0.001)		(0.005)		(0.007)		(0.007)
Second child age		-0.001		-0.000		-0.002*		0.005		0.008		0.001
		(0.001)		(0.001)		(0.001)		(0.005)		(0.006)		(0.007)
Household head age		-0.002***		-0.001***		-0.002***		-0.002		-0.001		-0.003*
		(0.000)		(0.000)		(0.000)		(0.002)		(0.003)		(0.002)
Household head school year		-0.001		-0.000		-0.001		-0.000		0.001		-0.001
		(0.000)		(0.001)		(0.001)		(0.003)		(0.003)		(0.004)
Mean depvar	0.20	0.20	0.21	0.21	0.19	0.19	1,906.14	1,906.14	1,985.19	1,985.19	1,838.78	1,838.78
Division FE	No		No	Yes	No	Yes	No		No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	1,693	1,687	795	792	898	893	1,693	1,687	795	792	898	893
R-Squared	0.001	0.319	0.003	0.314	0.000	0.353	0.002	0.208	0.006	0.259	0.000	0.192

Notes: Columns 1-6 of the table examine the impact of the second-born child's gender on the calorie share of the first-born relative to other household members, while columns 7-12 present the impact on the first-born's individual calorie intake amount. The regressions in columns 7-12 utilize the logarithmic transformation of the first-born's daily calorie intake as the dependent variable, with the 'Mean DepVar' statistic reporting the geometric mean consumption in kilocalories. "All" denotes sample for all first-born children, irrespective of gender, while "Male" and "Female" segregate the sample into first-born male and first-born female children, respectively. All child is between 5-16 years old. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

reaches 4.0 % ($p<0.05$) in the case where the first child is male.

We perform quantile regression at recurring intervals of 10 quantiles, starting from the 5th percentile. The quantile regression plot in Figure 1.4 and the results of the quantile regression analysis is presented in Table 1.B.5 in the Appendix. It provides evidence of a differential impact of the second-born child's gender on the distribution of the calorie share. The findings revealed significant effects in the 5th, 15th, 35th and 45th percentiles. The effect size diminishes and loses statistical significance as we progress towards the upper quantiles. Overall, the presence of a second-born female child is linked to a higher proportion of calories being allocated to first-born children in the lower range.

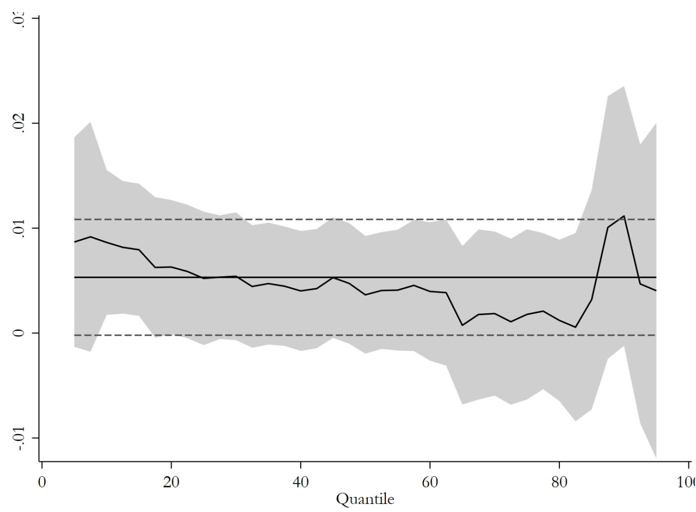


Figure 1.4: Quantile regression: First-child's calorie share in household

Notes: This figure displays quantile regression estimates of the impact of having a second-born female child on the calorie share of the first-born child, controlling for other covariates. The y-axis of the graph is depicting the coefficients of quantile regression. The dependent variable is daily calorie share of first-born in household total calorie consumption. The solid line represents OLS estimates, while the shaded area depicts the confidence interval from quantile regressions estimated at every 2.5 quantile between the 5th and 95th percentiles.

Change in household calorie consumption

In addition to analyzing the first-born's calorie share, our study also investigate the broader nutritional dynamics of the household in relation to the gender of a second-born child. The results shown in Table 1.4 indicate that the gender of the second-born child has a modest effect on the nutritional dynamics of the household. The presence of a second-born female is linked to a statistically significant drop in the amount of per capita calorie consumption in the household. Households, on average, experience a reduction of 57.27 kilocalories ($p<0.05$) in per capita consumption in the presence of a second-born female child. The decrease in calorie consumption per person is considerably more prominent, amounting to 121.62 kilocalories ($p<0.01$), in households where the first and second both children are female. Surprisingly, the presence of a second-born female child does not have a notable impact on the average per capita calorie consumption in households with first-born male children. A noticeable impact of the gender of the second-child on household total calorie consumption becomes more apparent in our further investigation. The result indicates that the total calorie consumption of households reduces by 290.26 kilocalories ($p<0.10$) if the gender of the second child is female. However, more precisely, households with both female child observe a decrease of 617.17 kilocalories ($p<0.01$) in total calorie intake. Households that have a first-born male show a rise in calorie intake, with an average increase of 110.46 kilocalories, although this change is not statistically significant.

These findings highlight a gender-specific response to household nutritional provision. The patterns suggest a noticeable gender bias, revealing a preference for sons when it comes to food consumption allocation. The fact that first-born receive a larger portion of food when there is a younger sister, along with the overall decrease in household consumption when both children are female. This indicates a consistent favor for male children.

1.5.2 Son preference and FTF intervention program outcome

Our third or final assumption states that large-scale intervention programs aimed at improving nutritional practice can result in distinct outcomes in households with different sibling gender compositions, and these differences are driven by the underlying dynamics of son preference. To examine our final hypothesis, we utilized two period data from the intervention program named Feed the Future (FTF) in the same households that we used in our previous analysis. We used triple difference (DDD) regressions to examine the impact of having a second-born female offspring among FTF households.[9] This investigation aimed to determine the impact of son preference on intervention programs, such as FTF, aimed at improving quality of life. The

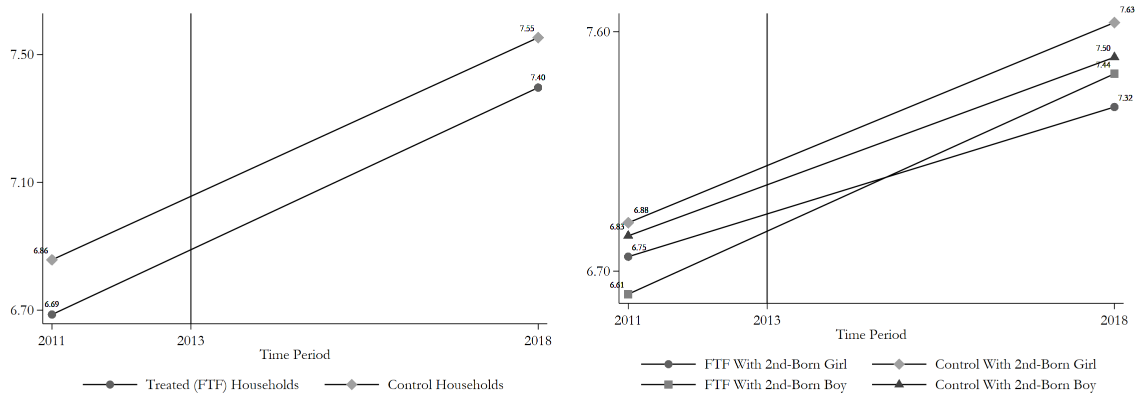
[9] In this part, we consider any second-born female children, including below the age of 5, up to the age of 16 at the time of the baseline survey as gender treatment. We assume that the gender of the child is a random phenomenon in Bangladesh.

Table 1.4: Household total and per capita calorie consumption

	Household per capita cal intake						Household total cal intake					
	All		Male		Female		All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Second child is female	-63.562**	-57.278**	36.518	14.628	-151.861***	-121.623***	-34.994	-290.263**	91.287	110.458	-142.385	-617.175***
	(28.190)	(26.662)	(40.790)	(36.824)	(38.693)	(39.213)	(169.724)	(128.875)	(245.939)	(173.870)	(234.292)	(194.839)
Mother's School Year		10.050*		3.766		12.959*		36.084		6.462		49.918
		(5.212)		(7.369)		(7.301)		(25.391)		(36.258)		(35.139)
Mother Age		7.649**		8.594		7.321		32.383*		39.987		28.003
		(3.801)		(5.455)		(5.432)		(19.392)		(26.222)		(28.617)
Household size		-106.414***		-93.408***		-115.547***		1697.388***		1826.229***		1605.272***
		(12.608)		(17.470)		(18.009)		(82.471)		(104.957)		(115.254)
Child Age		25.871***		31.323**		20.417		105.580***		136.460**		69.748
		(8.693)		(12.537)		(12.667)		(40.086)		(57.050)		(57.174)
Second child age		13.656*		10.555		12.326		64.457		34.481		72.442
		(8.171)		(11.440)		(11.781)		(40.292)		(55.355)		(57.707)
Household head age		0.988		1.487		0.664		9.755		6.669		13.579
		(2.669)		(3.788)		(3.685)		(14.050)		(18.000)		(20.513)
Household head school year		8.851*		7.202		9.399		44.325*		30.754		52.647
		(4.775)		(6.612)		(6.589)		(23.870)		(33.247)		(32.098)
Mean depvar	2,226.47	2,226.47	2,259.19	2,259.19	2,197.99	2,197.99	10,652.72	10,652.72	10,587.00	10,587.00	10,709.92	10,709.92
District FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	1,747	1,741	813	810	934	929	1,747	1,741	813	810	934	929
R-Squared	0.003	0.243	0.001	0.303	0.016	0.265	0.000	0.508	0.000	0.581	0.000	0.495

Notes: Columns 1-6 of this table report the impact of second-born's gender on the per capita calorie consumption in the household. Columns 7-12 shows the impact on household's total calorie consumption. "All" denotes sample for all first-born children, irrespective of gender, while Male and Female segregate the sample into first-born male and first-born female children, respectively. All child child is between 5-16 years old. The second child's age range is also 5-16. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

primary objective of the FTF initiative was to increase household food security and improve access to food in the areas where it was implemented.



(a) Change in HDDS before and after FTF

(b) Change in HDDS before and after FTF, based on second-born's gender

Figure 1.5: Change in household dietary diversity score Before and After FTF

Notes: This figure shows household dietary diversity score (HDDS), measured on the scale of 0-12 before and after FTF program, among program participants and non participants. Panel (a) shows the change in HDDS before and after the of the program execution. Panel (b) shows how FTF program participant's and control households' HDDS have changed with the presence of a second-born male and second-born female child in the family.

We first examine the changes in food consumption patterns from the period prior to the program's initiation in 2011 to the period following its implementation in 2018. Panel (a) in Figure 1.5 illustrates the changes in the Household Dietary Diversity Score (HDDS) between the two periods. Before the start of the FTF program, the average HDDS was 6.86 in the control group and 6.69 in the FTF-recipient (treated) group. The FTF region had experienced several household and market-level interventions since 2013 as part of the initiative. After the FTF intervention, average HDDS of the FTF household climbed to 7.40 in 2018, while the HDDS of the households in the control group increased to 7.55. By including a second-born female child as an additional factor as indicator variable of son preference in our analysis, we explore how it affects the household's food consumption practice. The second panel 1.5b of Figure 1.5 shows the comparison of dietary diversity scores between FTF and control households, with and without a second-born female. This comparison highlights the subtle impacts that family structure has on food preferences.

Panel (b) of Figure 1.5 shows a notable increase in HDDS for FTF households with a second-born male child. The score increased from 6.61 in the pre-intervention period to 7.44 in the post-intervention period. This sharp increase demonstrates the significant impact of the FTF program on enhancing food diversity in the households with a second-born male child. The FTF households with second-born females experienced a relatively flatter rise in their dietary diversity score after the intervention, from 6.75 to 7.32. The program effect on the HDDS of households with a second-born girl can be calculated by $[(7.32-6.75)-(7.63-6.88)]-[(7.44-6.61)-(7.50-6.83)] = -.34$, a decrease in HDDS.

Table 1.5: Gender specific household dietary outcomes of FTF

Dep. var:	Household Dietary Diversity Score (HDDS)					
	DiD			Triple Diff		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-0.171*			-0.219		
	(0.0946)			(0.133)		
Post	0.695***			0.671***		
	(0.0593)			(0.0687)		
Treat x post	0.0146	0.0146	0.0425	0.157	0.238	0.252
	(0.130)	(0.130)	(0.129)	(0.182)	(0.200)	(0.198)
Treat x post x 2nd born female				-0.347	-0.526*	-0.515*
				(0.214)	(0.269)	(0.268)
Post x 2nd born female				0.0813	-0.0134	-0.0510
				(0.0950)	(0.0938)	(0.0944)
Household Level Controls	No	No	Yes	No	No	Yes
Region FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Observations	3,920	3,920	3,920	3,920	3,920	3,920
Mean of Dep. Variable	7.18	7.18	7.18	7.18	7.18	7.18

Notes: The results displayed in columns (1-3) of the table are Difference-in-Differences (DiD) regression estimates, while columns (4-6) showcase the outputs of the triple difference (DDD) regression. The outcome is household dietary diversity score (HDDS). Treat is binary variable indicator of FTF households. Post is a time dummy that indicates before or after the the program execution. Treat x post gives DiD estimator, and Treat x post x 2nd born female gives the DDD estimator. Clustered standard errors in parentheses; Clustered around upazilas (sub-district). Household FE and Time FE implies household and survey wave fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We conduct this analysis under the Difference in Difference (DiD) and Difference in Difference in Difference (DDD) frameworks. Models 1-3 in Table 1.5 show DiD regression estimates, whereas models 4-6 show estimates of DDD regression. In these models, the coefficient of the interaction term 'Treat x post' provides DiD estimator, while the coefficient of the triple interaction term 'Treat x post x 2nd born female' gives the DDD estimator. We used three different variations of DiD and DDD models. We did not include any fixed effect and household-level control variables in the first model. In the second, we included the year and household fixed effect, and in the third, we added household-level controls such as household size, total owned land, and religion. All standard errors are clustered at upazila (sub-district) level, as FTF program was implemented at upazila level.

The DiD regression analysis assumes a parallel trend in the control and treatment groups; that is, in the absence of treatment, over time the trajectories of the control and treatment groups would show identical patterns. The scope of testing pararallel trend assumption within our data is constrained by the presence of only one data point from the pre-intervention period (baseline) before the initiation of the Feed the Future (FTF) program. To overcome this limitation of directly assessing parallel patterns, we used an alternative approach.[10] Following the approach, we utilised a supplementary data set, namely the Bangladesh Demographic and Health Survey (DHS) to check for pre-trend in the treated regions.[11] By matching the latitude and longitude of both (DHS and BIHS) data sets, we identified the treatment and control regions at the upazila (sub-district) level.[12] Then we used the household wealth index of DHS survey as a proxy for household food consumption to show the trend before FTF intervention. We employed the event study design, which shows that there are no significant differences in wealth index between the treated and control groups in the periods prior to the execution of the FTF. As wealth index and consumption are assumed to be closely correlated, consistent wealth index in pre-intervention period suggests consistent pre-treatment food consumption pattern. The Figure 1.C.4 and Table 1.C.9 in Appendix show the results of this event study analysis to validate the parallel trend assumption. In addition showing there was no pre-trend, we also conducted a difference in difference regression on the household's predetermined covariates to establish that households are not selected, and household characteristics did not change over time due to FTF, as shown in Table 1.C.10. Here we show that there was no significant change in Household heads age, education, religion, ethnicity due to FTF program.

[10] Suggested by David Mckenzie, Lead Economist, Development Research Group at World Bank: `https://blogs.worldbank.org/impactevaluations/what-do-about-parallel-trends-when-you-only-have-baseline-data`.

[11] DHS survey is nationally representative survey conducted in regular interval.

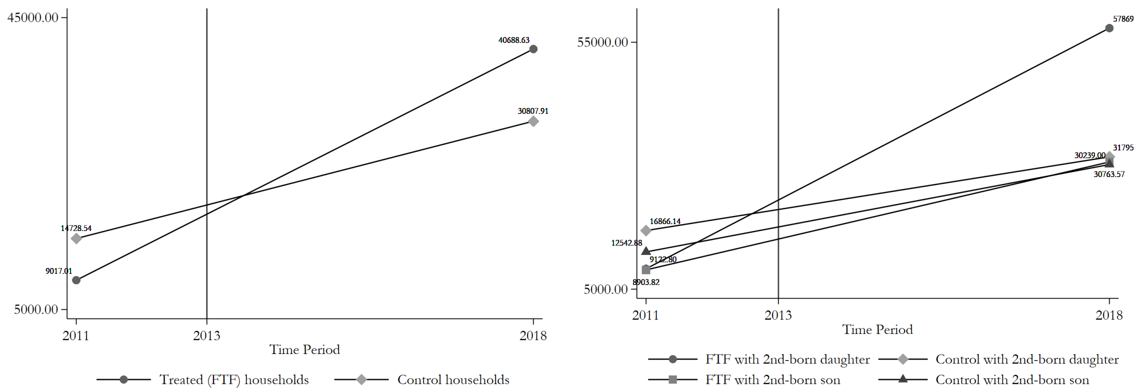
[12] FTF program was executed at upazila (sub-district) level.

The result of the DiD regressions (columns 1-3) in Table 1.5 shows that there is no significant impact of the FTF program on the household dietary diversity scores (hdds) between the treatment households. However, the DDD regression estimates (4-6) show that the presence of a second born daughter in the treatment household reduces the household dietary diversity score by 0.53 ($p<0.10$). HDDS measures the extent to which a household consumes a variety of food groups, ranging from 0 to 12, and a 0.52 decrease in the score indicates a substantial reduction in the variety of food groups consumed.

To determine whether the reduction in the household dietary diversity score originates from son preference induced consumption choices, not from lack of resources or savings, we analyze how total household savings fluctuate under both treatment conditions before and after FTF execution. Figure 1.6 shows that there is a positive change in savings due to program intervention. This pattern may indicate that households are either compiling additional savings specifically for the daughter's future, such as expenses related to marriage, or are bolstering their overall savings as a result of a perceived necessity for increased financial insecurity in the absence of a son. Although Table 1.B.2 indicates that the observed savings changes are not statistically significant, the trend still suggests that the decrease in the dietary diversity score is more likely influenced by son preference, rather than the limitations of the financial resources of the household. The differential outcome of large programs also supports the assumption of son preference, revealing a discernible bias against female children in rural Bangladesh. This bias is evident in household dietary decisions, a pattern that continues even in situations where households have experienced an improvement in their socioeconomic status due to external intervention programs such as FTF. This suggests that the influence of son preference persists in shaping household dietary decisions and the allocation of resources, notwithstanding general improvements in financial prosperity.

1.6 Conclusion

Our study causally demonstrates that in societies such as rural Bangladesh, with a predominant preference for male children due to socioeconomic factors, the presence of a female child can lead to inequalities in family priorities and allocation of resources. The results show that the nutritional inputs of a first-born child may vary based on the gender of the younger sibling. First-born children enjoy more nutritional care



(a) DiD (b) DDD

Figure 1.6: Change in household savings before and after FTF

Notes: Panel (a) shows the change in savings before and after the of the program execution. Panel (b) shows how FTF program participant's and control households' savings have changed in the presence of a second-born male and second-born female child in the family.

when the second-born is a female, which reflects through first-born's better health and anthropometric circumstances. We also reveal that if first-born and second-born children are both female, the overall per capita calorie consumption of the family substantially reduces. In this instance, first-born children get more share in foods; their health concerns are presumably managed better, and as a result, their height-for-age z-score increases. As mechanisms of these phenomena, we found empirical evidence of son preference as household size increases substantially when there is a second-born female in a family. It indicates that parents tend to have more children in expectation of a male child when second-born child is female.

Overall, the direction of the changes in the HFA z-scores, calorie intakes and household dietary diversity scores found in our study demonstrate the intricate relationship between social norms, gender dynamics, and family financial strategies, all of which seems to have a detrimental impact on the nutritional well-being of female children. The main strength of our paper is the use of multiple identification strategies to capture the effect of son preference on first-born children and households. Besides utilizing natural variation, such as randomness of second child's gender on the health outcome of first-born child, we also used other quasi-experimental methods to evaluate large-scale program's outcome on households with different sibling gender compositions. We also conducted a series of balance and robustnes checks for each part of our analysis. The major shortcoming of this study comes from the sample size, especially when we split it by gender and restrict first-born and second-born children to ages 5-16. This small sample size might end up estimating a less precise effect size. Moreover, the sample is representative of rural Bangladesh, from where it is hard to claim that the results are representative of all Bangladesh. So, based on our study, extensive research with larger sample size is necessary to have a more robust conclusion about the impact of son preference on children and households.

We distinguish how families with different sibling gender compositions diversify their food when households have better income and nutritional knowledge gained through intervention programs such as Feed the Future. We observed a significant reduction in the variety of foods consumed when families with a second-born daughter were part of a program aimed at improving financial barriers and nutritional condition. The results imply that large-scale external intervention programs, such as FTF, do not work in an identical way for all families. We report that the results of development initiatives can vary depending on the gender composition of target households. Our findings open up a new window for future research regarding intervention programs by taking consideration of son preference in society. We also anticipate that this study will make a valuable contribution towards the prevention of gender discrimination within households motivated by son preference.

In summary, the nutritional adequacy of Bangladeshi schoolchildren is closely linked with the gender of their siblings and the parental perception of the gender of the child. Our research reveals gender-based nutritional disparities among siblings, with daughters at a disadvantage. Stakeholders should prioritize awareness campaigns to address the neglect of female children's nutritional needs.

Appendix

1.A Background

1.A.1 Change in sex ratio in Bangladesh

Bangladesh is a least developed country (LDC) with a population of 164.6 million. It is the eighth most populous country worldwide, where 1,094 people live per square kilometre. With a large population and a relatively high poverty rate, reducing the deep-rooted gender disparity is one of the biggest development challenges for Bangladesh. Recent statistics indicate that Bangladesh has considerably improved in lowering its fertility rate from 6.3 childbirths per woman in 1976 to 2.1 children per woman in 2018. With continued efforts from the Bangladesh government and international development partners, gender differences or sex ratio at birth have also been reduced in Bangladesh.[13]

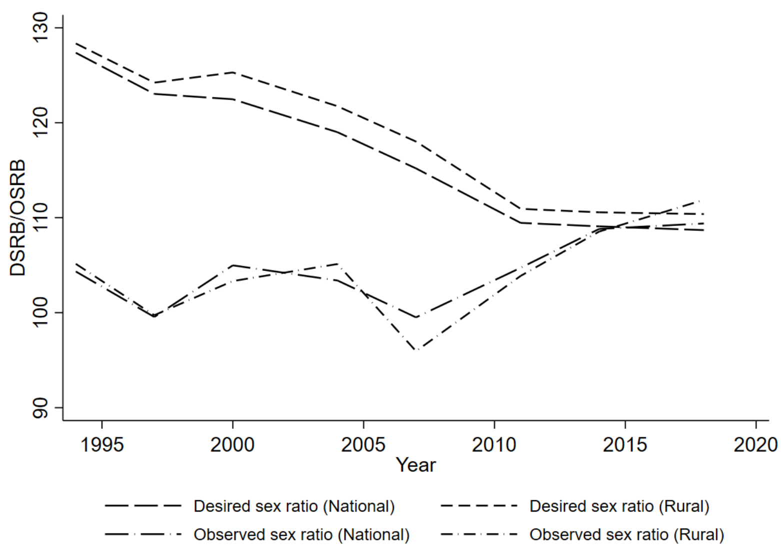


Figure 1.A.1: Observed and desired sex ratio in Bangladesh 1994-2018

Notes: This graph shows how desired and observed sex ratio in Bangladesh has changed at national level and rural level over the period 1994 to 2018. Source: Calculations are based on the Bangladesh Demographic and Health Surveys (1993-94 to 2017-18).

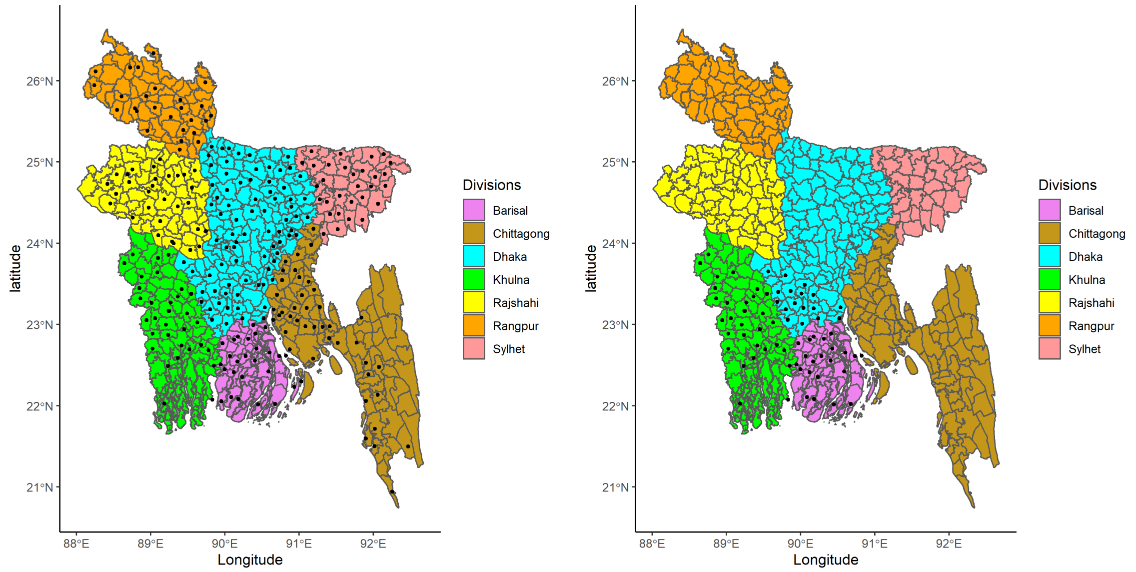
Using eight rounds of data from the Bangladesh Demographic and Health Survey from 1994 to 2018, we calculate the observed sex ratio and the desired sex ratio in

[13]The Sex ratio at birth (SRB) shows the number of male births per 100 female births in a period. The desired sex ratio at birth (DSRB) expresses the expected ratio of male to female children by parents. Meanwhile, the observed sex ratio (ORSB) is the actual male to female ratio at birth in a country.

Bangladesh which indicates how these two indicators have evolved in recent years. From the figure 1.A.1, it is clear that Bangladesh has improved rapidly in reducing (stated) preference for sons. In 1994, Bangladeshi parents wished for 127 male children for every 100 female children. In 2018, the desired sex ratio came down to around 108, indicating a balance in sex ratio with lower stated son preference. Interestingly, son preference is a stated preference by parents, while real statistics go in the opposite direction.

1.A.2 Feed the Future program detail

Feed the Future works worldwide in underdeveloped areas to improve the quality of life by engaging governments, private sectors, institutions, and civil society. FTF operates to achieve diversified economic growth, improved nutritional status, and reduced deprivation in the operating area. The FTF was implemented at both the household and policy levels. FTF projects focus on connecting agriculture and nutrition. FTF provided farmers and small enterprises with better market information, which helped them increase their income. In addition to improving the income of households, nutritious food was also promoted in the markets in the program area. High-yield, drought resistant and nutritious rice production was promoted to the government, NGOs, and farmers in the FTF zone. FTF program also encouraged high-nourishing other non-staple food items such as vegetables, fruits, livestock, and fish. Pregnant women and young children were especially subject to nutritional information delivery. Women were introduced through nutrition education and diverse homestead food production mechanisms. FTF initiative connected private sectors and government bodies to improve the market system and eliminate market constraints. FTF promoted updated farming practices in the program area by training farmers on fertilizer use, pesticide use, and proper irrigation practices. The achievement of gender equality was another goal of FTF. It targeted women in the family so that nutritional and agricultural knowledge is properly implemented. Behavioral and social change through community outreach was attempted for better nutritional practice in society. Feed the Future project puts extra effort into ensuring food security in the program region. The program was designed to protect dietary needs and reinforce a productive and healthy life in the program area. Before the implementation of the program, the FTF zone was surveyed. Figure 1.A.2 depicts the survey regions (FTF and national) in Bangladesh map.



(a) Survey Upazila (sub-districts): Country level (b) Survey Upazila (sub-districts): Feed the Future regions

Figure 1.A.2: Feed-the-future survey regions

1.B Data and Sample

1.B.1 Dataset

Bangladesh Integrated Household Survey is a combined effort from USAID, the International Food Policy Research Institute (IFPRI), and the Bangladesh government. USAID funded a project named *The Bangladesh Policy Research and Strategy Support Program (PRSSP)* in 2010 to understand and improve agricultural development and food security concerns in Bangladesh. It aimed to enhance overall analytical capacity along with policy assistance to government and development partners. This IFPRI-USAID initiative planned to conduct specially designed surveys at the individual, household, and community level, which would provide detailed information on food security and agricultural development situations. The sample for BIHS-2011 was collected from 325 primary sampling units, which were mainly villages. BIHS used two-stage stratified sampling with units being selected in the first stage, while in the second stage, households were selected from those primary sampling units based on the sampling frame of Bangladesh Population Census 2001. It also adjusted the sampling weight using Bangladesh Population Census 2011. Eight strata (Feed the Future Zone and seven administrative divisions) from where the primary sampling units were selected. Initially, eighty-seven primary sampling units were from the Dhaka division, 48 primary sampling units were from the Chittagong division, 36 primary sampling units were from Sylhet division, 29 primary sampling units were from the Rajshahi division, 27 primary sampling units were from Khulna division, 27 primary sampling units were from Rangpur division, 21 primary sampling units were from Barisal division and 50 PSUs were from FTF zone. A random selection of 20 households was made from each primary sampling unit. Later, it was noticed that the sample size selected primarily of 1,000 households in the FTF zone was not enough for robust program evaluation in the future, and then 52 more primary sampling units from FTF divisions (Dhaka, Barishal, and Khulna) were added. So, 52 more primary sampling units have been added with additional 1040 FTF households. After finalizing the sample size, the IFPRI team prepared a detailed questionnaire, which was peer-reviewed and validated by researchers from USAID and Bangladesh Government. The questionnaire had two parts, one part was for female respondents, and another part was for male respondents. A wide range of information was collected, ranging from household composition, income, assets to individual health, nutrition, and women empowerment status. Bangladesh Integrated Household Survey contains information on (i) individual and household income, expenditure, and agricultural production, (ii) Height and weight measurement for all the household members (iii) Household's dietary intake with names of food items and quantity consumed. (iv) Illness and health data of household members and Mother's nutritional knowledge and practice of micronutrient use in the household. In the first wave, the data collection process was started on October 2011, and the whole survey procedure was completed by March 2012.

1.B.2 Variable description

Table 1.B.1: Description of variables

Variable	Description
HFA z-score	The height-for-age z-score is a quantitative measure that indicates the deviation of a child's observed height from the median height of a reference group of children of the same age, expressed in terms of standard deviations. A positive HFA z-score indicates a height above the median, while a negative z-score indicates a height below the median. This score is an indicator of overall growth and nutritional status of children.
Stunting	The incidence of stunting is represented as a dummy variable derived from the height-for-age z-score. A child with a HFA z-score below -2 is considered stunted (assigned a value of 1, otherwise 0). A score below -3 indicates severe stunted.
Calorie-share (ind_share)	This variable represents the calorie share of the first-born child within the household. It is calculated as the proportion of the first-born child's calorie intake to the total household calorie consumption. This metric is useful for understanding the child's nutritional intake in the context of the household's overall calorie distribution, highlighting potential nutritional disparities within the family unit.
HDDS	Household Dietary Diversity Score (HDDS) capture the variation of foods that households have consumed. Using the available data on dietary intake of the households in the BIHS survey, we calculated HDDS employing the guideline of the Food and Agriculture Organization (FAO). Kennedy et al. (2011) provides detailed process and instruction for calculating HDDS and claims that HDDS delivers a holistic picture for nutrition and food security. High Dietary diversity scores reflect better nutritional adequacy and household food availability based on the variety of the last 24 hours of food intake. To calculate HDDS, all consumed food items by a household are brought under following 12 groups- x_1= Cereals; x_2= Tuber and Roots; x_3= Vegetables; x_4= Fruits; x_5= Meat, poultry; x_7= Fish and Seafood, x_8= Pulses, Nuts, Legumes; x_9= Milk and Milk products; x_{10}= Oils and fats; x_{11}= Sugar/honey; x_{12}= Miscellaneous. For consumption of any food item under one group, households receive a score of 1; otherwise, it gets 0. HDDS= $\sum(\mathrm{x}_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12})$ Finally, Household Dietary Diversity Score is calculated by adding up the number of food groups consumed, and it varies from 0 to 12.

1.B.3 Additional results

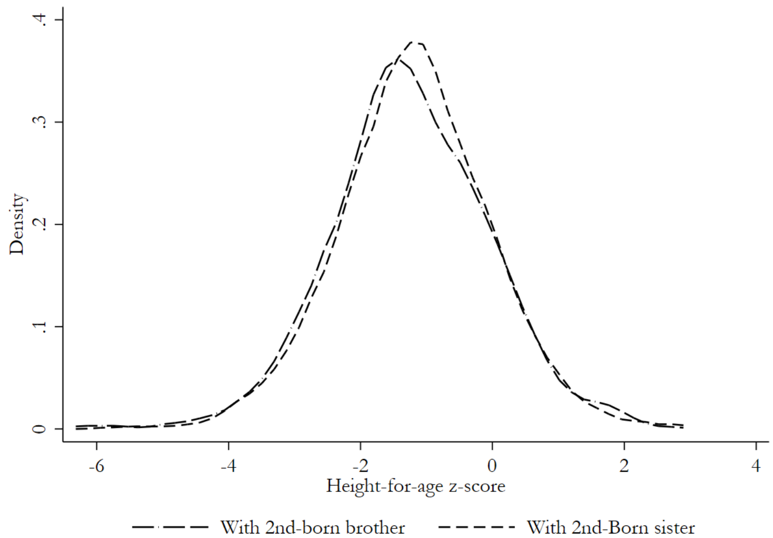


Figure 1.B.3: Density of HFA z-score by gender of second-born

Notes: This kernel density graph compares the HAZ score of first born children with a second-born brother and second-born sister. The figure indicates that HAZ score of first born is distributed slightly to the right side below zero if first born has a younger sister.

Table 1.B.2: Change in savings of due to FTF

Dep. var:	Household total yearly savings					
	DiD			Triple Diff		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-4678.6**			-3608.2		
	(2033.9)			(2618.9)		
Post	16726.0***			17210.6***		
	(2240.4)			(3189.4)		
Treat x post	17081.5	17081.5	18950.3	4484.3	3717.5	4790.9
	(12620.4)	(12618.8)	(13532.9)	(6562.0)	(9160.2)	(9146.6)
Treat x post x 2nd born female				33878.6	36824.2	38195.0
				(33712.8)	(39614.0)	(40028.8)
Post x 2nd born female				-370.9	1740.9	-1140.0
				(4580.3)	(4846.3)	(5013.3)
Household Level Controls	No	No	Yes	No	No	Yes
Region FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Observations	3,920	3,920	3,854	3,920	3,920	3,920
Mean of Dep. Variable	22,954.51	22,954.51	23,153.04	22,954.51	22,954.51	22,954.51

Notes: The results displayed in columns (1-3) of the table are Difference-in-Differences (DiD) regression estimates, while columns (4-6) show the outputs of the triple difference (DDD) regression. The outcome is household savings. Treat x post gives DiD estimator, and Treat x post x 2nd born female gives the DDD estimator. Clustered standard errors in parentheses; Clustered around upazilas (sub-district). Household FE and Time FE implies household and survey wave fixed effect. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.B.3: Change in per capita monthly food expenditure

Dep. var:	Household montlhy food exp					
	DiD			Triple Diff		
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	-53.86			-89.33		
	(53.96)			(73.84)		
Post	912.6***			1008.6***		
	(32.72)			(46.01)		
Treat x post	56.91	56.91	19.48	136.6	112.8	120.0
	(69.24)	(69.23)	(69.88)	(109.2)	(120.1)	(119.8)
Treat x post x 2nd born female				-200.8	-133.3	-145.2
				(155.5)	(150.9)	(151.7)
Post x 2nd born female				-265.9***	-266.7***	-260.4***
				(61.42)	(55.21)	(56.00)
Controls	No	No	Yes	No	No	Yes
Region FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
Observations	3,920	3,920	3,854	3,920	3,920	3,854
Mean of Dep. Variable	1,853.97	1,853.97	1,851.25	1,853.97	1,853.97	1,851.25

Notes: The results displayed in columns (1-3) of the table are Difference-in-Differences (DiD) regression analysis, while columns (4-6) show the outputs of the triple difference (DDD) regression. Outcome is Model Clustered standard errors in parentheses; Clustered around upazilas (sub-district). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.B.4 Quantile regressions

Table 1.B.4: HFA z-score quantiles: Impact of female second child

Dep. var:	Height-for-Age Z-Score							
Quantile	Q05 (1)	Q15 (2)	Q25 (3)	Q35 (4)	Q45 (5)	Q50 (6)	Q75 (7)	Q95 (8)
Second child is female	0.244* (0.135)	0.110 (0.101)	0.216*** (0.076)	0.176** (0.074)	0.132* (0.071)	0.098 (0.068)	0.049 (0.078)	-0.041 (0.100)
Mother's School Year	-0.008 (0.031)	0.022 (0.017)	0.018 (0.012)	0.017 (0.011)	0.014 (0.011)	0.010 (0.012)	0.008 (0.013)	0.016 (0.023)
Mother Age	-0.026* (0.016)	0.002 (0.017)	0.003 (0.014)	-0.001 (0.013)	-0.008 (0.010)	-0.006 (0.010)	0.008 (0.012)	0.012 (0.015)
Household size	-0.079** (0.038)	-0.050 (0.049)	-0.040 (0.029)	-0.066** (0.027)	-0.085*** (0.028)	-0.087*** (0.027)	-0.091** (0.035)	-0.107*** (0.037)
Child Age	-0.034 (0.061)	-0.042 (0.040)	-0.051* (0.028)	-0.062*** (0.021)	-0.083*** (0.021)	-0.085*** (0.020)	-0.154*** (0.023)	-0.250*** (0.032)
Child gender	-0.134 (0.184)	-0.019 (0.109)	0.171** (0.077)	0.150** (0.069)	0.113* (0.065)	0.182*** (0.068)	0.280*** (0.073)	0.412*** (0.094)
Second child age	0.051 (0.049)	0.032 (0.032)	0.017 (0.022)	0.020 (0.022)	0.041** (0.020)	0.039** (0.019)	0.051** (0.023)	0.068*** (0.022)
Household head age	0.011 (0.011)	0.000 (0.011)	0.002 (0.009)	0.001 (0.009)	0.009 (0.008)	0.011 (0.007)	0.001 (0.008)	-0.004 (0.010)
Household head school year	0.033 (0.025)	0.005 (0.018)	0.002 (0.012)	-0.007 (0.011)	-0.007 (0.011)	-0.002 (0.012)	0.008 (0.012)	-0.003 (0.019)
Division FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1599	1599	1599	1599	1599	1599	1599	1599
R-Squared	0.031	0.045	0.056	0.062	0.066	0.067	0.062	0.055

Notes: This table presents quantile regression results for the effect of having a second-born female child on the height-for-age z-score (HAZ) of the first-born child. The analysis uses data from households with at least two children, aged between 5 and 16 years old. The key independent variable indicating whether the second-born child is female.Administrative division fixed effects are included. Quantile regressions are estimated at the 5th, 15th, 25th, 35th, 45th, 50th, 75th, and 95th percentiles of the haz distribution. Standard errors are clustered at the upazila level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.B.5: Calorie share quantiles in household consumption: Impact of female second child

Dep. var:	Calorie share							
Quantile	Q05 (1)	Q15 (2)	Q25 (3)	Q35 (4)	Q45 (5)	Q50 (6)	Q75 (7)	Q95 (8)
Second child is female	0.009* (0.005)	0.008** (0.003)	0.005 (0.003)	0.005* (0.003)	0.005* (0.003)	0.004 (0.003)	0.002 (0.004)	0.004 (0.008)
Mother's School Year	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)	-0.000 (0.001)
Mother Age	0.000 (0.001)	-0.000 (0.000)	0.000 (0.001)	0.000 (0.000)	0.001 (0.000)	0.000 (0.000)	0.001** (0.001)	0.003*** (0.001)
Household size	-0.014*** (0.002)	-0.018*** (0.002)	-0.019*** (0.002)	-0.021*** (0.002)	-0.023*** (0.001)	-0.024*** (0.001)	-0.027*** (0.002)	-0.031*** (0.003)
Child Age	0.001 (0.002)	0.004*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.008** (0.003)
Child gender	0.014*** (0.005)	0.003 (0.003)	0.002 (0.003)	0.003 (0.003)	0.007** (0.003)	0.007** (0.003)	0.015*** (0.004)	0.021*** (0.007)
Second child age	0.002 (0.002)	0.000 (0.001)	-0.002* (0.001)	-0.002 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.003 (0.003)
Household head age	-0.000 (0.001)	-0.000 (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)	-0.002*** (0.000)	-0.004*** (0.001)
Household head school year	-0.001 (0.001)	-0.001 (0.001)	-0.001* (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.000 (0.001)
Division FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1687	1687	1687	1687	1687	1687	1687	1687
R-Squared	0.225	0.247	0.256	0.265	0.269	0.268	0.269	0.244

Notes: This table presents quantile regression results for the effect of having a second-born female child on the calorie share of the first-born child. The analysis uses data from households with at least two children, aged between 5 and 16 years old. The key independent variable indicating whether the second-born child is female.Administrative division fixed effects are included. Quantile regressions are estimated at the 5th, 15th, 25th, 35th, 45th, 50th, 75th, and 95th percentiles of the haz distribution. Standard errors are clustered at the upazila level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.B.6: Daily calorie intake quantiles of first-born: Impact of female second child

Dep. var:	Calorie intake							
Quantile	Q05 (1)	Q15 (2)	Q25 (3)	Q35 (4)	Q45 (5)	Q50 (6)	Q75 (7)	Q95
Second child is female	122.078** (60.465)	9.402 (48.815)	4.875 (27.703)	48.521 (29.486)	37.544 (29.955)	17.216 (39.001)	3.094 (59.890)	
Mother's School Year	-11.413 (14.016)	-1.888 (9.339)	-0.470 (5.538)	3.310 (5.935)	6.241 (7.020)	6.426 (7.222)	-16.601* (8.650)	
Mother Age	-6.699 (7.559)	3.292 (6.486)	2.158 (4.959)	2.541 (3.577)	5.748 (3.960)	12.939** (5.131)	13.077 (8.952)	
Household size	12.326 (35.805)	17.470 (26.283)	7.172 (12.399)	-2.597 (11.604)	4.645 (12.054)	37.817* (22.093)	2.618 (24.175)	
Child Age	59.300*** (21.378)	55.069*** (13.668)	73.442*** (8.483)	74.949*** (9.364)	80.583*** (10.565)	87.261*** (16.155)	85.186*** (16.858)	
Child gender	95.023* (57.373)	134.915*** (45.867)	122.769*** (27.369)	116.126*** (28.178)	131.692*** (29.398)	174.885*** (40.906)	212.014*** (68.324)	
Second child age	7.863 (22.104)	9.575 (12.276)	8.477 (9.464)	13.587 (9.589)	7.486 (10.558)	5.488 (16.538)	11.938 (17.013)	
Household head age	1.216 (4.963)	-0.065 (3.299)	-1.220 (3.389)	-1.650 (2.564)	-2.968 (2.685)	-7.953* (4.212)	-8.910* (5.083)	
Household head school year	8.166 (11.142)	-0.013 (7.657)	0.425 (5.231)	-0.468 (5.597)	-1.128 (6.086)	-3.526 (6.380)	13.368 (8.139)	
Division FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
N	1687	1687	1687	1687	1687	1687	1687	
R-Squared	0.138	0.169	0.183	0.183	0.185	0.181	0.155	

Notes: This table presents quantile regression results for the effect of having a second-born female child on the daily calorie intake of the first-born child. The analysis uses data from households with at least two children, aged between 5 and 16 years old. The key independent variable indicating whether the second-born child is female.Administrative division fixed effects are included. Quantile regressions are estimated at the 5th, 15th, 25th, 35th, 45th, 50th, 75th, and 95th percentiles of the haz distribution. Standard errors are clustered at the upazila level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.C Robustness

1.C.1 Regressions with clustered standard error

Table 1.C.7: Effect on first child's HFA z-score and incidence of stunting: with clustered se at the upazila level

	HFA Z-Score						Stunting					
	All		Male		Female		All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Second child is female	0.062	0.098*	0.061	0.083	0.061	0.099	-0.026	-0.031	-0.026	-0.026	-0.026	-0.031
	(0.059)	(0.058)	(0.082)	(0.083)	(0.079)	(0.079)	(0.021)	(0.022)	(0.029)	(0.029)	(0.030)	(0.032)
Mother's School Year		0.009		0.012		0.009		-0.002		-0.005		-0.001
		(0.011)		(0.015)		(0.015)		(0.004)		(0.005)		(0.006)
Mother Age		-0.001		-0.001		0.000		-0.000		0.001		-0.002
		(0.008)		(0.014)		(0.011)		(0.003)		(0.004)		(0.004)
Household size		-0.086***		-0.048		-0.117***		0.016*		0.003		0.027**
		(0.024)		(0.040)		(0.028)		(0.009)		(0.015)		(0.013)
Child Age		-0.110***		-0.089***		-0.137***		0.008		0.012		0.007
		(0.019)		(0.027)		(0.025)		(0.007)		(0.010)		(0.009)
Second child age		0.046***		0.072***		0.025		-0.004		-0.015		0.004
		(0.017)		(0.026)		(0.021)		(0.006)		(0.009)		(0.009)
Household head age		0.005		0.008		0.005		-0.000		-0.001		0.001
		(0.006)		(0.010)		(0.008)		(0.002)		(0.003)		(0.003)
Household head school year		0.006		0.007		0.006		-0.000		0.001		0.001
		(0.011)		(0.016)		(0.014)		(0.004)		(0.005)		(0.005)
Mean depvar	-1.21	-1.21	-1.13	-1.13	-1.28	-1.28	0.23	0.23	0.22	0.22	0.24	0.24
Division FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	1605	1599	734	732	871	867	1605	1599	734	732	871	867
R-Squared	0.001	0.075	0.001	0.071	0.001	0.109	0.001	0.033	0.001	0.056	0.001	0.043

Notes: Columns 1-6 of this table report the impact of second-born's gender on the height-for-age Z-score of the first-born child while columns 7-12 shows the impact on incidence of stunting. Stunting is a binary variable that indicates if HFA z-score falls below minus two standard deviations from the median of the World Health Organization (WHO) Child Growth Standards. "All" denotes sample for all first-born children, irrespective of gender, while Male and Female segregate the sample into first-born male and first-born female children, respectively. All child child is between 5-16 years old. Second child age range is also 5-16. Standard errors are clustered at upazila level. Controlled for seven administrative division fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.C.8: Effect on first child's HFA z-score and incidence of stunting: clustered se at the village level

	HFA Z-Score						Stunting					
	All		Male		Female		All		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Second child is female	0.062	0.099*	0.061	0.087	0.061	0.097	-0.026	-0.032	-0.026	-0.030	-0.026	-0.030
	(0.058)	(0.057)	(0.084)	(0.084)	(0.076)	(0.075)	(0.021)	(0.021)	(0.030)	(0.030)	(0.030)	(0.031)
Mother's School Year		0.010		0.012		0.008		-0.002		-0.005		-0.000
		(0.011)		(0.015)		(0.014)		(0.004)		(0.005)		(0.006)
Mother Age		-0.000		-0.000		0.001		-0.001		0.001		-0.002
		(0.008)		(0.013)		(0.011)		(0.003)		(0.004)		(0.004)
Household size		-0.086***		-0.048		-0.112***		0.016*		0.004		0.025*
		(0.024)		(0.039)		(0.029)		(0.010)		(0.015)		(0.013)
Child Age		-0.109***		-0.090***		-0.135***		0.008		0.012		0.007
		(0.019)		(0.027)		(0.025)		(0.007)		(0.010)		(0.009)
Second child age		0.047***		0.073***		0.026		-0.004		-0.015*		0.004
		(0.016)		(0.026)		(0.021)		(0.007)		(0.009)		(0.009)
Household head age		0.005		0.007		0.004		0.000		-0.001		0.001
		(0.006)		(0.009)		(0.008)		(0.002)		(0.003)		(0.003)
Household head school year		0.006		0.007		0.007		0.000		0.001		-0.000
		(0.011)		(0.015)		(0.014)		(0.004)		(0.005)		(0.005)
Mean depvar	-1.21	-1.21	-1.13	-1.13	-1.28	-1.28	0.23	0.23	0.22	0.22	0.24	0.24
Division FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	1605	1599	734	732	871	867	1605	1599	734	732	871	867
R-Squared	0.001	0.072	0.001	0.069	0.001	0.101	0.001	0.028	0.001	0.048	0.001	0.033

Notes: Columns 1-6 of this table report the impact of second-born's gender on the height-for-age z-score of the first-born child while columns 7-12 shows the impact on incidence of stunting. Stunting is a binary variable that indicates if HFA z-score falls below minus two standard deviations from the median of the World Health Organization (WHO) Child Growth Standards. "All" denotes sample for all first-born children, irrespective of gender, while Male and Female segregate the sample into first-born male and first-born female children, respectively. All child child is between 5-16 years old. Second child age range is also 5-16. Standard errors are clustered at village level. Controlled for seven administrative division fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.C.2 Event study

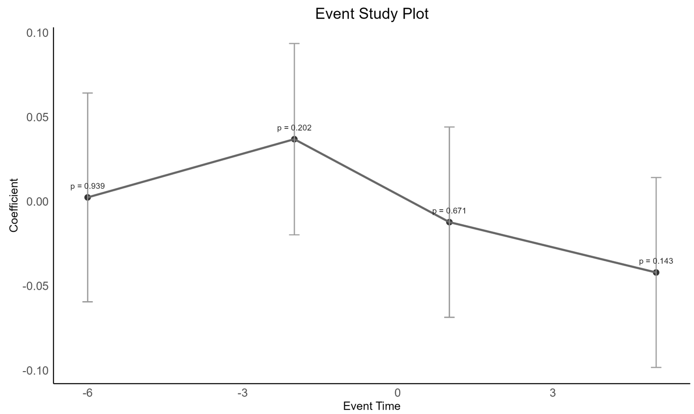


Figure 1.C.4: Event study plot: trends in wealth index

Notes: This figure presents an event study plot examining the trend in household wealth index before the Food for the Future (FTF) intervention. The wealth index from the DHS survey is used as a proxy for household food consumption. The x-axis represents event time, with 0 indicating the start of the FTF intervention. The y-axis shows the coefficient estimates for the interaction between treatment status and time relative to the intervention. Error bars represent 95% confidence intervals. P-values for each estimate are provided above the corresponding point. The plot demonstrates no statistically significant differences in wealth index between treated and control groups in the pre-intervention periods six years and three years from the intervention started (-6, -3), supporting the parallel trends assumption.

Table 1.C.9: Event study regression

	Dependent variable:	
	Wealth Index	
	(1)	(2)
Treatment x Event Time (Factor) -6	0.002	0.005
	(0.032)	(0.031)
Treatment x Event Time (Factor) -2	0.037	0.042
	(0.029)	(0.029)
Treatment x Event Time (Factor) 1	−0.012	−0.012
	(0.029)	(0.029)
Treatment x Event Time (Factor) 5	−0.042	−0.042
	(0.029)	(0.028)
Model Controls	No	Yes
Two-Way Fixed Effects	Yes	Yes
Observations	46,748	46,748
R^2	0.0002	0.017

Notes: This table presents results from an event study analysis testing for pre-trends in the Demographic and Health Survey (DHS) data. The dependent variable is the Wealth Index. Coefficients shown are for the interaction between treatment status and event time Column (1) shows results without additional controls, while Column (2) includes model controls. Both models include two-way fixed effects. Standard errors are reported in parentheses. No statistically significant differences are observed in pre-treatment periods, supporting the parallel trends assumption for difference-in-differences analysis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

1.C.3 FTF impact on predetermined covariates

Table 1.C.10: Impact of FTF predetermined covariate

Dep. var:	Religion	Language	HH School	Marital	Mother Age	Mother Edu
	(1)	(2)	(3)	(4)	(5)	(6)
Treat x post	0.001	-0.001	-0.047	0.001	-0.100	-0.085
	(.0276)	(.0017)	(.3481)	(.0012)	(.5865)	(.3061)
Treat	-0.037*	0.003**	0.679***	0.001	1.280***	0.409*
	(.0195)	(.0014)	(.2451)	(6.2e-04)	(.41)	(.216)
Post	-0.001	0.001	0.100	-0.001	6.269***	0.069
	(.0101)	(.0017)	(.1333)	(.0012)	(.2367)	(.1233)
Constant	0.909***	0.997***	3.152***	0.999***	31.836***	3.526***
	(.0071)	(.0014)	(.0946)	(6.2e-04)	(.1579)	(.0879)
Observations	3920	3920	3920	3920	3887	3885
R^2	0.002	0.001	0.004	0.001	0.180	0.002

Notes: This table presents difference-in-differences estimates for predetermined household characteristics. Each column represents a separate regression with the dependent variable indicated in the column header. The analysis demonstrates that the Feed for the Future (FTF) program did not significantly affect household's predetermined characteristics such as religion, language, household head's school attendance, marital status, mother's age, and mother's education. This Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 2

Marriage Market Induced Health Investments in Women: Evidence from a Regression Discontinuity Design in India*

Joint with: Daniel Wiesen † and Johannes Schmieden ‡

Abstract

We examine the changes in the various health related indicators among Indian women around the age of 18, which is their legally specified threshold age for marriage in India. Using data from Indian Demographic and Health Survey-2019/20 and utilizing the local randomization-based regression discontinuity design with a 3-month bandwidth on either side of this cut-off age, we found a significant jump both in body mass index (0.40 kg/m^2) and bmi-for-age Z-score (0.13 SD). The blood hemoglobin, blood glucose, and blood pressure levels also show discontinuity at this point. As a mechanism of this change, we found a significant shift in consumption patterns with a higher likelihood of having meat and fish on a daily meal after reaching age 18. The results indicate that families increase their investments in female health once they reach the legal age of marriage. Our findings underscore that the legal marriage age has a profound impact on women's health.

*We are grateful for valuable comments and suggestions from Abu S. Shonchoy, Jishnu Das, Lorenz Goette, Arijit Ghosh, Juliane Hennecke, Mohammad Jakaria, Ingo Isphording, Akib Khan, and Arndt Reichert. We also express our appreciation to the seminar and conference participants at the dggö annual conference in Hannover, iHEA in Cape Town, and the Workshop of the dggö Committee "Allocation and Distribution" in Oldenburg.

†Department of Healthcare Management and Center for Social and Economic Behavior (C-SEB) University of Cologne and Erasmus School of Health Policy and Management (ESHPM), Erasmus University of Rotterdam

‡University of Bonn

2.1 Introduction

Child marriage remains a widespread problem globally, with a notable prevalence in least developed and developing nations of Asia and Africa.[1] Various factors are deeply rooted in the practice of child marriage, such as poverty (Otoo-Oyortey and Pobi, 2003; Field and Ambrus, 2008), education (Duflo et al., 2015; Paul, 2019), conflict (Mourtada et al., 2017; Hunersen et al., 2021), cultural standards (Chowdhury, 2004; Birech, 2013), and climatic conditions (Alston et al., 2014; Asadullah et al., 2021a). It is well known that female children are the main victims of child marriage.[2] In impoverished families, the key motivations for parents to instigate a child marriage is to alleviate the perceived burden of an additional family member, especially when that member is not contributing to any income-generating activities. Consequently, there is a marked inclination for parents to hasten the marriage of their female offspring as soon as they reach reproductive maturity. Countries establish legal age restrictions to safeguard children from becoming victims of child marriage. In India, the legal minimum age for marriage is 18 years for women.[3] In this paper, we assume that individuals enter the marriage market once they reach this cut-off age set by the government.[4] Using a nationally representative sample of Indian women aged around 18 years and utilizing their birth month information, we examine whether women experience marriage market-specific health accumulation, driven by the expectation of enhancing their prospects of securing a more desirable and higher quality marriage partner.

Numerous studies have reported that early marriage negatively impacts women's health and the children they give birth to (Field and Ambrus, 2008; Raj et al., 2009; Santhya et al., 2010; Sekhri and Debnath, 2014; Chari et al., 2017). Therefore, governments, international organizations, and development agencies are actively implementing measures to address the problem of child marriage. International laws play

[1] Child marriage is defined as any marriage or union involving children or between a child and an adult (UNICEF, 2005; Wodon et al., 2017).

[2] The disproportionate burden of child marriage mostly falls on girls, with 720 million girls out of 900 million yearly child marriage (Raj et al., 2018).

[3] According to Indian law, marriage below age 18 for women is not permissible and can result in imprisonment for up to two years and a fine of up to one Lakh for individuals involved in performing, conducting, directing, or assisting in child marriage.

[4] The marriage market is defined as a framework in which individuals engage in searching and choosing their match for marriage. Here, the objective of an individual is to maximize their own utility. They achieve this by evaluating potential partner's various attributes, such as socioeconomic status, beauty, education, and other relevant factors (Becker, 1973, 1974, 1993; Mansour and McKinnish, 2018).

a crucial role in reducing the number of child marriages around the world (Gaffney-Rhys, 2011). Following the international regulations and conventions aimed at reducing child marriage, most countries have instituted a legally defined minimum age for marriage. For example, utilizing data from 12 Sub-Saharan countries, Maswikwa et al. (2015) revealed that those nations which consistently enforced prohibitions against child marriage, witnessed a 40% reduction in its occurrence compared to counterparts without such steadfast regulations. When it comes to the legal impact of marriage laws, Bharadwaj (2015) shows that stricter application of marriage-related legislations contributes to delays in marriage and consequently defers fertility.

In general, the implementation of legal minimum marriage ages has become a prominent policy tool worldwide in addressing the issue of child marriage.[5] The map in Figure 2.1 shows the minimum legal age for countries around the world. The majority of countries have implemented a minimum marriage age, with only a few exceptions. The effectiveness of the legal minimum marriage age laws are dependent

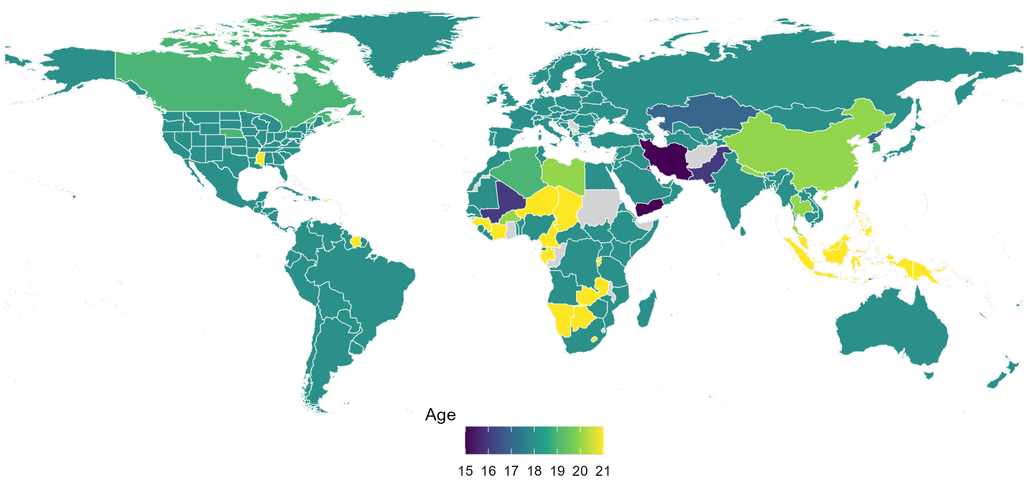


Figure 2.1: Global map for minimum legal age for marriage (female)

Notes: The graph shows the minimum legal age of marriage around the world for females. In some countries, there are state-level differences in marriageable age; we have considered the age which is practised in the majority of states. In many countries, minimum legal age for marriage is different with parental consent, but in India, both age is same.

[5]The minimum age at which a person can lawfully get married, and all marriages performed before that age are illegal in that nation or territory.

on the level of enforcement by countries. These laws are rigorously enforced in many countries, whereas in other cases, enforcement is less strict. Overall, a good number of countries exhibit a positive discontinuity in marriage rate at the legal marriage age (Collin and Talbot, 2023). The discontinuity in the marriage rate at the legal age of marriage suggests that many parents wait to reach their daughter at the legally defined age of marriage of that country to avoid legal complications. Adams and Andrew (2019) demonstrated that parents in India generally prefer to delay their daughters' marriages until they reach the age of 18 years, without extending the waiting period any further. The general practice suggests that an individual enter the marriage market for the first time at the legal age of marriage established by the government.

The institution of marriage continues to play an important role in the formation of human societies, and in most cases marriage involves the complex process of matchmaking. The primary components of any marriage market are the brides and grooms, and the matchmaking process depends on their individual attributes, such as wealth, education, attractiveness, and health status (Becker, 1973). But there are notable differences in the roles within households and the decision-making abilities of women of different societies around the world (Agarwal, 1997; Boserup, 1989; Duflo, 2012). As a result of this difference, distinct marriage market mechanisms emerged in developing countries are different from those observed in developed ones. In developed societies, it is common for the bride or groom to select their own partner, whereas in the developing world, the families typically arrange the matchmaking process.(Anukriti and Dasgupta, 2017). So, the process of matching for marriage is of significant importance, particularly in cultural contexts like India, where the prevalence of child marriages is coupled with alarmingly high rates of malnutrition among women.[6] We study what happens to a woman's health when she enters the marriage market and match-making process starts in case of India.

Due to historical practices such as dowry and complexity of a stratified social structure, the Indian marriage market exhibits unique characteristics. In India, despite many institutional reforms, women's marital prospects are influenced and shaped by conventional gender-stratified marriage rules (Banerjee, 1999). In addition, the caste system has historically played an important role in the matching process of the Indian marriage market (Corwin, 1977; Banerjee et al., 2013; Ahuja and Ostermann, 2016). The intersection of traditions, poor nutritional status of Indian women, and societal

[6] India is home to one-third of the world's child brides (UNICEF, 2023).

norms of matchmaking in India is intricate and worthy of further investigation. In South Asia, arranged marriages are social norms (Dube, 1997) and it is rational to assume that women, who are not malnourished, have more demand on the marriage market than the malnourished one. This could potentially lead parents to prioritize the daughters' well-being and consequently take better care of them by providing better food and nutrition as they approach entering the marriage market. The causal link between a woman's nutritional status and her transitions into this market has not been well understood. Recognizing this, we address a pressing question: In a country like India, where a large number of women live with malnutrition, how their nutritional and health needs are addressed, especially at the state-defined legal age of entrance into the marriage market? So far, research underscores the impact of marriage on multiple facets of an individual's life, ranging from socioeconomic dynamics to physical and emotional well-being (Williams and Umberson, 2004; Umberson et al., 2006; Carlson, 2012; Uecker, 2012). Due to social, cultural and religious factors, marriage is the only medium of childbearing and start of the reproductive years for women in many countries. This makes marriage a inevitable tie with the woman's identity and status to her role as a mother. In terms of the link between maternal health and socioeconomic development, Gill et al. (2007) showed that women's health has a direct implication in the long-term well-being of children. The female group who gets married earlier is relatively less educated and has less bargaining capability regarding her needs.[7] Verbrugge (1979) showed American married people are healthier than single and divorced people. Umberson (1992) tried to hypothesize the fact that marriage is beneficial to health because couples try to positively control each other's health behaviors. Waite and Gallagher (2001) discusses the point that married people are happier, healthier, and financially well-off than unmarried people. From a psychological point of view, Kiecolt-Glaser and Newton (2001) found marital functioning has significant positive impact on health outcomes. Marriage is also beneficial for the individual in terms of financial aspect (Loh, 1996; Blanchflower and Oswald, 2004; Becker and Becker, 2009), while the benefit of marriage is higher for men than for women (Gove et al., 1983). In general, there is a good amount of literature available, specifically in the context of developed countries, discussing the impact of marriage on personal well-being, but very little research on the health dynamics of women at the legal age of marriage for developing countries.

[7] Field and Ambrus (2008) showed 0.22 additional year of more schooling of girls keep them away from marriage one more year. They also found that delayed marriage is correlated with use of preventive health care service.

Our study examines variations in various health and nutritional indicators of women immediately before and after their 18-year age. At this point in their life, we analyze if there are any alterations in anthropometric measurements such as body mass index (BMI) and BMI-for-age z-score.[8] Furthermore, we evaluate variations in blood parameters, such as blood glucose and haemoglobin level, to understand the dynamics taking place during this transitional period. These blood indicators are directly related to dietary intake and help us understand the quality of food intake.

To investigate the causal impact of reaching the age of 18 and legally entering the marriage market on women's health outcomes, we used a sample of 50,745 women aged 17 to 19 years from the Indian Demographic and Health Survey (DHS). In our analysis we used a quasi-experimental approach, namely the regression discontinuity design (RDD). The RDD method estimates the causal effects of an intervention, event, or policy by comparing potential outcomes immediately below and just above a predetermined cut-off point. This cutoff is the value of a variable called the running variable or assignment variable, which decides whether an individual receives treatment or not. We conducted our analysis according to the latest guidelines on the regression discontinuity design method provided by Cattaneo et al. (2019). Given that our running variable is discrete (women's age in months), we used the local randomization method after determining the robust bandwidth.[9] The formal test recommended by Cattaneo et al. (2016) for an ideal bandwidth with a discrete running variable suggests a window size of 3 month on both sides of the cutoff. Moreover, we calculated the data-driven bandwidth as proposed by Imbens and Kalyanaraman (2012), which yielded same window size. In addition to inferences based on local randomization, we also incorporated local polynomial regression discontinuity design (RDD) estimations in our analysis.

Having balanced covariates at the cut-off point is the initial prerequisite for implementing a local randomization approach of RDD.[10] In our setting, the predetermined covariates are evenly balanced between the treatment and control groups at the cut-off. That is, those who have just turned age 18 and those who have not have

[8]The Body Mass Index is calculated by dividing a person's weight in kilograms by the square of their height in meters. It is an indicator of individual's underweight, overweight and some certain diseases. BMI-for-age z-score shows how a person's BMI varies against the reference population set up by the World Health Organization for any specific age cohort.

[9]The "bandwidth" refers to the range surrounding the cutoff score that determines which observations to be included.

[10]This balance guarantees that any differences observed in outcomes between the groups are the treatment effect instead of pre-existing differences. see Lee and Lemieux (2010).

similar characteristics. Our findings underscore the significant influence of legal age for marriage, thus entry to marriage market on various health indicators of women. Specifically, we observed a noticeable jump in a woman's health indicators shortly after she turns 18. The BMI jumps by 0.404 kg/m^2 immediately after reaching this legal marriage age. We noticed distinct trends based on urban and rural residences. For the rural population, the BMI increased by 0.401 units ($p < 0.01$) when reaching 18, suggesting a particularly pronounced effect in these areas. However, for the urban subset of our study population, the change in BMI was not statistically significant. Furthermore, we analyzed the BMI-for-age z-score, which measures relative weight adjusted for a individual's age and provides a more nuanced view of their nutritional status. We found that turning age 18 led to a significant increase in this z score of 0.129 standard deviations ($p<0.01$). This effect was more pronounced in the rural population, where the increase was 0.133 standard deviations ($p<0.01$).

Switching our focus to blood markers, we identified interesting patterns. The transition into age 18 is associated with an increase in blood glucose levels by 0.09 percent, although not significant when controls, time and state-level fixed effects are added. On the other hand, our analysis reveal marginally significant reduction in hemoglobin levels by 1.01 percent ($p< 0.10$). Sudden elevation in BMI along with increased blood glucose levels and reduced hemoglobin level indicate a potential change in consumption pattern, specially increase in carbohydrate intake. We do not have data on carbohydrate consumption, but we have data on meat and fish consumption. As mechanism of such change in anthropometric and blood markers, we analyzed daily fish and meat consumption and found significant jump in the likelihood of daily consumption of meat by 0.08 percentage point ($p<0.10$) and fish by 1.3 ($p < 0.05$) percentage point after reaching legal marriage age. Our findings provide compelling evidence of the profound influence that the entry to marriage market has on women's nutritional intake and health. In general, better investments in healthcare care brings an improved nutritional status, leading to gains in body mass index and changes in blood glucose levels.

We incorporated multiple checks and methodologies into our analysis, which ensures the robustness of our study. In addition to the local randomization-based regression discontinuity approach, we also added local polynomial-based regression discontinuity estimations for each outcome to strengthen the reliability of our results. In the local polynomial based design, we clustered our standard errors around the running variable. This clustering technique is a broadly practiced in studies where the running variable is discrete (Lee and Card, 2008). To determine the optimal bandwidth for the local randomization method, we adopted various data-driven methods. Additionally, our findings show that the notable change at the age of 18 is specific to females while the BMI z-score for the male sample does not exhibit any significant shift during this age transition.

Our study contributes to several streams of literature. First, our study contributes to the literature discussing the health accumulation of young women. It discusses a distinct perspective of women's health at the significant moment of entering the marriage market at a legally defined age. The existing literature focuses mainly on how marriage-related laws are enforced (Collin and Talbot, 2023) and what is the impact of such laws on child marriage (McGavock, 2021). To the best of our knowledge, we are the first to explore the health impact of entry Indian marriage market at the legal marriage age for women. We also add to the literature that investigates human capital investments in connection with the delay in marriage. Previously, Bharadwaj (2015) analyzed the impact of a law designed to delay marriage in Mississippi, resulting in an increase in educational enrollment. Our research explores the implications of the delayed marriage till legal marriage age on health investments.[11] In the complex cultural setting of India, it is imperative to examine health dynamics that may deviate from western paradigms. Therefore, our study not only enhances the current scholarly conversation on delayed marriage and health investment, but also expands our understanding of the impact of marriage age on health dynamics and broader dimensions of longterm well-being. Furthermore, our study contributes to the existing literature on the inequalities experienced by women in specific social norms and their potential consequences (Duflo, 2012; Jayachandran, 2015). We examine the link between a broader social norms of early marriage and the developmental obstacles faced by individuals- related to health, marriage, and age. This analysis presents a comprehensive understanding of the disparities in women's health experienced during these transition phases.

Our results are most relevant to the discussion regarding the policy change in the legal age for marriage, as in recent years, two of the world's most populous developing countries have taken the initiative to alter their legal marriage age for women. In 2020, Indonesia already shifted the minimum age of marriage for girls from 16 to 19 years.[12] India, which has the second largest population in the world, compris-

[11]We assume minimum legal age for marriage delays the entry to marriage market.
[12]https://www.unicef.org/press-releases/unicef-welcomes-recent-amendment-indonesias-marriage-act

ing about 18% of the global population, introduced a proposal in December 2021 to raise the minimum age for marriage for women from 18 to 21 years. However, it is important to note that this legislation has not yet been enacted until now (July 2024). According to UNICEF, India is the home of 18 million child brides, which is the largest number in the world.[13] It was our another motivation to select India as our study country, where we endeavoured to investigate what happens to women's health when they reach legal marriage age and enter the marriage market. The issue of the potential amendment of the minimum legal age for marriage for women in India, namely the proposal to raise it from 18 to 21 years, carries substantial implications for policy. The proposed alteration will result in a prolonged duration in which several women continue to be under the guardianship of their parents, due to the firmly entrenched patriarchal characteristics of the country's socio-economic framework. Moreover, within the context of developing nations, there is a significant disparity in literacy rates and female participation in the labor force which is considerably limited. Women frequently encounter obstacles when attempting to pursue career opportunities beyond their household responsibilities. In addition to existing difficulties, the limited financial resources of low-income households can result in neglect of the food and nutritional needs of female children. Furthermore, the prevalent matter of dowry entails that parents might save and invest limited resources for the purpose of their daughters' wedding. As the proposed policy changes are associated with more days under parental supervision, the potential delay in health investments should be addressed, which is related to the delayed entry into marriage market.

2.2 Background

2.2.1 Legal marriage age in India

The Child Marriage Restraint Act-1929, was an initial legislation to establish age limitations for marriage among male and female, during the period when India was under the governance of the British colonial administration. The legislation was enacted in 1930 and was applicable to all regions of India, with the exception of the state of Jammu and Kashmir.[14] The minimum age of marriage for females was set at age 14 while the minimum age for male was established at age 18 in this Act. Immediately after independence from British rule, in 1949, this Act was amended and

[13] https://www.unicef.org/india/what-we-do/end-child-marriage

[14] https://wcd.nic.in/child-marriage-restraint-act-1929-19-1929

the minimum age of marriage for women was increased to 15, while the marital age for males was not changed. The second amendment of this Act was in 1978, when the minimum age for marriage was increased for female to 18 years and for male to 21 years. In 2006, the 2006 Prohibition of Child Marriage Act was introduced without increasing the minimum age for marriage, but the strengthening of penalties for individuals who facilitate, encourage, or officiate child marriages. Recently there

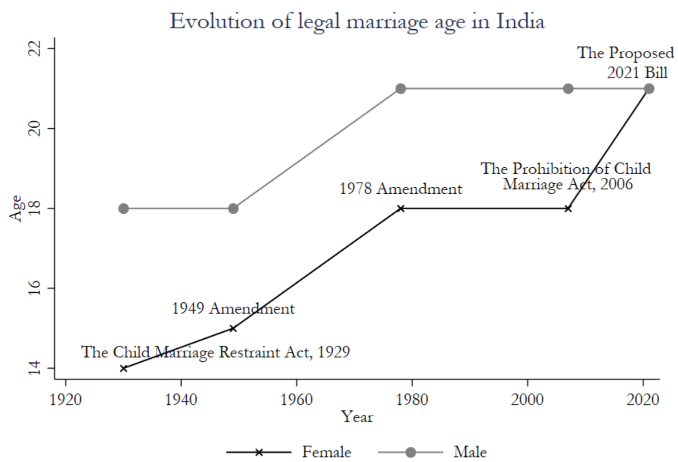


Figure 2.2: Evolution of legal marriage age in India

Notes: The graph shows the overtime development of legal marriage age in India for male and female. Age of individual is shown in Y axis while Year of the law introduced is in X axis.

has been an amendment initiative of the Prohibition of Child Marriage Act, 2006, with the introduction of The Prohibition of Child Marriage (Amendment) Bill, 2021. The proposed amendment seeks to establish a uniform legal age of 21 for marriage eligibility, applicable to individuals of both male and female. The Figure 2.2 shows the development of legal marriage age in India in the past century.

2.2.2 Child marriage in India

In India, child marriage is a significant social problem, with government and non-governmental organizations making constant efforts to decrease its prevalence. Although child marriage rates have significantly declined in recent years, India still has the highest incidence of child marriage among all countries.[15]

[15] `https://www.girlsnotbrides.org/learning-resources/child-marriage-atlas/atlas/india/`

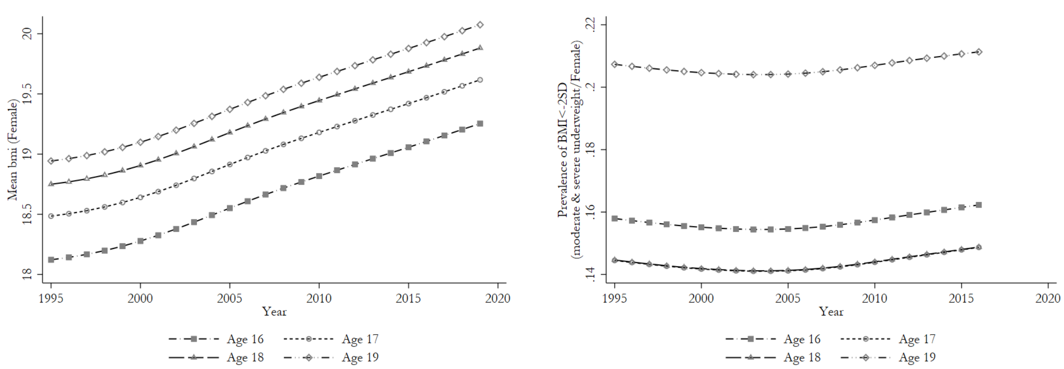
Figure 2.3: Percentage of Indian women aged 20-24 married before age 15 and 18

Notes: This graph illustrates the trend of child marriage in India for two age groups: below 18 years and below 15 years. The x-axis represents year, while the y-axis displays the percentage of marriages falling into each age category. The two lines trace the proportion of marriages involving brides under 18 years of age (solid line) and under 15 years of age (dashed line) over time. The graph illustrates a gradual decline in the prevalence of child marriage in India for both age groups. Source: The child marriage database, UNICEF (2024).

Sourced from the Indian National Family Health Survey (NFHS) and the Demographic and Health Survey (DHS), UNICEF provides data on the prevalence of early marriage among Indian women aged 20-24, specifically focusing on the percentages of those married before the ages of 15 and 18. Figure 2.3 shows a substantial decrease in early marriage over the last 25 years, from 1996 to 2021. In 1996, 21.81% Indian women, who fall within the age range of 20-24, were reported to have married before the age of 15. By 2001, a slight decrease in this statistic was observed, with a value of 20.65%. The next five-year period witnessed a notable decline, as the percentage decreased to 17.02% by 2006. The aforementioned decline continued with more intensity during the following decade, reaching a low of 12.99% in 2011 and further decreasing to 7.77% by 2016. The percentage then decreased to a promising 4.79% by the year 2021. Similarly, Figure 2.3 shows the occurrence of marriages before the age of 18, we also observe a persistent decline. In the year 1996, 53.15% female aged between 20-24 was married before they reached 18, and then decreased to 48.77% by the year 2001. In the subsequent five-year period, there was a decrease to 45.74% by the year 2006. The downward trend persisted, culminating in a figure of 37.95% in the year 2011. As of 2016, the prevalence of early marriage among Indian women aged 20-24 was recorded at 29.50%, indicating that fewer than one-third of this demographic had entered into marriage prior to reaching the age of 18. The aforementioned percentage experienced a subsequent decline, reaching a value of 23.25% in the year 2021. The data exhibits an optimistic trend, indicating the potential for a future in which the prevalence of early marriages among women could further decrease. Figure 2.3 illustrates the trend.

2.2.3 Female malnutrition in India

According to the World Health Organization (WHO), malnutrition is the deficits, excesses, or imbalances in an individual's consumption of energy and/or essential nutrients. Malnutrition among women is prevalent in developing nations, resulting in significant financial burdens and substantial consequences for future generations. India is no exception with a remarkably high prevalence of malnutrition among women(Narayan et al., 2019). The most common form of malnutrition among adolescent women are underweight and anemia.[16]



(a) BMI (b) Underweight

Figure 2.4: Overtime change of nutritional status by age 16-19

Notes: Figure (a) displays the change in body mass index (BMI) among Indian females aged 16 to 19 across the time span from 1995 to 2019. The second panel (b) presents the fraction of women with a body mass index (BMI) two standard deviations below the mean, indicating moderate to severe underweight. This change is observed across the time period from 1995 to 2016, within the same age group. *source:* The Lancet, (Abarca-Gómez et al., 2017)

[16]Underweight refers to a state in which an individual's body mass index (BMI) falls below the normal range. In the context of adults, a body mass index (BMI) that falls below 18.5 kg/m^2 is typically considered indicative of being underweight. Anemia is a condition that arises due to a lack of essential nutrients such as iron, vitamin B12, or folic acid in body. In anemia, the quantity red blood cells reduces, causing to decrease in the capacity of the blood to transport oxygen.

Panle (a) of Figure 2.4 depicts the mean body mass index of adolescent Indian females aged 16 to 19 years between the years 1995 and 2019.[17] The graph indicate a substantial increase in BMI across various age groups over time, indicating an improvement in the country's nutritional condition. However, panel (b) shows there was no significant decrease in the prevalence of moderate to severe underweight over time. It is worth noting that the age cohorts of 17 and 18 demonstrate the most minimal prevalence of underweight. The decrease in the prevalence of underweight individuals in this specific age group signifies the increased focus and concern directed towards women approaching the legal age of marriage in India.

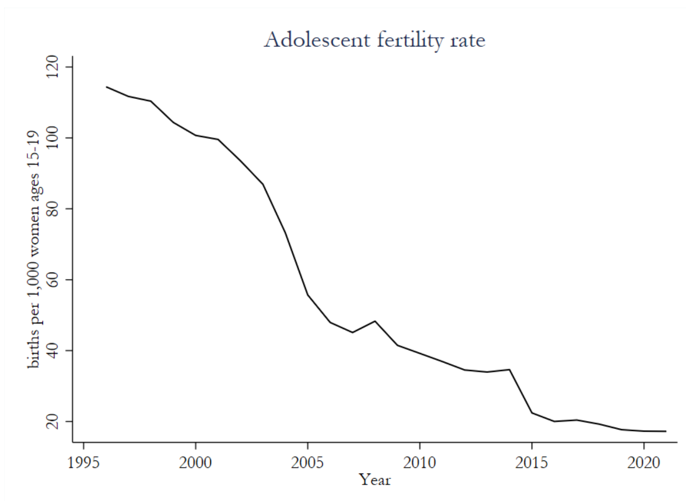


Figure 2.5: Adolescent fertility rate over time

Notes: The graph shows the adolescent fertility rate of Indian women in the age group 15-19 from the year 1995 to 2021. The vertical axis represents birth per 1,000 women and hprozontal axis represents year. Source: Gender statistics, The World Bank Group (2023)

Maternal mortality is strongly associated with nutrition, with the risk of death among females being highest in the lowest 10% of the BMI decile distribution (< 16.39), as demonstrated by Hosegood and Campbell (2003). Figure 2.5 illustrates the adolescent pregnancy rate spanning the years 1995 to 2021. It presents the changing socio-cultural norms within society, specifically pertaining to the practice of early marriages and childbearing. Over the course of a quarter century, there has been

[17]The numbers are derived from a comprehensive dataset given in a Lancet article by Abarca-Gómez et al. (2017), where the authors collected anthropometric data of 128.9 million individuals aged 5 years and older. Among this population, 31.5 million individuals fell within the age range of 5 to 19 years. Using a Bayesian hierarchical model, the authors examined the trends in BMI from 1975 to 2016 in 200 countries. We have only used Indian data.

a notable decrease in the birth rate per 1,000 women within the age range of 15 to 19. Commencing from a marginally higher figure of 110 births in the year 1995, the birth rate had a substantial decline, reaching approximately 20 births by the year 2020. This phenomenon is possibly indicative of the changing attitudes about early marriages and the importance of adhering to the legally established age to enter into marriage market.

2.3 Method

2.3.1 Conceptual framework

Our primary assumption posits that adolescent women enter the marriage market at legal marriage age of 18 in India. We assume women with below average anthropometric and health status are less attractive and have less demand in the marriage market. We consider statistical evidence that suggests that, on average, Indian women have a high likelihood of having a nutritional status below the recommended reference values. So, we conclude that when a woman turns 18 years and enters the marriage market, parents immediately pay more attention than before to her health and welfare so that her nutritional condition improves and the women get a better match. This increased attention aims to improve the woman's nutritional condition, thereby enhancing her prospects of securing a better marital match. Failing to provide such special care may necessitate paying a higher dowry to compensate for the woman's suboptimal health and nutritional status. Following assumptions of the theory of marriage market (Becker, 1973), we develop a model which incorporates the costs associated with both health investments and dowry payments. In general, investments in health at the legal marriage age are determined by:

$$M(I) > M(I - \delta) - c(I - \delta) - d(I) \tag{2.1}$$

Here, I stands for the woman's investment in her health and M for her prospective match quality in the marriage market. M is rising in I, and a woman must invest a certain amount in her health to find a high-quality match. where, $M = f(I)$ can be expressed as the relationship between health investment and match quality, where $f' > 0$ and $f'' < 0$. $c(I)$ is the cost of health investment. $d(I)$ is the additional dowry cost if the bride is malnourished. δ is a small amount of additional health investment, representing the marginal benefit of investment. Our model implies that the decision to invest in health and nutrition when reached at legal marraige age is determined by comparing the expected benefits (that is, improved match quality) and the expected costs of the investment in health.

2.3.2 Model identification

To identify the effect of reaching legal marriage age and entering the marriage market we used the regression discontinuity (RD) design. RDD is a quasi-experimental approach used to estimate the causal effects of treatment by leveraging pre-specified cutoff points. This method has recently had a resurgence as a tool for conducting causal studies, despite its origins being traced back to the initial proposition of (Thistlethwaite and Campbell, 1960). This approach remained relatively dormant until the influential study conducted by (Angrist and Lavy, 1999), which analyzed the causal effect of marginal reductions in class sizes on academic performance measured by test scores. Hahn et al. (2001) expanded methodology for the more robust identification and estimation of regression discontinuity designs. Imbens and Lemieux (2008) and Lee and Lemieux (2010) offered extensive guidelines on bandwidth selection and model specification, which contributed to improve the reliability and validity of the RDD estimators. Calonico et al. (2014) examined concerns related to bandwidth selection and bias correction in RDD design and presented robust estimators for constructing confidence intervals for average treatment effects at the cutoff in various types of RDD. Cattaneo et al. (2019) provides further breakdown on multidimensional aspects of the standard sharp regression discontinuity design.

In this study, we analyze how the entrance to the marriage market at the legal age of marriage affects women's health in India. The task of accurately measuring the health impact of reaching legal marriage age and thus entering the marriage market can pose significant difficulties with the presence of potential endogeneity issue. There exist several potential unobserved factors, such as education, peer pressure, and family support, that can influence the health outcomes of young women during their transition to 18 years of age, which we cannot control. The feasibility of implementing a randomized controlled trial (RCT) in this particular context is limited and not rational. As a purely experimental design is not feasible in our study context, we resort to the quasi-experimental RDD. The analytical context of our study mostly resembles the method of RDD, as it involves a cut-off of legal marriage age, below which everyone is assumed in control group as they can not be legally a part of marriage market and over this cut-off point everyone is part of it.

Here, the age of women, denoted by X_i, measured in months, serves as running variable. This age variable is central in our regression specification as it determines the treatment assignment:

$$D_i = \begin{cases} 1 & \text{if } X_i \geq c \\ 0 & \text{if } X_i < c \end{cases} \tag{2.2}$$

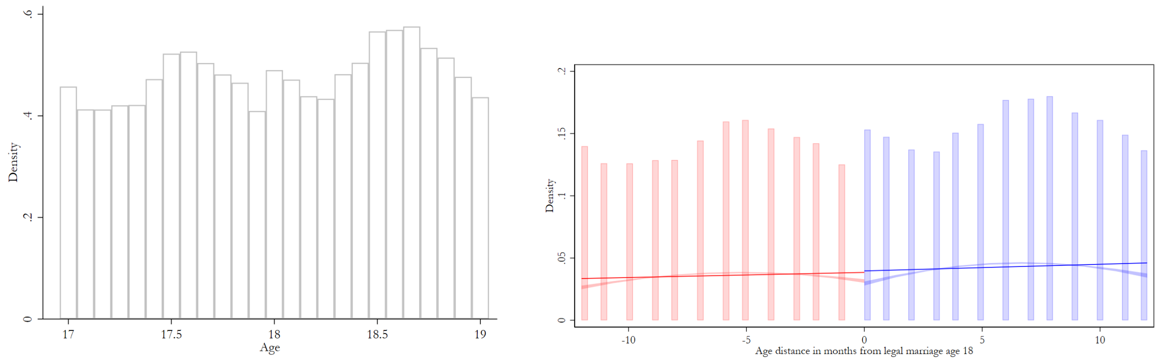
where c is the cutoff, set at 18 years. Thus, if a woman's age X_i is greater than or equal to the cutoff, she is considered to have legally entered the marriage market and thus is in the treatment group. On the other hand, if $X_i < c$, she remains in the control group. Formally, the treatment D_i can be viewed as a function of the running variable:

$$D_i = f(X_i) \tag{2.3}$$

For the efficacy of the RDD design, it is essential that the age variable X_i is observed accurately in the data, and the predetermined threshold c is well-acknowledged. Both satisfy in our case, which confirms the suitability of using the RDD approach in our context, providing a robust mechanism to discern the causal effects of entering the marriage market at government-defined age on subsequent health outcomes.

All required assumptions of RDD are also satisfied in this context. Independent assignment to treatment is a necessary condition for qualifying as an RDD study. Legal entry into the marriage market is exclusively based on the cut-off age and is not determined by potential health outcomes, which ensures independent assumption. The second assumption to be fulfilled is the continuity of conditional regression functions. The expected health outcomes for women $E[H(0)|X_i = x]$, given their age just before and after turning 18 years (for those who are not yet in the marriage market) and $E[H(1)|X_i = x]$ (for those who are already in the marriage market) , are continuous functions of age, represented by X_i. It implies that, at the threshold, the only difference in outcome between the two sides of that cutoff is the treatment itself. Exogeneity of the cutoff is also satisfied as the age barrier (18 years) to entry in the marriage market is not chosen by the women or their families, rather it is externally imposed from government.

It is also necessary that there is no manipulation at the cut-off point. Women cannot cannot strategically choose/change their age to enter or bypass entering the marriage market. Age data collected in the Indian Demographic and Health Survey (DHS) are considered reliable and free of any intentional manipulation. In our scenario, the running variable is not subject to manipulation. Figure 2.6 shows an early assessment of potential manipulation in the running variable. Figure 2.7 suggests a generally uniform dispersion of data points on both sides of the central point.



(a) Age distribution (in Month)

(b) Manipulation testing plot

Figure 2.6: Change in bmi within months of entering marriage market

Notes: Panel (a) illustrates the density distribution of ages in our sample of females aged 17 to 19 years old, measured in months. The distribution is evenly spread across these ages, indicating a well-balanced sample suitable for conducting regression discontinuity analysis. Panel (b) shows the manipulation testing plot (N=50,739) proposed by Cattaneo et al. (2020).

Formal manipulation tests are necessary to test for endogenous sorting or self-selection of units exposed to the treatment assignment. Panel (b) of Figure 2.6 illustrates the the manipulation test introduced by Cattaneo et al. (2020), which is also valid for discrete running variables (Cattaneo et al., 2018).

The plot demonstrates the density distributions of age measured in month on either side of the 18 year threshold, with red bars representing those below the age of 18 years and blue bars representing those who are 18 years or older. Out of a sample size of 50,739 observations, 23,283 observations are found to be below the threshold, while 27,456 observations are located above it. By employing local polynomial density estimation, with an effective bandwidth of 12 months, the obtained t-statistic is -1.59 (p = 0.11). This finding indicates that, there is not enough evidence to conclude that there was manipulation or aberrant clustering around the legal marriage age 18 in our sample.

Furthermore, we ensure that all other factors that could potentially influence women's health outcomes, household age, household head sex, education, electricity access, residence, wealth index do not show a "jump" at the 18-year threshold. The lack of abrupt changes in these variables at the cut-off ensures the absence of potential bias in our analysis. In the following, Figure 2.7 plots the smoothness of several important covariates that are available in our dataset. We can see that in our predetermined covariates are smooth and shows no manipulation at the age of 18.

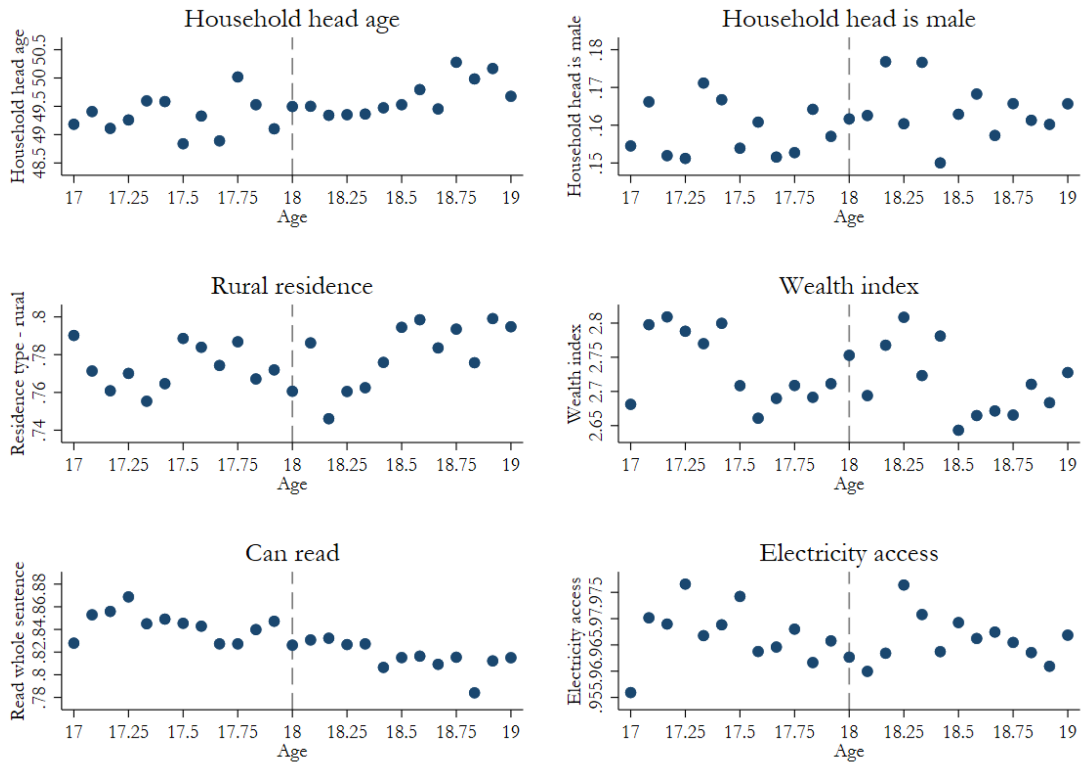


Figure 2.7: Smoothness in covariates

Notes: This figure shows smoothness in several variables at cut-off age 18 (N=50,739). Overall, it shows that there is no immediate jump at the legal marriage age 18 for these variables.

Figure 2.7 indicates predetermined covariates exhibit a regular and uninterrupted pattern. Significantly, it is evident that there is no observable "leap" or sudden jump around the age of 18. This implies that our data does not exhibit indications of manipulation concerning the marriage age threshold.

2.3.3 Model selection

The conventional RDD method relies on the continuity assumption of the running variable. The most recent RD guideline of Cattaneo et al. (2023) suggests that the direct applicability of the continuity-based local polynomial method is limited when the running variable is not continuous, and then the application of this method mostly depends on the number of mass points. Then it is suggested to use the local randomization-based method in a setting with discrete running variable. The primary methodological selection of our study comes from the discrete running variable, specifically the measurement of age in months. And following Cattaneo et al. (2023), we use local randomization based approach of sharp regression discontinuity design. Additionally, we provide the local polynomial-based RDD analysis, clustering standard errors by the running variable. This approach is widely used in the prominent economic literature that features a discrete running variable (Card et al., 2008; Chetty et al., 2013; Frandsen, 2017).

In the context of studying the effect of entering the marriage market at legal marriage age on health outcomes using a sharp regression discontinuity design (RDD) with local randomization, the model is:

$$\alpha_{\text{LR}} = E[H_{1i} - H_{0i} \mid X_i \in W] = \frac{1}{N_1} \sum_{\substack{X_i \in W \\ D_i = 1}} H_i - \frac{1}{N_0} \sum_{\substack{X_i \in W \\ D_i = 0}} H_i \tag{2.4}$$

Here, α_{LR} is the local average treatment effect (LATE) within the window W. H_{1i} and H_{0i} are health outcomes for female i under treatment and control group, respectively. $E[H_{1i} - H_{0i} \mid X_i \in W]$ is the expected difference in health outcomes between treated and untreated women within the window W. W is the range of values near the cutoff point where the assumption of random assignment of the treatment can be considered valid. X_i is our running variable: age in month for individual i. We determine the assignment to control or treatment groups based on its value relative to a threshold. D_i is treatment indicator variable for individual i. It is 1 if the individual is in the treatment group and 0 otherwise. within the window W, N_1 and N_0 is the number of treated and untreated individuals, respectively. $\frac{1}{N_1} \sum_{\substack{X_i \in W \\ D_i = 1}} H_i$ and $\frac{1}{N_0} \sum_{\substack{X_i \in W \\ D_i = 0}} H_i$ represent the average health outcomes within the window W for treated and untreated individuals, respectively.

Additional assumption necessary for local randomization based RDD is inside the selected window on both side of the cutoff $W = [c - w, c + w]$, treatment assignment is as good as random. That means we establish a covariate balance around the selected window. The local randomization method satisfies all the other assumptions mentioned in the previous section and gives a local average treatment effect (LATE) at the threshold. The regression we estimate is-

$$H_i = \alpha + \delta D_i + \varepsilon_i \tag{2.5}$$

With δ being LATE, representing the average difference in health outcomes for those who just entered and those who did not enter into the marriage market.

2.3.4 Window selection for local randomization

We need to determine an ideal window or threshold in which the samples have similar characteristics. Cattaneo et al. (2015) provide a methodology to determine the ideal window, within which local randomization is deemed to be valid. Treatment assignment within this test is anticipated to be "as if random", which guarantees that both pre-intervention variables and other post-intervention outcomes, unaffected by the treatment, exhibit similar distributions for both treated and control units. It starts with an initial "small" window. For each covariate, it test if the treatment has no effect. If this smallest p-value is less than a predetermined threshold (0.15), the initial window is considered too large and then reduced. This step repeats until a window is found where the p-value is below the set threshold.

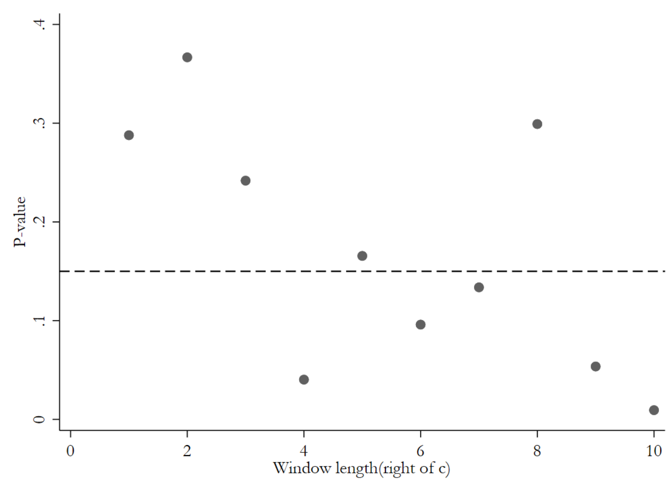


Figure 2.8: Window selection test

Notes: The graph shows window selection test (Cattaneo et al., 2016) with three predetermined covariates age of the household head, sex of the household head and household electricity access.

We conduct our test based on the module using the STATA command rdwinselect (Cattaneo et al., 2016). To calculate the bandwidth, we used three predetermined covarates: age of the household head, sex of the household head, and electricity access to households. The result of the formal test for an ideal window with discrete running variable gives window size 3 month on both sides, as shown in Figure 2.8.

Furthermore, we calculated the data-driven bandwidth for the local linear regression suggested by Imbens and Kalyanaraman (2012), which also gives same bandwidth of 3 months.

Table 2.1: Balance table

	All sample	Experiment sample		
	All	Control	Treatment	Difference
	(1)	(2)	(3)	(3)-(2)
Age of household head	49.51	49.57	49.45	0.12
	(11.09)	(11.02)	(11.16)	(0.21)
Household size	5.76	5.79	5.73	0.05
	(2.32)	(2.31)	(2.33)	(0.04)
Rural residence	0.77	0.78	0.76	0.01
	(0.42)	(0.42)	(0.42)	(0.01)
Can read	0.83	0.84	0.83	0.01
	(0.37)	(0.37)	(0.38)	(0.01)
Muslim religion	0.13	0.13	0.13	-0.00
	(0.34)	(0.34)	(0.34)	(0.01)
Household has electricity	0.96	0.97	0.96	0.00
	(0.19)	(0.18)	(0.19)	(0.00)
Wealth index combined	2.72	2.70	2.74	-0.03
	(1.34)	(1.34)	(1.35)	(0.02)
Education in single years	9.56	9.51	9.61	-0.10*
	(2.90)	(2.77)	(3.03)	(0.05)
Observations	11,655	5,735	5,920	

Notes: This table presents the summary statistics of key covariates for the full sample and separately for the control and treatment groups. It includes only samples of individuals who turned 18 years old within a 3-month window. The treatment group includes respondents who became 18 years old, controls are below 18. The last column reports the difference in means between the treatment and control groups. Standard deviation are in parentheses. Residence is a binary indicator, with 1 representing rural areas and 0 representing urban areas. Can read is a binary indicator, with 1 indicating the ability to read and 0 indicating the inability to read. Religion is a binary indicator, with 1 representing Muslim households and 0 representing non-Muslim households. Electricity access is a binary indicator, with 1 indicating that the household has access to electricity and 0 indicating no access. Wealth index combined is a composite measure of the household's wealth status. Education in single years is years of schooling by respondends. The last row shows the total number of observations for each column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.1 presents the balance of covariates around the cutoff point, which is essential for our reliance on the local randomization method. The table demonstrates that key covariates, including age and sex of the household head, household size, place of residence (rural/urban), religion, literacy, wealth index, and access to electricity, are well-balanced within the 3-month window around the age of 18. This balance in baseline characteristics provides evidence that the treatment assignment near the cutoff is as good as random, supporting the validity of our empirical strategy.

2.4 Data and sample construction

We use data from the India Demographic and Health Survey (DHS), 2019-2020. The data-set is unique due to its extensive coverage and comprehensive level of detail, it contains information pertaining to an impressive total of 636,699 households. The data set presents a characteristics of a large sample of individuals, consisting of 724,115 women and 101,839 men. From this dataset we used a subset of 50,761 non-pregnant young women, falling within the age range of 17 to 19 years. One notable feature of the DHS dataset is the inclusion of anthropometric measures, which are carefully recorded for both married and unmarried women. Furthermore, the DHS data set has several individual-level and household-level details, which facilitates a comprehensive examination of demographic patterns and health metrics.

Firstly, we consider body mass index (BMI) and BMI-for-age z-score of our sample. Body mass index (BMI) is an anthropometric measurement to asses an individual's nutritional status. More precisely, BMI represents the fatness or metabolic efficiency through which the body's energy level can be assessed. BMI is calculated by dividing an individual's weight in kilograms by the square of their height in meters. BMI below 18.5 indicates the presence of potential chronic energy deficiency (CED) (James et al., 1988). Similarly, a BMI below 18.5 indicates underweight, a BMI between 18.5 and 24.9 is considered normal, a BMI between 25 and 29.9 is categorized as overweight, and a BMI over 30 is classified as obese (WHO Global InfoBase team, 2005).

While BMI is a general anthropometric measurement, growth reference is recommended to use for analyzing nutritional status of adolescents (de Onis et al., 2007).[18] As our respondents are under 19 years old, we used the BMI-for-age Z score as an additional measure. This score represents the number of standard deviations by which an individual's BMI deviates from the mean or median BMI of a reference population of the same age and sex. According to WHO, Individuals (between age 5-19) with a z-score exceeding +1 standard deviation (SD) are classified as overweight, which translates to a BMI of 25 kg/m^2 at 19 years of age. A score above +2 SD is denoting of obesity, equating to a BMI of 30 kg/m^2 when the individual reaches 19 years. In contrast, a score below -2 SD is a sign of thinness.

The use of BMI and BMI-for-age z-score gives us a proxy measure of health and nutritional care as Body mass index is directly related to the quality and quantity of

[18]Growth reference is a statistical summary of anthropometric measurements in a reference group of children. It is typically presented as a frequency distribution across various age groups (Cole, 2012).

energy intake for an individual (Shetty and James, 1994; James et al., 1988; Herrera et al., 2003; Kurpad et al., 2005) If there is a sudden increase in health and nutritional care for any individual, it is anticipated that her body mass index will rise. We expect to capture any increased health and nutritional care during entry into the marriage market through BMI analysis.

Dietary components play a substantial and clinically significant role in the regulation of blood glucose levels in the human body (Wheeler and Pi-Sunyer, 2008; Russell et al., 2016). The blood glucose level serves as a marker for diabetes.[19] Therefore, the level of blood glucose can act as a vital marker of an individual's metabolic health and carbohydrate consumption. If there is an increased dietay care, it can be easily captured through blood glucose level. The blood glucose data in the DHS survey are collected random time of the day, not as fasting blood glucose.

2.5 Results

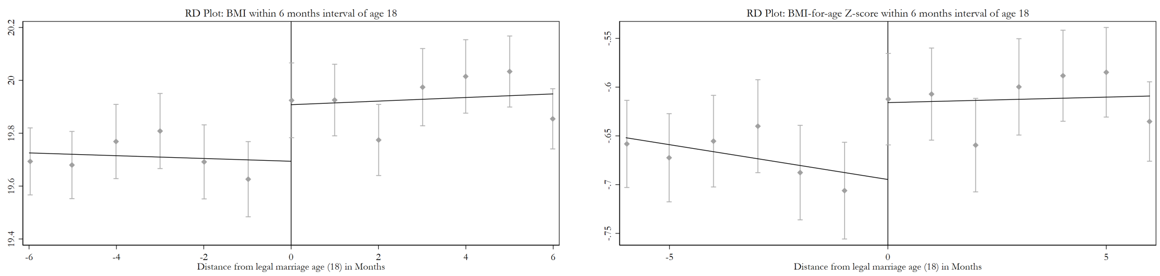
2.5.1 Impact on anthropometric indicators

To evaluate the consequences of transition to legal marriage age, we analyze changes in both body mass index and age-adjusted BMI-for-age z-scores.[20] Regression discontinuity plot in Figure 2.9 shows a noticeable jump in the BMI and BMI-for-age z-score indicators at the month they become age 18. Table 2.2 shows result of the local randomization-based regression discontinuity design with a calculated bandwidth of 3 months on each side of the legal marriage age threshold. We find a consistent and noticeable jump in both indicators at the month females become 18. Although, the magnitudes exhibit variation across subgroups based on place of residence, the full sample displays an increase of 0.166 kg/m^2 (p<0.01) (0.404 kg/m^2 (p<0.01) when controls are added and state and time fixed effect are applied) in body mass index who just became 18 and consequently entered the marriage market. More pronounced pattern is observed among females in rural areas with 0.196 kg/m^2 increase (p<0.01) in BMI after reaching to age 18. The effect is 0.401 kg/m^2 (p<0.01) when controls, time

[19]Normal level of blood glucose is between 72 mg/dL (4 mmol/L) and 108 mg/dL (6 mmol/L). Elevations in glucose levels, namely prediabetes, are called when the blood glucose level is 100 to 125 mg/dL (5.7 to 6.4 mmol/L) after at least 8 hours of fasting. If random blood glucose is more that 11.1 mmol/L (200 mg/dL) or fasting glucose level is 126 mg/dL (7 mmol/L) or above, it is considered as diabetes (Mathew et al., 2024).

[20]BMI-for-age z-scoresis calculated using growth references provided by the World Health Organization (de Onis et al., 2007).

and state fixed effects are applied. The effect in urban sample is relatively smaller and statistically not significant. The inclusion of controls and state-fixed effects in the does not change the scenario for this sub-sample.



(a) BMI

(b) BMI-for-age z-score

Figure 2.9: Change in BMI and BMI-for-age z-score within 6 months around the age 18

Notes: This figure illustrates a discontinuity analysis of absolute BMI in Panel (a) and BMI-for-age z-score in Panel (b) within a 6-month interval around the legal marriage age of 18. The plot includes a first-order polynomial fit (p=1) and 95% confidence intervals (CI).

The mean BMI-for-age z-score of our sample is -0.65, when we compare it with the WHO reference group, it becomes evident that average BMI of Indian women falls below the worldwide median for their age group. This variation suggests the presence of potential malnutrition concerns or distinctive health patterns that are specific to the age group in India. Urban areas, with a score of -0.52, exhibit a lower proximity to the worldwide median. Similarly, women from rural region display even lower score of -0.69, indicating potential discrepancies in dietary consumption, lifestyle, or cultural expectations.

Table 2.2 also highlights the changes in BMI-for-age z-score among Indian women when they reach the legal age of marriage. As women undergo the transition to this crucial stage of life, there is a significant and consistent increase in the bmi-for-age z score values. In the total sample, reaching the age of 18 is associated with an increase of 0.049 standard deviation in BMI-for-age z-score ($p<0.01$). Applying additional controls and fixed effectsthe estimate becomes 0.129 sd ($p<0.01$). The increase is 0.057 standard deviation ($p<0.01$) in the rural sample, which is 0.133 if additional controls and state fixed effects are applied. Parallel to local randomization-based RD estimates, we employed the local polynomial-based regression discontinuity design method with clustering standard errors by the running variable age. Table 2.B.3 and Table 2.B.4 presents the local polynomial based regression result for outcomes

Table 2.2: Change in body mass index

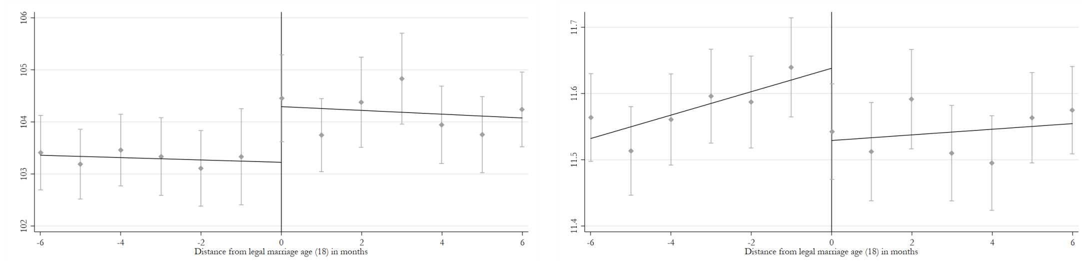
	BMI						BMI-for-age- z-score					
	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.166***	0.404***	0.196***	0.401***	0.046	0.334	0.049**	0.129***	0.057**	0.133***	0.019	0.086
	(0.057)	(0.121)	(0.062)	(0.130)	(0.141)	(0.294)	(0.020)	(0.043)	(0.022)	(0.047)	(0.047)	(0.099)
Age of household head		0.008***		0.005*		0.019**		0.003***		0.002*		0.005**
		(0.003)		(0.003)		(0.008)		(0.001)		(0.001)		(0.002)
Household size		-0.025*		-0.017		-0.035		-0.007		-0.006		-0.005
		(0.013)		(0.015)		(0.032)		(0.005)		(0.005)		(0.012)
Household head female		0.160*		0.122		0.262		0.053*		0.038		0.100
		(0.081)		(0.088)		(0.202)		(0.030)		(0.032)		(0.071)
Household has electricity		-0.170		-0.162		0.076		-0.059		-0.060		0.060
		(0.143)		(0.146)		(0.703)		(0.050)		(0.051)		(0.247)
Wealth index combined		0.241***		0.202***		0.255***		0.074***		0.065***		0.079***
		(0.025)		(0.031)		(0.065)		(0.009)		(0.011)		(0.022)
Age (in months)		-0.090**		-0.081**		-0.102		-0.030**		-0.029**		-0.024
		(0.035)		(0.038)		(0.084)		(0.013)		(0.014)		(0.029)
Mean depvar	19.84	19.84	19.72	19.72	20.28	20.28	-0.63	-0.63	-0.67	-0.67	-0.51	-0.51
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11,117	10,858	8,612	8,391	2,505	2,467	11,134	10,875	8,626	8,405	2,508	2,470
R^2	0.001	0.045	0.001	0.039	0.000	0.065	0.001	0.045	0.001	0.042	0.000	0.063
Adjusted R^2	0.001	0.040	0.001	0.032	-0.000	0.044	0.000	0.040	0.001	0.035	-0.000	0.041

Notes: This table presents the local randomization based regression results with 3 month bandwidth on each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on bmi and bmi-for-age Z-score. treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns (e.g., 2, 4, 6, etc.) incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of BMI and BMI-for-age z-score. The estimates from polynomials based approach is consistent with the local randomization-based approach with controls and fixed effects. Besides BMI, we also did similar analysis other anthropometric outcomes such as Rohrer's index and arm circumference of women in our target age group.[21] The estimates shown in Table 2.A.2 in the Appendix gives clear indication of these indicators at the age 18, when a women enters to marriage market. All the outcome indicates a better care and health related investment during the time a women becomes 18 and enters to marriage market.

2.5.2 Impact on blood markers

We analyze two important blood markers, namely blood glucose and blood hemoglobin levels, which provide us valuable insights into the health and nutritional quality of our sample around the legal age of marriage. Figure 2.10 shows the discontinuity in the blood glucose and hemoglobin level in the month the woman reaches 18 years old.



(a) Change in blood glucose

(b) cCange in hemoglobin level

Figure 2.10: Change blood markers within 6 months bandwidth

Notes: This plot presents the discontinuity (RD) analysis of blood glucose (Figure (a)) and hemoglobin level (Figure (b)) centered around the legal marriage age of 18. The x-axis represents the distance in months from the age of 18, covering a period of 6 months before and after. The y-axis represents the respective blood marker. The plot includes a first-order polynomial fit (p=1) with 95% confidence intervals (CI) for visualizing the potential discontinuity at the cutoff age of 18.

Employing the calculated bandwidth of three months on either side of the marriage age cut-off, in Table 2.3 we analyze changes in blood indicators for the entire sample and stratified subsets of individuals residing in rural and urban settings. The regression results reveal a positive association between reaching the marriage age, a proxy for entry into the marriage market. For the entire sample, we observe an

[21] Rohrer index also known as corpulence index or ponderal index is ratio of weight (M) and height (L); $R = M/L^3$.

increase of 0.9 percent ($p<0.01$) in glucose levels, but this significance does not sustain when we add controls and fixed effects. The impact on blood glucose is more pronounced in the rural subset, where the jump was 1 percent ($p<0.01$) in analyses without control variables. Upon the inclusion of controls and accounting for state fixed effects and time (year x month) fixed effects , the magnitude of these effects moderated to 0.80 percent, without statistical significance. Specifically, the jump in glucose levels in urban areas is not statistically significant.

We found negative impact on blood hemoglobin level when someone enters to marriage market. For treatment group, we found a 1.1 percent ($p<0.1$) decline (without control and fixed effects, 0.6 percent ($p<0.05$) in hemoglobin level compared to the control group. The rural samples without controls shows a decline of 0.06 percent. When we add control and state fixed effect, the result is not significant. Also, the urban sub-sample does not show any change in hemoglobin level. Regarding hemoglobin levels, the effect of reaching age 18 and entering the marriage market is comparatively less uniform. Results for change in blood glucose and blood hemoglobin with the local polynomial-based regression discontinuity design with clustering standard errors by the running variable are presented in Table 2.B.5 and Table 2.B.6. The result is consistent with the local randomization-based approach.

We also examine the impact of reaching 18 years of age on blood pressure measurements. The results indicated a decrease in systolic and diastolic measures for the treatment group. we found a 0.7 percent decrease ($p<0.1$) in systolic blood pressure for full sample and 1.4 percent decline ($p<0.1$) in urban subsample. We find reduced but not considerable impact on the diastolic blood pressure level of our treatment indicator. The local polynomial regression discontinuity design, which clusters standard errors by the running variable (age), provides consistent results as shown in Table 2.B.7 and Table 2.B.8 in the Appendix.

Table 2.3: Change in Blood markers

Dep. var	Glucose level (log)						Hemoglobin level (log)					
Sample	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.009***	0.009	0.010***	0.008	0.006	0.007	-0.006**	-0.011*	-0.006*	-0.010	-0.007	-0.015
	(0.003)	(0.006)	(0.003)	(0.007)	(0.006)	(0.012)	(0.003)	(0.006)	(0.003)	(0.007)	(0.006)	(0.013)
Age of household head		0.000		0.000		0.000		0.000		0.000		0.001*
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Household size		-0.000		-0.000		-0.001		-0.000		0.000		0.000
		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
Household head female		0.004		0.003		0.005		-0.002		-0.002		-0.005
		(0.004)		(0.005)		(0.008)		(0.004)		(0.005)		(0.009)
Household has electricity		0.011		0.010		-0.004		-0.001		-0.000		0.014
		(0.008)		(0.008)		(0.031)		(0.008)		(0.008)		(0.032)
Wealth index combined		-0.001		0.003*		-0.007**		0.007***		0.007***		0.006*
		(0.001)		(0.002)		(0.003)		(0.001)		(0.002)		(0.003)
Age (in months)		-0.001		-0.001		-0.001		0.003		0.003		0.004
		(0.002)		(0.002)		(0.004)		(0.002)		(0.002)		(0.004)
Mean depvar	102.57	102.57	102.57	102.57	102.57	102.57	11.42	11.42	11.42	11.42	11.42	11.42
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11022.000	10766.000	8551.000	8333.000	2471.000	2433.000	11011.000	10755.000	8543.000	8325.000	2468.000	2430.000
R^2	0.001	0.062	0.001	0.061	0.000	0.090	0.000	0.046	0.000	0.050	0.001	0.057
Adjusted R^2	0.001	0.057	0.001	0.055	-0.000	0.068	0.000	0.040	0.000	0.044	0.000	0.035

Notes: This table presents the local randomization based regression results with 3 month bandwidth on each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on blood glucose and and blood hemoglobin level. treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.6 Mechanisms and robustness check

The obtained result supports our hypothesis that there is a positive association between a woman's legal marriage age and anthropometric outcomes. The empirical findings indicate a substantial and statistically significant increase in important anthropometric measures, such as BMI, BMI-for-age z-score, within a period of 3 months when she reach legal marriage age. To investigate the potential mechanisms driving these improvements, we analyzed the dietary habits and health-seeking behaviors of unmarried women nearing this critical age threshold. Figure 2.B.9 depicts positive and statistically significant jump in the probability of daily consumption of meat and fish. For the full sample, we found 0.80 percentage point increase ($p<0.10$) in probability of consuming meat in daily basis for the treatment group. Similarly, increase in probability of daily fish consumption is 1.3 percentage point ($p<0.05$) when just reached 18, compared to those yet to reach this age. Given the unavailability of specific carbohydrate consumption data, we infer an associated increase in the intake of rice and other carbohydrates, typically consumed alongside protein-rich foods in India. This inferred dietary pattern is likely contributing to the observed enhancements in anthropometric measures. The increase in the probability of frequent consumption of meat, fish, and eggs, could potentially be ascribed to intensified parental attention towards dietary decisions among women who have attained the age of legal marriage. Furthermore, our data suggest a slight, yet statistically insignificant, increase in healthcare facility visits by the sample, indicating a possible elevation in healthcare engagement upon reaching legal marriage age. The details of our mechanism analysis are presented in section 2.B.2 in the Appendix.

We performed several robustness analysis and falsification tests to validate our regression discontinuity design specification. First, in Table 3.1 we provide evidence that the predetermined covariates are well-balanced at just below and above the female legal marriage age of 18 years, within a 3-month window. This balance in covariates ensures that any observed differences in health outcomes around the 18-year threshold can be attributed to the impact of entering the marriage market at the legal age of marriage, rather than confounding factors. Secondly, we employed placebo tests using two "fake cut-offs" to further validate our empirical strategy. In Table 2.B.12 and Table 2.B.13, we set artificial cut-offs at 17 years and 6 months, and 17 years and 9 months, respectively. We then conducted local randomization regression discontinuity analyses on the anthropometric outcomes using these placebo cut-offs. None of the anthropometric outcomes exhibited a significant treatment effect at these fake cut-off points. The absence of any significant impact on the outcomes when using artificial thresholds that do not correspond to the actual legal age at marriage provides strong evidence against the presence of confounding factors or underlying trends that could potentially drive our main results. By demonstrating that the observed effects are specific to the true cut-off of 18 years and do not manifest at other arbitrary age cut-offs, we reinforce the causal interpretation of our findings and the robustness of our empirical approach.

Finally, in Table 2.B.11, we performed a placebo analysis on a male sample to determine whether the observed treatment effects are simply an age-related phenomenon or specific to the legal age at marriage for females. This allows us to disentangle the impact of reaching the legal age at marriage from any potential age-specific factors that may influence the outcomes of interest. In this placebo test, we replicated our empirical strategy using a sample of males aged around 18 years. We implemented the regression discontinuity design where treatment group is defined as males who had turned 18 years old. We also used a 3 month window on each side of cut-off. We then examined the impact of this age threshold on the same set of anthropometric outcomes (BMI and BMI-for-age z-score) that we analyzed for the female sample. The results did not reveal any significant impact on the anthropometric measures of males around the age of 18. The lack of a discernible effect on the outcomes of interest for the male sample provides evidence that the improvements observed in our main analysis are indeed driven by the legal age at marriage for females, rather than general age-related factors.

2.7 Conclusion

The health, welfare and poverty levels of women are closely correlated with their marital status. Marriage before the age of 18 is considered a violation of the human right of self-determination. The implementation of legal age restrictions for marriage has emerged as a widely embraced approach in addressing the issue of child marriage. Due to the strict implementation of marriage age-related regulations, parents often defer their daughters' marriage until they reach the legally defined age. Thereby, match-making takes place at a legally defined age by the authority. While the legal age barrier delays marriage in many developing countries, societal norms continue to emphasize childbearing and household management as primary responsibilities after marriage. The prevalent belief holds that healthier and women who are not suffering from malnutrition are better equipped to perform these roles. As a result, healthier women are typically more sought-after in the marriage market.

Our study presents evidence for a significant improvement in general health and nutritional condition of women upon their entry into the marriage market at the age of 18. We hypothesize that this improvement in health is initiated by extra care or health investment induced by the entry to marriage market. Using a local randomization-based regression discontinuity design, we demonstrate that in India, upon a woman reaching the legal marriage age of 18 years and entering the marriage market, there is a significant boost in her anthropometric measurements and blood glucose indicators. Our further analysis reveals that these changes are the result of better quality food consumption and care. All these findings are in line with our hypothesis that there is a marriage market induced health investment in women, motivated by the desire to obtain a better match.

We made a unique contribution related to adolescent female health using a sample of young Indian women who just approached legal marriage age. Given the prevalence of malnutrition and early marriage in developing countries, our study specifically addresses women's nutritional and health condition at the period leading up to their participation in the marriage market.Moreover, our research has significant policy implications, particularly in the context of legislative modifications concerning the minimum age for marriage. When existing law or regulation undergoes changes, they inevitably influence the choices made within families and the behaviors observed within communities. Modifications to marriage laws have the potential to prolong the duration in which unmarried women live with their parents. This, in turn, might alter the trajectory of health investments, such as better care and nutrition may shift by the timing of entrance to the marriage market. Our analysis highlights the significance of considering health dynamics when formulating marriage age related laws. Although one might anticipate consistency in women's health outcomes throughout their lives, our objective is to discern how constant this health trajectory is. Moreover, we assume there are periods in a woman's life, especially previous years of entering marriage market, she might require additional support for their health needs. The findings indicate the potentials for government intervention programs designed to promote efficient health investment practices in society, thus reducing this discontinuity at the age of entry to the marriage market.

Appendix

2.A Supplementary analysis

Table 2.A.1: Change in Rohrer's index and arm circumference

	Rohrer's index		Rohrer's index-rural		Rohrer's index urban		Arm-circum all		Arm-circum rural		Arm-circum urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.213***	0.259*	0.257***	0.273*	0.054	0.142	0.118**	0.185*	0.128**	0.077	0.072	0.498**
	(0.072)	(0.145)	(0.082)	(0.158)	(0.153)	(0.328)	(0.049)	(0.102)	(0.054)	(0.113)	(0.109)	(0.233)
Age of household head		0.002		-0.002		0.015**		0.007***		0.004		0.016***
		(0.004)		(0.004)		(0.007)		(0.002)		(0.002)		(0.006)
Household size		-0.001		0.004		-0.006		-0.010		0.008		-0.066**
		(0.016)		(0.016)		(0.044)		(0.011)		(0.012)		(0.026)
Household head female		0.145		0.103		0.269		0.003		0.069		-0.197
		(0.116)		(0.124)		(0.276)		(0.067)		(0.074)		(0.150)
Household has: electricity		0.024		0.046		0.007		0.014		-0.006		0.755
		(0.119)		(0.125)		(0.460)		(0.128)		(0.131)		(0.537)
Wealth index combined		0.080***		0.045		0.150**		0.186***		0.179***		0.249***
		(0.027)		(0.034)		(0.064)		(0.021)		(0.026)		(0.052)
Age (in months)		-0.024		-0.017		-0.034		-0.021		0.019		-0.144**
		(0.046)		(0.052)		(0.095)		(0.030)		(0.033)		(0.068)
Mean BMI												
Mean BMI-for-age Z-score												
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11134	10875	8626	8405	2508	2470	11103	10845	8601	8381	2502	2464
R-Squared	0.001	0.018	0.001	0.015	0.000	0.047	0.001	0.052	0.001	0.055	0.000	0.065

Notes: This table presents the local randomization based regression results with 3 month bandwidth in each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on rohrer's index and arm circumference. The treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.A.2: Change in blood pressure level

Dep. var:	Systolic bp (log)						Diastolic bp (log)					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	-0.002	-0.007*	-0.001	-0.005	-0.003	-0.014*	-0.002	-0.004	-0.004*	-0.002	0.004	-0.009
	(0.002)	(0.004)	(0.002)	(0.004)	(0.004)	(0.008)	(0.002)	(0.004)	(0.002)	(0.005)	(0.004)	(0.009)
Age of household head		0.000		-0.000		0.000		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Household size		0.000		0.000		-0.001		-0.000		0.000		-0.000
		(0.000)		(0.000)		(0.001)		(0.000)		(0.001)		(0.001)
Household head female		-0.004		-0.003		-0.004		-0.004		-0.004		-0.001
		(0.003)		(0.003)		(0.005)		(0.003)		(0.003)		(0.006)
Household has electricity		-0.004		-0.005		-0.006		0.003		0.002		-0.012
		(0.005)		(0.005)		(0.022)		(0.005)		(0.005)		(0.025)
Wealth index combined		-0.006***		-0.004***		-0.007***		-0.001		0.000		-0.004*
		(0.001)		(0.001)		(0.002)		(0.001)		(0.001)		(0.002)
Age (in months)		0.002**		0.002		0.004		0.001		0.000		0.004
		(0.001)		(0.001)		(0.002)		(0.001)		(0.001)		(0.003)
Mean depvar	111.35	111.35	111.78	111.78	109.81	111.78	73.58	73.58	73.62	73.62	73.44	73.62
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11,099	10,841	8,599	8,379	2,500	2,462	11,099	10,841	8,599	8,379	2,500	2,462
R^2	0.000	0.072	0.000	0.062	0.000	0.098	0.000	0.040	0.000	0.040	0.000	0.066
Adjusted R^2	0.000	0.067	-0.000	0.056	-0.000	0.077	0.000	0.035	0.000	0.033	-0.000	0.044

Notes: This table presents the local randomization based regression results with 3 month bandwidth in each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on blood pressure level. Columns 1-6 represent the results of the log-transformed mean systolic blood pressure, while columns 7-12 display the log-transformed diastolic blood pressure. The treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.B Robustness

2.B.1 Robustness: local polynomial based RD analysis

Table 2.B.3: Change in BMI at age 18: Local polynomial regression with conventional and robust standard error

	(1) bmi
Conventional	0.363***
	[0.019]
Bias-corrected	0.414***
	[0.019]
Robust	0.414***
	[0.048]
Robust 95% CI	[.32 ; .507]
Kernel Type	Triangular
Observations	48,416
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in BMI at age 18. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.4: Change in BMI-for-age at age 18: Local polynomial regression estimates

	(1) zbfa
Conventional	0.108***
	[0.003]
Bias-corrected	0.125***
	[0.003]
Robust	0.125***
	[0.023]
Robust 95% CI	[.079 ; .171]
Kernel Type	Triangular
Observations	48,496
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in BMI-for-age z-score at age 18. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.5: Change in glucose level: Local polynomial regression with clustered se

	(1) blood_glucose
Conventional	1.037***
	[0.265]
Bias-corrected	1.017***
	[0.265]
Robust	1.017***
	[0.289]
Robust 95% CI	[.45 ; 1.584]
Kernel Type	Triangular
Observations	47,988
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in blood glucose level at age 18. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.6: Change in hemoglobin level: Local polynomial regression with clustered standard errors

	(1) hemoglobin_level
Conventional	-0.127***
	[0.019]
Bias-corrected	-0.130***
	[0.019]
Robust	-0.130***
	[0.027]
Robust 95% CI	[-.183 ; -.076]
Kernel Type	Triangular
Observations	47,925
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in blood hemoglobin level at age 18. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.7: Change in systolic bp: Local polynomial regression with clustered standard errors

	(1) log_sys
Conventional	-0.007***
	[0.001]
Bias-corrected	-0.008***
	[0.001]
Robust	-0.008***
	[0.001]
Robust 95% CI	[-.011 ; -.006]
Kernel Type	Triangular
Observations	48,340
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in systolic blood pressure level at age 18. The outcome variable is natural log of systolic blood pressure of the respondent. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.8: Change in diastolic bp: Local polynomial regression with clustered standard errors

	(1) log_dys
Conventional	-0.005***
	[0.001]
Bias-corrected	-0.004***
	[0.001]
Robust	-0.004***
	[0.001]
Robust 95% CI	[-.006 ; -.003]
Kernel Type	Triangular
Observations	48,337
Conventional p-value	0.00
Robust p-value	0.00
Order Loc. Poly. (p)	1.000
Order Bias (q)	2.000
Covariates	No
Cluster	Running variable-Age

Notes: This table presents results of regression discontinuity analysis using local polynomial regression, examining the change in diastolic blood pressure level at age 18. The outcome variable is natural log of diastolic blood pressure of the respondent. Coefficients for conventional, bias-corrected, and robust estimations are shown, with respective standard errors in parentheses. The analysis did not include additional covariates. Standard errors are robust and clustered by the running variable (age). The polynomial regression is of order 1, with a triangular kernel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.B.2 Mechanisms

Table 2.B.9: Mechanism: Daily meat and fish consumption

Dep. var:	Daily meat cons						Daily fish cons					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.001	0.008*	0.001	0.001	0.000	0.032***	0.002	0.013**	-0.001	0.010	0.009	0.024
	(0.002)	(0.004)	(0.002)	(0.004)	(0.006)	(0.011)	(0.003)	(0.007)	(0.004)	(0.007)	(0.009)	(0.016)
Age of household head		-0.000		-0.000		-0.000		0.000		0.000		0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Household size		0.000		-0.000		0.002		0.000		0.001		-0.002
		(0.000)		(0.000)		(0.002)		(0.001)		(0.001)		(0.002)
Household head female		0.000		-0.001		0.002		-0.004		-0.006		-0.002
		(0.003)		(0.003)		(0.009)		(0.004)		(0.004)		(0.012)
Household has electricity		-0.002		-0.004		0.016		0.004		0.007		-0.033
		(0.005)		(0.006)		(0.010)		(0.007)		(0.007)		(0.049)
Wealth index combined		0.004***		0.003***		0.004		0.002*		0.001		0.000
		(0.001)		(0.001)		(0.003)		(0.001)		(0.002)		(0.004)
Age (in months)		-0.003**		-0.001		-0.011***		-0.003*		-0.004*		-0.003
		(0.001)		(0.001)		(0.003)		(0.002)		(0.002)		(0.005)
Mean depvar	0.01	0.01	0.01	0.01	0.02	0.02	0.04	0.04	0.03	0.03	0.05	0.05
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11655.000	11384.000	8975.000	8746.000	2680.000	2637.000	11655.000	11384.000	8975.000	8746.000	2680.000	2637.000
R^2	0.000	0.019	0.000	0.016	0.000	0.042	0.000	0.200	0.000	0.190	0.000	0.248
Adjusted R^2	-0.000	0.014	-0.000	0.009	-0.000	0.022	-0.000	0.196	-0.000	0.184	0.000	0.232

Notes: This table presents the local randomization based regression results with 3 month bandwidth in each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on probability of eating daily meat and fish consumption. The dependent variables are dummy. treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.10: Mechanism: Health facility visit in last 3 months

Dep. var:	Health facility visit					
Sample:	All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)
treat	0.004	0.002	0.005	-0.008	-0.001	0.047
	(0.007)	(0.015)	(0.008)	(0.018)	(0.016)	(0.032)
Age of household head		0.000		0.001		-0.000
		(0.000)		(0.000)		(0.001)
Household size		-0.001		-0.002		0.000
		(0.002)		(0.002)		(0.004)
Household head female		0.010		0.010		0.009
		(0.010)		(0.012)		(0.022)
Household has electricity		0.033*		0.042**		-0.143
		(0.018)		(0.018)		(0.105)
Wealth index combined		0.011***		0.014***		-0.003
		(0.003)		(0.004)		(0.008)
Age (in months)		0.001		0.004		-0.013
		(0.005)		(0.005)		(0.010)
Mean depvar	0.18	0.18	0.18	0.18	0.21	0.21
State FE	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes
N	10,759	10,510	8,260	8,050	2,499	2,460
R^2	0.000	0.031	0.000	0.029	0.000	0.065
Adjusted R^2	-0.000	0.025	-0.000	0.022	-0.000	0.043

Notes: This table presents the local randomization based regression results with 3 month bandwidth in each side of age 18 for female. The table showing the effect of turning 18 years and entering marriage market on probability of visiting health facility in last three month. The Dependent variable is dummy. The treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.B.3 Placebo tests

Table 2.B.11: Change in male BMI at age 18

	BMI male	BMI-z male
	(1)	(2)
treat	-0.129	-0.084
	(0.151)	(0.060)
Mean BMI	20.01	
Mean BMI-for-age Z-score		-0.84
State FE	No	No
Controls	No	No
N	1565	1565
R-Squared	0.000	0.001

Notes: This table presents the local randomization based regression results with 3-month bandwidth on each side of placebo cut-off age 18 years for a male. It shows the effect of turning 18 on BMI and BMI-for-age z-score. The treat variable indicates whether individual is already at this age. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.12: Placebo cut-off: Change in female bmi at 17 years and 6 month cut-off with 3 months bandwidth

Dep. var:	BMI						BMI-for-age- z-score					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$treat17_5$	-0.057	-0.021	-0.039	-0.028	-0.088	0.022	-0.024	-0.015	-0.023	-0.024	-0.019	0.015
	(0.059)	(0.059)	(0.063)	(0.063)	(0.145)	(0.147)	(0.021)	(0.021)	(0.023)	(0.023)	(0.049)	(0.049)
Age of household head		0.004		0.003		0.007		0.001		0.001		0.003
		(0.003)		(0.003)		(0.007)		(0.001)		(0.001)		(0.002)
Household size		-0.032**		-0.033**		-0.025		-0.009*		-0.009*		-0.007
		(0.015)		(0.016)		(0.035)		(0.005)		(0.005)		(0.012)
Household head female		-0.046		-0.053		0.030		-0.017		-0.011		-0.019
		(0.083)		(0.090)		(0.199)		(0.030)		(0.033)		(0.067)
Household has electricity		-0.192		-0.142		-0.808		-0.053		-0.029		-0.370*
		(0.163)		(0.171)		(0.571)		(0.057)		(0.060)		(0.204)
Wealth index combined		0.265***		0.234***		0.323***		0.083***		0.073***		0.101***
		(0.027)		(0.033)		(0.076)		(0.010)		(0.012)		(0.026)
Education in single years		-0.038***		-0.042***		-0.005		-0.014***		-0.016***		0.003
		(0.012)		(0.012)		(0.040)		(0.004)		(0.004)		(0.013)
Mean depvar												
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11,191	10,972	8,689	8,501	2,502	2,469	11,214	10,995	8,708	8,520	2,506	2,473
R^2	0.000	0.048	0.000	0.048	0.000	0.055	0.000	0.050	0.000	0.052	0.000	0.057
Adjusted R^2	-0.000	0.042	-0.000	0.041	-0.000	0.033	0.000	0.045	0.000	0.046	-0.000	0.035

Notes: This table presents the local randomization based regression results with 3-month bandwidth on each side of placebo cut-off age 17 years and 6 months. It shows the effect of turning 17 years and 6 months on BMI and BMI-for-age z-score for a female. The treat_5 variable indicates whether individual is already at this age. Rural and urban implies the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.13: Placebo cut-off: Change in female BMI at 17 years and 9 month cut-off with 3 months bandwidth

Dep. var:	BMI						BMI-for-age- z-score					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$treat17_9$	0.025	0.059	-0.029	0.014	0.210	0.194	-0.011	-0.000	-0.026	-0.012	0.041	0.032
	(0.056)	(0.056)	(0.061)	(0.061)	(0.139)	(0.137)	(0.020)	(0.020)	(0.022)	(0.022)	(0.047)	(0.047)
Age of household head		0.006**		0.003		0.016**		0.002**		0.001		0.005**
		(0.003)		(0.003)		(0.007)		(0.001)		(0.001)		(0.002)
Household size		-0.035***		-0.024		-0.056*		-0.010**		-0.008		-0.015
		(0.013)		(0.015)		(0.031)		(0.005)		(0.005)		(0.011)
Household head female		0.057		0.091		-0.038		0.022		0.039		-0.031
		(0.080)		(0.088)		(0.192)		(0.029)		(0.032)		(0.067)
Household has electricity		0.008		0.092		-0.490		-0.000		0.025		-0.121
		(0.128)		(0.130)		(0.699)		(0.050)		(0.051)		(0.247)
Wealth index combined		0.272***		0.191***		0.464***		0.088***		0.063***		0.142***
		(0.027)		(0.032)		(0.067)		(0.009)		(0.011)		(0.023)
Education in single years		-0.045***		-0.047***		-0.029		-0.015***		-0.016***		-0.007
		(0.011)		(0.012)		(0.032)		(0.004)		(0.004)		(0.011)
Mean depvar												
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	11,970	11,741	9,359	9,162	2,611	2,579	11,991	11,762	9,375	9,178	2,616	2,584
R^2	0.000	0.048	0.000	0.043	0.001	0.072	0.000	0.048	0.000	0.047	0.000	0.062
Adjusted R^2	-0.000	0.043	-0.000	0.037	0.000	0.052	-0.000	0.043	0.000	0.041	-0.000	0.041

Notes: This table presents the local randomization based regression results with 3-month bandwidth on each side of placebo cut-off age 17 years and 9 months for female. It shows the effect of turning 17 years and 9 months on BMI and BMI-for-age z-score. The treat_9 variable indicates whether individual is already at this age. Rural and urban implies the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.B.4 Additional results: Heterogeneity analysis

Results for only unmarried sample

Table 2.B.14: Change in body mass index for unmarried sample

Dep. var:	BMI						BMI-for-age- z-score					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.177***	0.411***	0.188***	0.377***	0.097	0.458	0.049**	0.123***	0.050**	0.117**	0.036	0.125
	(0.060)	(0.127)	(0.065)	(0.136)	(0.144)	(0.301)	(0.021)	(0.045)	(0.024)	(0.049)	(0.049)	(0.101)
Age of household head		0.007**		0.004		0.018**		0.002**		0.002		0.005**
		(0.003)		(0.003)		(0.008)		(0.001)		(0.001)		(0.003)
Household size		-0.027*		-0.018		-0.039		-0.007		-0.006		-0.006
		(0.014)		(0.016)		(0.034)		(0.005)		(0.006)		(0.012)
Household head female		0.167**		0.125		0.264		0.060*		0.044		0.104
		(0.084)		(0.090)		(0.206)		(0.031)		(0.033)		(0.073)
Household has electricity		-0.240		-0.236		0.073		-0.088*		-0.090*		0.056
		(0.153)		(0.157)		(0.709)		(0.053)		(0.054)		(0.250)
Wealth index combined		0.242***		0.195***		0.265***		0.075***		0.063***		0.082***
		(0.026)		(0.032)		(0.067)		(0.009)		(0.011)		(0.023)
age_{frac}		-1.134**		-0.949**		-1.547		-0.354**		-0.329*		-0.395
		(0.442)		(0.479)		(1.037)		(0.158)		(0.176)		(0.352)
Mean depvar												
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	10,159	9,997	7,764	7,632	2,395	2,365	10,174	10,012	7,776	7,644	2,398	2,368
R^2	0.001	0.045	0.001	0.039	0.000	0.065	0.001	0.046	0.001	0.042	0.000	0.062
Adjusted R^2	0.001	0.039	0.001	0.031	-0.000	0.042	0.000	0.040	0.000	0.035	-0.000	0.040

Notes: This table presents the local randomization based regression results with 3-month bandwidth on each side of cut-off age 18 for female. It shows the effect of turning 18 years and entering marriage market on BMI and BMI-for-age z-score. The treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.15: Change in rohrer's index and arm circumference for unmarried sample

Dep. var:	Rohrer index						Arm circumference					
Sample:	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.198***	0.201	0.226***	0.175	0.087	0.218	0.140***	0.228**	0.140**	0.116	0.115	0.538**
	(0.075)	(0.144)	(0.086)	(0.153)	(0.159)	(0.329)	(0.051)	(0.107)	(0.057)	(0.118)	(0.112)	(0.237)
Age of household head		0.002		-0.001		0.015**		0.006**		0.003		0.015**
		(0.003)		(0.004)		(0.007)		(0.002)		(0.003)		(0.006)
Household size		0.001		0.009		-0.006		-0.007		0.015		-0.070***
		(0.017)		(0.017)		(0.047)		(0.012)		(0.013)		(0.027)
Household head female		0.176		0.142		0.275		0.012		0.093		-0.237
		(0.123)		(0.131)		(0.290)		(0.069)		(0.077)		(0.152)
Household has electricity		-0.026		-0.003		0.007		0.022		-0.004		0.737
		(0.125)		(0.131)		(0.462)		(0.133)		(0.137)		(0.541)
Wealth index combined		0.079***		0.031		0.157**		0.188***		0.179***		0.252***
		(0.028)		(0.033)		(0.067)		(0.022)		(0.027)		(0.053)
age_frac		-0.149		0.040		-0.531		-0.412		0.068		-1.796**
		(0.574)		(0.650)		(1.134)		(0.375)		(0.417)		(0.833)
Mean depvar												
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	10,174	10,012	7,776	7,644	2,398	2,368	10,145	9,983	7,754	7,622	2,391	2,361
R^2	0.001	0.021	0.001	0.017	0.000	0.049	0.001	0.051	0.001	0.056	0.000	0.065
Adjusted R^2	0.001	0.015	0.001	0.010	-0.000	0.026	0.001	0.046	0.001	0.049	0.000	0.042

Notes: This table presents the local randomization based regression results with 3 month bandwidth on each side of age 18 for female. It shows the effect of turning 18 years and entering marriage market on rohrer's index and arm circumference. treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.B.16: Change in Blood markers for unmarried sample

Dep. var:	Glucose level (log)						Hemoglobin level (log)					
Sample	All		Rural		Urban		All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
treat	0.009***	0.008	0.011***	0.008	0.005	0.005	-0.006*	-0.009	-0.006*	-0.007	-0.005	-0.018
	(0.003)	(0.006)	(0.003)	(0.007)	(0.006)	(0.012)	(0.003)	(0.006)	(0.004)	(0.007)	(0.006)	(0.013)
Age of household head		0.000		0.000		0.000		0.000		-0.000		0.001*
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Household size		0.000		0.000		-0.000		0.000		0.000		0.000
		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)		(0.001)
Household head female		0.006		0.005		0.006		-0.003		-0.004		-0.005
		(0.004)		(0.005)		(0.008)		(0.004)		(0.005)		(0.009)
Household has electricity		0.009		0.009		-0.003		-0.001		0.000		0.013
		(0.008)		(0.009)		(0.032)		(0.008)		(0.009)		(0.032)
Wealth index combined		-0.001		0.002		-0.006**		0.007***		0.006***		0.006*
		(0.001)		(0.002)		(0.003)		(0.001)		(0.002)		(0.003)
age_frac		-0.010		-0.004		-0.012		0.023		0.016		0.060
		(0.022)		(0.026)		(0.043)		(0.022)		(0.026)		(0.045)
Mean depvar	102.27	102.27	102.27	102.27	102.27	102.27	11.45	11.45	11.45	11.45	11.45	11.45
State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year x Month FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	10,073	9,912	7,712	7,581	2,361	2,331	10,065	9,904	7,706	7,575	2,359	2,329
R^2	0.001	0.064	0.001	0.064	0.000	0.092	0.000	0.044	0.000	0.051	0.000	0.055
Adjusted R^2	0.001	0.059	0.001	0.057	-0.000	0.069	0.000	0.039	0.000	0.044	-0.000	0.032

Notes: This table presents the local randomization based regression results with 3 month bandwidth on each side of age 18 for female. The table showing the effect of turning 18 years and entering marriage market on blood glucose and and blood hemoglobin level. Treat variable indicates whether individual is already 18 years old. Rural and urban refer to the subsample from rural and urban region. The even-numbered columns incorporate state fixed effect, time fixed effect, and control variables. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 3

Addressing Antimicrobial Resistance and Self-Medication Practice in Developing Countries: A Field Experiment in Bangladesh*

Joint with: Daniel Wiesen † and Lorenz Goette ‡

Abstract

Antimicrobial resistance is a global public health concern. We investigate how an information intervention about the use of safe medicines can reduce unnecessary antibiotic uptake. We conducted a clustered randomized control trial in rural Bangladesh with evermarried female respondents. While the treatment group got information regarding antimicrobial resistance, they also randomly received color-coded drug boxes as visual cue or a monthly subsidy for qualified doctor visits. After two months of intervention, we find that providing information along with visual cue or subsidy increases registered doctor visit by 0.37 relative to the control group. Our finding implies that an innovative information intervention with causal selections can reduce channels of unnecessary antibiotic consumption.

*This project was supported by Center for Social and Economic Behavior (C-SEB), University of Cologne through C-SEB junior start-up grant (Rd14-2022-JSUG-Imran) and C-SEB seedcorn grant (Rd18-2024-SG-Wiesen). We are grateful for valuable comments and suggestions from Ingo Isphording, Joshua McCormick, Akib Khan, and seminar and conference participants at C-SEB early idea workshop, 6th FHM+ Development Economics Workshop.

†Department of Healthcare Management and Center for Social and Economic Behavior (C-SEB) University of Cologne and Erasmus School of Health Policy and Management (ESHPM), Erasmus University of Rotterdam

‡Department of Economics, National University of Singapore

3.1 Introduction

Antimicrobial resistance (AMR) is a global public health threat. It renders antibiotic drugs ineffective in the human body due to the emergence of resistance mechanisms by microorganisms.[1] Murray et al. (2022) estimated that the global number of deaths caused by bacterial antimicrobial resistance in 2019 was 1.27 million. AMR has emerged as a global concern within the realm of public health and is sometimes characterized as a slow-motion pandemic since it results in a large number of deaths. In near future, lower and middle-income countries are likely to bear the greatest direct and indirect burden of AMR (O'Neill, 2016; Sulis et al., 2022). The development of antimicrobial resistance is mostly attributed to the improper usage and excessive application of antimicrobial agents (Prestinaci et al., 2015). The general population of Bangladesh lacks sufficient knowledge about AMR (Azim et al., 2023).

In order to avoid the improper use of medications, the laws in Bangladesh impose restrictions on the sale of antibiotics without a prescription. However, due to inadequate enforcement, it is still possible to obtain antibiotics over the counter (Ahmed and Islam, 2012; Ahmed et al., 2017). In addition, self-medication, drug shop salespersons, and quacks play a significant role as health service providers in rural regions, despite their little understanding of antimicrobial resistance (Ahmed and Hossain, 2007; Akhtar et al., 2021).[2] The limited understanding and lack of awareness of the negative health consequences of inappropriate antibiotic use contribute to the increasing prevalence of antimicrobial resistance in society (Dadgostar, 2019).

Communities where people frequently use non-prescription drugs have a high prevalence of antimicrobial-resistant bacteria (Morgan et al., 2011). Bangladesh is one of those countries among which non-prescription antibiotic usage is very high (Tangcharoensathien et al., 2018). Azim et al. (2023) examined attitudes, knowledge and practices on antibiotic use in Bangladesh, finding that the general public does not have adequate knowledge about AMR. Bepari et al. (2023) indicated unqualified village medical practitioners and drugstore sales person have moderate to poor knowledge about the use of antibiotic drugs. Apparently, the fight against AMR is a supply-side and demand-side issue in Bangladesh, where the main challenge is the

[1] Microorganisms comprise of microscopic or submicroscopic organisms, including viruses, bacteria, fungi, algae, and small protozoa (Chanda and Joshi, 2022).

[2] Quacks are also known as "Kobiraj" or "Village Doctors", they operate widely in Bangladesh, particularly in rural and urban underserved areas. These individuals provide medical treatments without the formal qualifications or registration required by the government.

massive information gap on the antimicrobial resistance mechanism. Many cherry-picked drugs are already ineffective in Bangladesh as a consequence of antimicrobial resistance (Ahmed et al., 2019). To improve the situation, there is no alternative to restricting unnecessary consumption and misuse of antibiotics among the general population. As frequent visits to unlicensed healthcare professionals and individual self-medication practices are the primary contributors to unnecessary antibiotic use in rural Bangladesh, focus on these two behaviors.

We investigate how an information intervention about the risk of antimicrobial resistance can restrict the channels of unnecessary antibiotic uptake, especially in areas where antibiotic drugs are easily accessible. In addition, it also attempts to identify the reasoning behind the recurring visit to unqualified medical professionals who prescribe antibiotics more frequently even though government facilities provide service at negligible cost. Despite efforts by the government and other organizations to raise awareness about the adverse consequences of AMR, a significant portion of the rural population in Bangladesh remains in danger of AMR.

With our research question, we analyze whether information intervention strategies targeting rural women (i) with a behavioral nudge (separate color-coded boxes for safe and unsafe medicines) or (ii) with a transport subsidy reduce visits to unqualified healthcare providers and increase visits to qualified doctors. In other words, our research question investigates how we can successfully address unnecessary antibiotic consumption, which is the main reason for AMR.

Unlike developed countries, few studies focus on the public choice and decision-making processes of healthcare services in developing nations such as Bangladesh. Here, the healthcare system operates under a different structure from that of developed counties, specifically with high out-of-pocket expenses (Molla and Chi, 2017; Sarker et al., 2021) and a limited prevalence of the health insurance scheme (Adams et al., 2013). Primary health care service providers comprises of government-operated health institutions, non-profit organizations, and for-profit hospitals. Physicians and healthcare service providers from these institutions are adequately educated to provide quality treatment. In addition to these, a significant number of people, particularly in the rural region, rely on healthcare practitioners who lack formal qualifications, such as village doctors or quacks, homeopath practitioners, and pharmacy sales staff. The quality of qualified physicians are undoubtedly better than that of unqualified practitioners. In our research, we explore the determinants of health-seeking behavior of individuals through a field experiment where we randomly provide information about the consequences of AMR and provide transport cost to qualified doctors.

Information experiments are common in various domains of economics, including the field of health. By providing information, the choices of the economic agent are studied by manipulating their beliefs and constraints, allowing the addressing of questions relevant to different policies (Haaland et al., 2023). In healthcare sectors, information provision experiments are very useful tool to understand the choice of health insurance plans (e.g., Bundorf et al., 2019; Bhargava et al., 2017). In recent years, especially since the Covid-19 outbreak, the utilization of information provision experiments has increased to understand behaviors and preferences related to public health. Our study, which provides a public health message regarding the risk of AMR, builds on the finding that public health messages were effective in promoting a good understanding of Covid-19 safety measures in Italy (Barari et al., 2020). Previously, the impact of a house to house information campaign on arsenic contamination in Bangladesh was causally documented by Madajewicz et al. (2007) which showed that information can succesfully help to change behavior of people who face a health risk.

In addition to information provision, the significant impact of nudges on shaping behaviors and practices related to healthcare is widely acknowledged (e.g., Thaler and Sunstein, 2009; Rice, 2013). Nudging leverages bounded rationality and can discretely drive individuals toward healthy behaviors without imposing restrictions on their choices. This is achieved through the use of intuitive design along with simple visual cues that are in line with natural cognitive processes. Recognizing the constraints of human decision-making ability, the use of nudge can improve health outcomes by simplifying different health-related choices. Blackwell et al. (2017) have demonstrated that the use of visual cues, such as arrows pointing towards sinks, can enhance public health behaviors, such as hand washing. In the same vein, Prasetyo et al. (2022) provided evidence that implementing posters and visible footprints as nudges in Indonesian government offices significantly increased hand washing practices. Another commonly used visual cue is the health warnings displayed on the cigarette packages. The purpose of these warnings is to underscore the dangers linked to smoking, with a design aimed at heightening public awareness and deterring nicotine use. The effectiveness of tobacco warning labels is enhanced by graphic pictorial images (Fong et al., 2009).

In developing countries, it is common to provide health services and products either free of charge or at a minimal cost. However, the results of these free health product provisions are varied. Dupas and Miguel (2017) conducted a randomized controlled trial in Kenya that offers a valuable understanding regarding the dynamics of health product distribution, demonstrating the varied efficacy of voucher-based systems compared to cost-sharing and free offering in encouraging actual product usage. They found adding a non-financial cost—like the ability to redeem vouchers—can greatly improve the targeting effectiveness of health interventions and guarantee that the benefits are distributed to people who actually use the services. Similarly, government hospitals in Bangladesh offer services for a nominal fee (around 5 cents for a ticket and complimentary medication). Although hospital services are almost free, public health research regarding utilization of hospital service often overlooks the expense associated with transportation. In our study, we considered transportation expenses to the qualified physician, which is a potential determinant in choosing a healthcare provider or self-medication.

In our paper, we combined information provision intervention with nudging or transportation expenditure subsidies for visiting qualified physicians to tackle the issue of unnecessary antibiotic usage by the general public. An unbiased assessment of any form of intervention is best achieved through a randomized control trial (RCT), which is considered as the gold standard of evaluation. Miguel and Kremer (2004) is one of the early randomized interventions in the public health domain that examined the impact of deworming drugs on school absentinism of school-age children in Kenya. Combining information module with other interventions in the field experiment is common in case of developing country setting. In a field experiment on improving preventive health behaviors in rural Kenyan women, John and Orkin (2022) investigated the effectiveness of visualization and planning treatments, which were supplemented with an information module on the advantages of water chlorination. They found that the intervention had long-lasting favorable effects by considerable increase in household water chlorination, reduction in child diarrhea episodes and improved savings practices. We integrated a simple information experiment with a nudge-based intervention, namely a visual cue and a subsidy intervention in transport costs spent on registered doctors. With the risk information regarding consequence of AMR, one of the respondent groups received two different boxes with red and blue lids to organize their daily medications. We ask them to keep all drugs prescribed by registered physicians in the blue lid box, and drugs purchased from other sources (e.g., self-purchased, provided by quacks, etc.) in red lid boxes. The other intervention group receives same information along with subsidies in transport costs that are spent on medical facilities to visit registered doctor's. They receive up to 100 bdt (equivalent to approximately 80 cents in June 2024) if they choose to visit registered doctor's facility for healthcare purposes. Besides these two groups, we have a pure control group that does not receive any information regarding AMR, nudge, or subsidy. According to Kremer et al. (2019), most information experiments in developing countries mainly examine changes in health behavior rather than beliefs. In light of this, Haaland et al. (2023) proposes that eliciting beliefs could provide valuable insights into why certain information interventions are more effective than others in promoting behavior change. To our knowledge, our study is the first attempt to gather insights on beliefs and observe behavioral changes in the context of medication usage and available healthcare service choices in the community. Employing the same information module on AMR risk with nudge and subsidy interventions separately enables us to perform a precise evaluation of each scheme on healthcare behavior. This extends our understanding of how information influences decision-making in healthcare provider choice in conjunction with economic and psychological incentives.

Government medical facilities and private clinics where qualified physicians are available are located mainly in urban areas.[3] However, unqualified medical professionals live very close to the villagers. Accessing a qualified medical service provider incurs a transportation expense, whereas financially challenged villagers can consult with unqualified practitioners without any transportation costs. The transport cost subsidy is anticipated to alleviate the financial burden of accessing qualified medical professionals, resulting in a decrease in perceived loss from out-of-pocket expenses. This reduction in out-of-pocket expense is also expected to re-adjust the reference point. With the updated belief based on the information provided on AMR and the rearranged reference point from the subsidy, the action on the choice of the healthcare provider is expected to move towards a more safe option considering the long-term health effect.

On the other hand, we used boxes as visual cues to distinguish drugs from qualified and unqualified sources, aiming to nudge respondents towards safer medication behavior. The boxes are expected to influence their prospective memory.[4] Furthermore, our provided medicine storage boxes are color coded, with a blue lid and a

[3]Union subcenters and community clinics are available in rural regions but only provides limited and primary medical services, specialist doctors are also not available in these places.

[4]Prospective memory refers to the cognitive ability to recall and execute planned actions at a predetermined period in the future. It is considered a type of memory that encompasses the strategic planning and implementation of future actions; see Brandimonte et al. (1996) and McDaniel and Einstein (2007).

red lid, which is also expected to enhance the sailence of information. Bordalo et al. (2022) broadly discusses salience and economic behavior noting that salient stimuli capture human attention due to high contrast, startling character, or prominence. In our design, we have used red and blue color-coded boxes to automatically attract the attention of respondents when they encounter them in their daily lives during medication. This boxes are expected to work as salient stimuli using a bottom-up approach.[5]

To conduct our experiment we collaborated with a local NGO and initiated a free health and nutritional service provision program named "Kemon Achen". Under this program, women health workers, whom we call guides, visited selected villages in two sub-districts of Dinajpur district. They approached ever-married women from randomly chosen households to participate in the program. The program offered a monthly free service such as height, weight, BMI measurement, free counseling according to their nutritional condition, regular blood pressure checkup, they received oral saline, free handwashing soaps, etc. We finally recruited 1,129 women from 62 villages and completed our baseline survey in October 2023. Women were chosen because they play a crucial role in household health management, particularly in overseeing child health and can effectively convey health-related messages to other household members. In addition, it is often observed that women could potentially derive significant advantages from receiving enhanced health education. The "Kemon Achen" program is designed with the objective of enhancing the health and nutritional status of rural people. This is achieved through the provision of complimentary monthly anthropometric measurements, educational resources, and information on a range of health-related issues. In general, our goal is to reduce unnecessary antibiotic consumption in the community through innovative interventions targeting women, which in turn has the potential to benefit the entire household.

In November 2023, we completed our baseline survey, collecting comprehensive data, with a primary focus on household demography, medical history, medicine-related knowledge, choice of healthcare providers, and geo-coordinates of each household, etc. Given the potential for spillover and contamination in our information experiment, we opted for cluster randomization. To ensure robust estimation, we divided the study area into 700x700 meter (0.49 square kilometer) grid clusters.[6] We

[5]The bottom-up approach captures attention through its unique and striking features, provoking immediate and sensory-based responses, neutral from previous knowledge or expectations(Bordalo et al., 2022).

[6]Clusters containing less than five individuals were combined with the closest cluster.

then pair-matched the grids based on the average number of households, average age of members and average visits to government health facilities and unqualified healthcare professionals over the last six months. One grid from each pair was randomly selected as the control and the other as a treatment cluster. In the treatment grids, half of the members randomly received boxes, while the other half received a subsidy of up to 100 BDT (approximately 80 cents) for transportation if they visited and presented prescriptions in each month. All members of the treatment group (box or subsidy recipient) received similar information regarding the risks of AMR.

In January 2024, the main experiment was initiated and the both treatment was implemented. After two months (March 2024), data on different medical facility visits of the participants was collected, which we use in this paper. Firstly, using baseline data, we examine the relationship between proximity to government hospitals and the number of visits to different healthcare facilities. The selection of a health facility has a direct influence on the risk of AMR due to its association with self-medication and dependence on unqualified medical practitioners, as these factors are the main drivers of unnecessary antibiotic usage. To account for the presence of spatial autocorrelation in our data, we used both the spatial error model (SEM) and the spatial autoregressive model (SAR) in our analysis. Our results show that the farther away someone lives from the closest Upazila Health Complex (UHC) or Medical College Hospital, the more often they visit unlicensed doctors, also known as "quacks," over the course of six months.[7] Each additional mile from the residence to the closest Upazila Health Complex or Medical College Hospital was associated with a 0.38 increase in visits to quacks in the OLS model, a 0.24 increase in the SAR model, and a 0.37 increase in the SEM model over the past six months (all $p<0.01$). Conversely, the distance to the nearest government facility exhibits a negative, but non-significant, effect on the visits to these government health establishments.

The result shows that our interventions have a promising impact on increasing the number of visits by qualified healthcare providers after two months. The treated individuals demonstrated a significant increase in visits to qualified healthcare professionals, with 0.37 more visits ($p <0.01$) compared to the control group. The impact varied based on the specific intervention type. Participants who received visual cues (colored boxes) along with risk information about AMR showed the highest increase in visits to qualified physicians at government facilities, with 0.39 more visits ($p<0.01$) than the control group. Those who received a travel subsidy alongside risk informa-

[7]Upazila Health Complex (UHC) are primary government hospitals at sub-dristric level.

tion also increased their visits to qualified professionals in government facilities by 0.27 ($p<0.01$) compared to the control group. Notably, neither intervention significantly reduced visits to quacks during this two-month period.

We contribute to the economics literature in several ways. Firstly, we are the first to implement a randomized control trial in a least-developed country that targets household women to reduce antibiotic consumption. In rural India, Das et al. (2017) conducted a randomized experiment targeting informal healthcare providers, where their intervention (9 month long training) successfully increased correct case management, but did not reduce unnecessary drugs and antibiotics prescription. We add to this discussion by providing information on negative consequence of AMR to the consumers, instead of educating healthcare providers. Rodgers et al. (1999) evaluates Stroke Education Program (SEP) for stroke patients and their informal carers which found higher stroke knowledge for treated group. Lovell et al. (2010) conducted rct with standardized educational intervention (video and booklet) for patients with cancer pain which found significant improvements in pain scores and reduced barriers to pain management. Although educating patient is relatively common in academia and has been found to have a considerable impact, we provide a benchmark that represents how educating an individual instead of healthcare providers can reduce unsafe medication usage, thus reducing the risk of AMR.

Secondly, we contribute to the literature discussing the effect of risk information on health behavior. Dupas (2011) is the leading literature on risk information in the context of the least developing countries. They conducted a randomized experiment on information provision about the relative risk of HIV infection based on partner's age, resulting in a 28 percent reduction in teenage pregnancies in rural Kenya. Similarly, in an RCT setting, Ciancio et al. (2024) showed that how informing individuals about population-level mortality risk in a high-HIV-prevalence environment leads to safer sexual practices. Chinkhumba et al. (2014) found that both price and information are important determinant of adult medical male circumcision as HIV preventive intervention. We discuss how providing information on risk from AMR influences safe medicine behavior in rural Bangladesh. As previous literature claims that information intervention has mixed effect, we combined risk information with two other widely used policy tools-visual cue and subsidy.

Thirdly, we contribute to the literature on the use of behavioral interventions to shape healthcare behavior. We utilize visual cues as nudges to promote safe medication practices in healthcare. Visual cues have been found to have considerable impact to shape hand-washing behavior by (e.g., Blackwell et al., 2017; Nevo et al., 2010; Prasetyo et al., 2022). Some research applied behavioral interventions to reduce inappropriate antibiotic prescription by healthcare professionals. Meeker et al. (2016) showed that nudges, such as suggestions for non-antibiotic treatments, the requirement for accountable justifications, and peer comparison emails were effective in decreasing the rate of improper antibiotic prescription among primary care practices in the United States. In the context of the United Kingdom Hallsworth et al. (2016) targeted both general practitioners (GP) and patients to reduce antibiotic usage. It showed that a clinician-focused intervention (feedback letter from England's chief medical officer on the high antibiotic prescription rate) effectively reduced antibiotic dispensing among general practitioners, whereas a patient-focused intervention (display of posters and leaflets in the patient waiting areas and GPs consulting rooms) did not significantly alter prescribing patterns, highlighting the difficulties in changing patient behaviors through education alone. The previous intervention mostly targeted GPS, but we targeted the general population.

Finally, to our knowledge, we are the first to conduct a causal study to understand the driving force of antibiotic consumption at the micro-level. There are several other observational and quasi-experimental studies that discuss the determinants of antibiotic consumption at the macro level; for example Koya et al. (2024) explores the relationship between antibiotic consumption and macro-level indicators such as gross domestic product, girls' enrollment rate, and government spending on health in India. Similarly, Klein et al. (2018) reports the causes of increased antibiotic consumption at the global level using data from 76 countries and distinguishes the difference in antibiotic consumption between high-income versus lower- or middle-income countries. Using data from 16 European countries Blommaert et al. (2014) demonstrates that increased overall antibiotic usage correlates with factors such as relative humidity, healthcare expenditure as a percentage of GDP, distrust sentiments, the percentage of the population aged 65 and older, and the presence of treatment guidelines. We provide evidence that removing financial barrier to access healthcare service can successfully reduce the channel of unnecessary antibiotic consumption.

The result of this paper is based on the baseline survey and the evaluation of two months. The endline survey is planned in August 2024. The rest of the paper is organized as followings: Section 3.2 gives a detailed background and context of the study including overview of the Bangladeshi health sector, Section 3.3 provides research design, Section 3.4 discusses data hypothesis and empirical approach, Section

3.5 presents detailed results and Section 3.6 is the policy implication, limitation, and conclusion of the study.

3.2 Background

Overview of health sector in Bangladesh

Bangladesh is a least developed country with a significantly large population with a very high density.[8] Providing healthcare for everyone is a huge challenge for the nation. The health system in Bangladesh is shaped and operated by four main contributors: the government, the private sector, NGOs, and donor agencies, where the government plays the main role in the provision of healthcare for the general public (World Health Organization, 2015). Section 3.A.1 displays government healthcare system hierarchy in Bangladesh.

In addition to these four key actors, informal healthcare practitioners also provides health care services to a large number of people, especially in rural areas.[9] The healthcare system in Bangladesh faces significant difficulties, such as a nearly non-existent health insurance system, high out-of-pocket medical expenses of patients, a lack of public health facilities and a shortage of healthcare workers (Islam and Biswas, 2014). In this section, we discuss some challenges faced by Bangladesh's health sector which are very crucial to address in the fight against AMR. In general, AMR will put more burden on Bangladeshi population if these challenges are not addressed properly.

High prevalence of AMR

According to multiple studies, the level of antimicrobial resistance in Bangladesh is extremely concerning. In a systematic review, Ahmed et al. (2019) reported that many front-line medications have already largely lost their effectiveness against pathogens in Bangladesh. Jain et al. (2021) showed an alarming 98% prevalence of multi-drug resistance (MDR) among clinical E. coli isolates in Bangladesh. Another review article by Hoque et al. (2020) reports that the development and transmission of AMR in

[8] More than 160 million people with 1,119 people per square kilometers (Bangladesh Bureau of Statistics, 2022).

[9] Informal practitioners consist of individuals who do not have formal institutional education or a specific curriculum, directly receive payments from patients, usually function without being registered or monitored by government or other official bodies, and if they are part of professional associations, these organizations are mainly involved in networking and commercial activities with minimal self-regulation (Sudhinaraset et al., 2013).

Bangladesh is deep-rooted in irrational use of antimicrobials, supply-side issues such as poor regulatory controls and demand-side problems like self-medication and non-compliance. Bonna et al. (2022) highlights that antibiotic resistance in Bangladesh is aggravated by poor healthcare practices, improper and excessive use of antibiotics, and insufficient regulatory measures.

Multidimensional disparity in healthcare

Healthcare services in Bangladesh exhibit multiple dimensions of disparity, spanning from rural-urban differences to private-public inequities. There is a substantial shortage of qualified healthcare professionals in Bangladesh, with the majority of these professionals being concentrated in urban areas (Ahmed et al., 2011). Rural-urban disparity in healthcare facilities is a major concern. The urban areas of Bangladesh are strongly biased towards the availability of qualified healthcare professionals (doctors, nurses, and dentists), as indicated by Ahmed et al. (2013). They reported that urban areas have 18.2 physicians, 5.8 nurses, and 0.8 dentists for every 10,000 inhabitants, whereas rural regions have only 1.1 physicians, 0.8 nurses, and 0.08 dentists per 10,000 inhabitants. Moreover, rural areas exhibit a greater number of traditional and community-based health providers than urban areas. Specifically, there are 42.2 traditional birth attendants and 13.8 village doctors per 10,000 people in rural areas, in contrast to urban areas, which have only 6.0 traditional birth attendants and 8.8 village doctors per 10,000 people. The availability of unqualified healthcare providers carelessly prescribes antibiotic drug which increases the danger of AMR more among rural population. Andaleeb et al. (2007) found that patient satisfaction is higher in private healthcare facilities in terms of reliability, responsibility, tangibility, professionalism, waiting time, and environment. There is also a huge disparity in the utilization of public and private healthcare facilities by poor and rich, where poor people go primarily to public facilities (Mannan, 2013). Each types of disparities poses a substantial threat to fight against AMR.

Low health literacy

Joarder et al. (2019) underscores the importance of patient education as a key strategy to overcome the challenges to universal health coverage in Bangladesh, pointing out its ability to improve community involvement and confidence in health systems. Our study aims to educate the rural population, especially women, to make them understand the consequences of AMR so that they do not further fall into additional burden. Das et al. (2017) discovered that in rural Bangladesh, 86% of the surveyed individuals identified quack as the most familiar health facility, emphasizing a substantial lack of awareness about the use of formal healthcare services and accentuating the dependence on uncertified practitioners for medical care.

High Out-of-pocket expenditure

Out-of-pocket expenditure (OOP) continues to be the primary means of financing healthcare in Bangladesh, which poses a substantial obstacle to achieving universal health coverage for the population (Rahman et al., 2022). High OOP expenditures pose a substantial financial burden on the impoverished population. According to Hamid et al. (2014), annual out-of-pocket expenses in healthcare contribute to increasing the poverty rate of Bangladeshi households by 3.4%. Excessive out-of-pocket healthcare expenses result in catastrophic health expenditure (CHE) across households, precipitating the economic impoverishment of millions of individuals, predominantly in rural rather than urban areas (Xu et al., 2003; Khan et al., 2017). Specially in the rural households, the incedence of CHE has increased persistantly between period 2005 to 2006 (Rahman et al., 2024). Rahman et al. (2022) reported 14% of Bangladeshi population faced forgone healthcare and major reason of this was treatment cost. Overall, AMR will pose more burden on rural population if people do not change their healthcare behavior and continues to receive medication and health advice from unsafe and unregistered service providers.

3.3 Research design

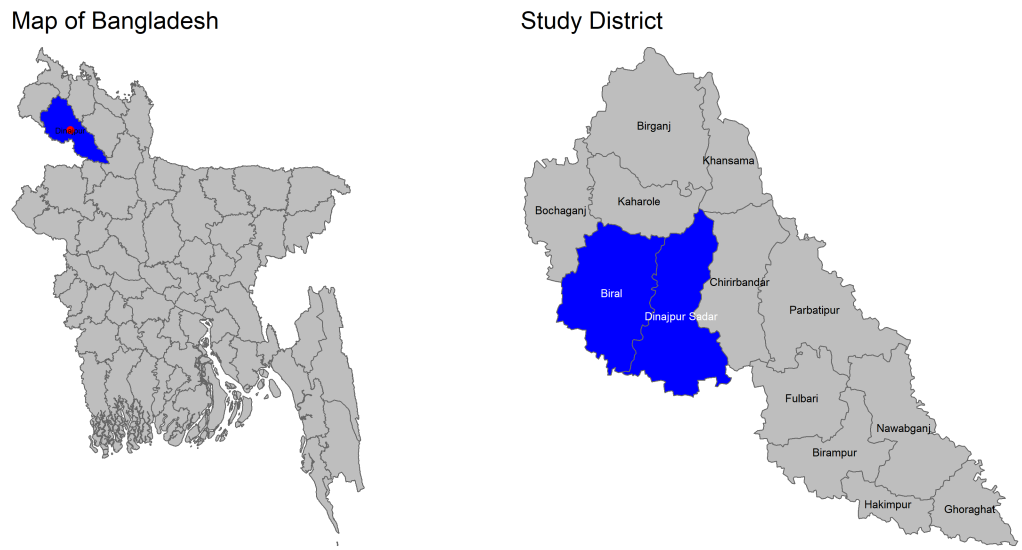
We designed our study to foster behavioral changes among the rural population regarding their choice of healthcare providers and limit the consumption of unsafe medications from available alternatives, which is essential to address the issue of antimicrobial resistance. To achieve our objective, we conduct a field experiment in Bangladesh. Governments and other organizations have already launched campaigns about AMR awareness, so in our design we assume that people have some prior risk information from other sources. Our intervention is not only designed to capture the impact of an information campaign that informs about the dangers of antimicrobial resistance; it instead attempts to identify additional barriers to safe healthcare practices after receiving any risk information regarding health. We examine additional factors that can lead people to choose safer medication options after receiving information. More specifically, instead of an information-only intervention, we focus on both the economic and psychological aspects that shape healthcare behavior in rural regions.

3.3.1 Program execution

To implement the study, we have partnered with a local NGO named the Education Health and Human Development Foundation (EHHD Foundation). This organization is registered with the Bangladesh government as a non-profit development organization. In partnership with the EHHD Foundation, we launched a health outreach initiative called "Kemon Achen?" which translates to "How are you?" in English. The program area was two upazilas (sub-district) of Dinajpur district named dinajpur sadar and Birol depicted in Figure 3.1. The program encompassed a total of sixty-two villages that were dispersed among seven Union Parishads.[10] This distribution was carefully planned to cover a wide geographical area to avoid of information spillover of our planned information treatment.

The "Kemon Achen?" program was set up to offer complimentary health assessment and nutritional services to women residing in study villages. Health assessments encompass essential measurements such as blood pressure and blood glucose levels. In addition, the nutritional services provided include basic measurements such as height, weight, and calculations of Body Mass Index (BMI). The EHHD Foundation recruited and provided rigorous training to female field workers, who are addresed as guides, in order to facilitate the delivery of services offered by the "Kemon Achen?" program. The guides visited selected villages in two sub-districts of Dinajpur district and approached evermarried women from randomly chosen households to participate in the program. They invited women to be a member of the program by explaining how the membership of the program plans to improve women's health and nutritional well-being through free monthly anthropometric measurements and educational information on different health-related topics. Enrollment in the program was voluntary and provided at no cost, with only one female participant per household approached only once. Upon expressing her consent to join the program, she was required to complete a form containing personal details. In the end, she was given a unique membership number. Upon the following month, she began receiving complimentary services provided by the guides.

[10] The smallest administrative unit at the rural level in Bangladesh, typically comprising several villages.



(a) (b)

Figure 3.1: Study area

Notes: Panel (a) shows the country map of Bangladesh, with the blue area highlighting the location of the study district Dinajpur. Panel (b) presents a detailed map of the Dinajpur district, where the blue highlighted area represents two sub-districts, namely Dinajpur Sadar and Biral, where the research is conducted.

The program started in April 2023 and continued through September 2023. During this period, enrolled members received monthly visits conducted by guides who were equipped with height measurement tapes, weight scales, and blood pressure monitors. Based on the outcomes of these assessments, the guides provided personalized health related advice. For example, in cases where a participant's blood pressure readings were found to be elevated, the guides recommended that they seek medical advice immediately from a physician. Likewise, if a participant's BMI deviated from the recommended thresholds, they were notified of their condition and provided with guidelines on how to modify their diet accordingly.

3.3.2 Randomization

During the household visits, it became apparent that rural households have very close community bonds and strong social connections, which are prone to substantial spillover effects between control and treatment groups. Due to the high degree of social interaction and information exchange among community members contamination is anticipated. Contamination is an important factor to consider cluster randomization as mentioned by (Eldridge and Kerry, 2012; Donner, 1998; Hayes and Moulton, 2017). Cluster randomization provides several other advantages, such as better representation of real-world scenarios, increased convenience and enhanced ethical considerations Campbell and Walters (2014). These considerations drove us to shift towards cluster randomization instead of individual randomization. It is also common that the cRCT requires a higher number of observation than individual randomization (Arnup et al., 2017; Hemming and Taljaard, 2023). From power calculation, we found that the existing number of villages was insufficient to conduct a robust clustered randomized controlled trial (cRCT). This lead us to add 310 more members from additional 22 villages. In the end, we had 1,129 members across 62 villages. The size of the villages varied in area and number of members reside on in them. Some villages are small while some villages are relatively larger and have several parts called "para", where people live very closely and have strong community sense.

Hayes and Moulton (2017) reports the use of arbitrary geographical zones as clustering units in areas where the population is dispersed and lacks well-defined communities. With widely scattered household in rural Ghana,Binka et al. (1996) conducted a clustered RCT, where they divided the area into 96 clusters of geographically adjacent households, with an average of about 1400 persons per cluster. Bousema et al. (2016) divided 100 sqkm area into 500 $\times$ 500 m cells in to study the impact of Hotspot-Targeted interventions on malaria transmission in Kenya. Consequently, to have more robust analysis with higher statistical power, we redefined our cluster units by adopting the use of arbitrary geographical zones. Utilizing the geographical coordinates collected during baseline survey, we created square grids over the operational area, each grid measuring 700 x 700 meters. Finally, the delineation of each grid cell encompassed an area measuring half a square kilometer, and a unique identifier was assigned to each grid. All households located within a single grid area were grouped into the same cluster. This reorganization resulted in each grid serving as a distinct cluster. Nevertheless, there was variation in the number of members across different grids; grids with less than five members were merged with the nearest adjacent grid. Following this reorganization, There were 68 grids on average 17 members in each grid. Figure 3.A.2 in Appendix shows the grids in the map of the study area.

Literature emphasizes the use of pair matching while using cluster randomization. For instance, Imai et al. (2009) recommends the using matching in cluster-randomized trials to enhance efficiency, power, robustness, research cost effectiveness, applicable to both large and small sample sizes. Arnold et al. (2024) demonstrated that geographic pair matching significantly enhances statistical efficiency in cluster randomized trials, and that trials without matching would require doubling the number of clusters to reach the same precision. Given the variability in the number of members and other characteristics across these grids, we decided to implement a match-pair randomized controlled trial (RCT) design. We matched our square grids based on average number of households in each grid, average age of the members, average visit in government health facilities and unqualified healthcare professionals in last six months. This entailed matching grids with similar characteristics. From each pair, one grid was randomly assigned as the control and the other as the treatment group. 3.A.3 shows the randomized grids in the map.

Following the initial cluster randomization (random assignment of grids into treatment or control groups), additional randomization was carried out within the treatment arms among members of the treatment grids. In addition to all members of the treatment groups receiving the same information about AMR, participants in this group were randomly assigned to receive one of two additional interventions: (I) received boxes to segregate medications by sources, or (II) received a subsidy for their travel expenses to registered physicians. This method ensured that both interventions were assessed separately for their efficacy, while preserving a common information component on AMR throughout the treatment groups.

Color coded boxes

Half of the randomly selected participants in the treatment group received two color-coded boxes as visual cue to distinguish the sources and associated dangers of unsafe medications. One box had a blue lid, provided for storing medications prescribed by registered doctors, reflecting a trusted and safe source. The other box, with a red lid, was for keeping medications that were self-purchased, prescribed by non-qualified practitioners/quacks, or suggested by drugstore salespersons. The use of red color in framing risk messages is well acknowledged. Pravossoudovitch et al. (2014) discusses communication value of the red color and recommends using red in conveying danger-relevant information. Gerend and Sias (2009) examines how red color, which is often associated with blood and danger, can enhance the effectiveness of loss-framed messages by serving as a peripheral threat cue that primes perceptions of threats. Our guides provided risk information that highlighted that medications stored in the red lid box could pose health risks, including the possibility of developing antimicrobial resistance. So, we used red lid boxes to symbolically represent the dangers associated with drugs bought from unsafe sources. Participants were clearly instructed to keep their medications in the respective boxes according to the source of the medicine. During the monthly visit, the guides ensured that they are using these boxes for keeping medicines and not use for other purposes. Each box was costed around 40 cents and the box treatment costs less than a dollar, which the respondents received only once. These cheap color-coded boxes were served as visual cues for motivating towards safe healh care practices. Section 3.A.3 contains the information sheet and color-coded boxes which worked as visual cues.

Transport cost subsidies

The remaining half of the treatment group received a cash transfer on the transport cost, conditional upon their visit to a qualified physician. Each member of this group was eligible to receive a reimbursement up to 100 BDT (approximately 80 cents) per month for this purpose. The allocation was valid for visit to both private and public health centers with registered doctors. Transport cost subsidies were provided to participants for expenses they incurred in last month during subsequent regular visits conducted by our health workers. The necessary requirement to obtain this subsidy was the submission of a dated prescription from a registered physician, which served as evidence of medical consultation.

The purpose of this cash transfer was to foster incentives to seek medical care from qualified healthcare providers instead of quacks, drugstore sales persons, or self-medication. This intervention sought to assess whether transportation costs restrict access to quality healthcare, leading people to seek service from unqualified healthcare professionals that leads to the emergence of antimicrobial resistance (AMR). We explore whether promoting better access to qualified professionals by reducing the loss from out-of-pocket expense lowers the dependence on unreliable health-care providers and diminishes the risk factors linked to AMR. In addition, in Bangladesh, the skilled health professionals are concentrated in urban areas, coupled with the fact that the majority of the population resides in rural region. The provision of transport costs highlights the impact of distance on the utilization of healthcare services located in urban areas. Section 3.A.4 in Appendix contains the information sheet and physician prescriptions that served as evidence of visit to registered doctors.

3.4 Data, hypotheses, and empirical methods

3.4.1 Data

We conducted our baseline survey in two phases, first among 820 participants from the end of September 2023 to early October 2023. We added an additional 310 participants to our study in October 2023 and conducted a survey on them in November 2023. In the baseline survey, the trained enumerators visited the household and collected data on household demographics, socioeconomic characteristics, medication behavior, preference of healthcare provider, choice of healthcare provider for different illnesses, and their knowledge about antibiotics. All of our respondents are female. For our final analysis, we analyze variables related to the number of visits to health facilities of different types, self-medication practice, source of the consumed medicine, the type of medicine, knowledge about AMR, trust in government and private health facilities. Our secondary outcome variables include quality of life, healthcare expenditure, community awareness and behavioral changes. The intervention started in January 2024. We collected data on changes in medication behavior after first and second month of intervention. We recorded member's visit to government health centers, private facilities, and quacks on February and March 2024, which we report in this paper. Including these, we schedule to collect the rest of the outcome variables in August 2024 through our endline survey.

3.4.2 Outcomes

Our main outcome of interest is changes in medication behavior of the study participants. We define the term medication behavior as a diverse range of practices and choices that individuals engage in when seeking and utilizing medication in response to their perceived health needs. This includes both the selection of sources for obtaining medications and the activities associated with this process. Qualified sources refer to registered healthcare providers, who operate within government facilities, private clinics or their own private practices.[11] Unqualified sources refer to individulas and establishments that lack official certification to offer medical advice or prescriptions but still provides services. This group ranges from unlicensed practitioners, commonly known as quacks, to sales personnel in drugstores, as well as instances of self-medication, where individuals buy and use medications by themselves. We

[11]Those Who are registered as doctor in Bangladesh medical and dental council (BMDC) with a registration numbers.

track medication behavior by measuring the number of times our participants visit various healthcare providers throughout the study. While visits to qualified facility are recorded through provided prescription documents, in many cases visits to unqualified providers and self medications are self reported as they do not provide any document. Besides medication behavior, we capture the impact of distance to government healthcare facilities from the respondent residents on their their visits to those facilities.

3.4.3 Hypothesis

It is anticipated that the implementation of our innovative interventions, namely the utilization of colored boxes to keep medicines and the provision of travel subsidies, in conjunction with the dissemination of risk information related to antimicrobial resistance, will result in a decrease in the unnecessary consumption of antibiotics among general people. Giné et al. (2010) showed that the provision of information on health hazards associated with smoking did not yield a substantial increase in smoking cessation rates. Nevertheless, the intervention demonstrated greater effectiveness when implemented in combined with commitment contracts, thereby emphasizing the necessity for supplementary mechanisms beyond the provision of information. In our case, providing cues and travel subsidies is expected to increase the effectiveness of the provided risk information and reduce the hazard of AMR. Educating women can have direct impact on the improved child health outcomes (Duflo, 2012). We expect that our intervention, which educates women in households about antimicrobial resistance, will directly impact the health of the family members. We also anticipate an increase in the utilization of government healthcare facilities over unqualified healthcare providers, which will lead to the overall enhancement in public health quality.

3.4.4 Empirical Approach

We estimate the following regression model-

$$Y_{ig} = \alpha + \beta_1 C_{ig} + \beta_2 T_{ig} + \gamma X_{ig} + w_g + \epsilon_{ig} \tag{3.1}$$

where Y_{ig} is the outcome variable from cluster g for member i after one and two months of intervention. C_{ig} indicates the dummy variable if the member receives visual cue treatment (blue and red boxes to keep medicine). T_{ig} is also a treatment dummy variable that indicates whether a member received a monthly travel cost subsidy up to bdt 100. X is vector of control variables that includes the age of the member, the level of education, the religion, the marital status, etc. w_g is grid fixed effect. We cluster our standard errors at the grid level. We estimate intent-to-treat (ITT) effects of our both treatment.

3.4.5 Balance check and summary statistics

Table 3.1 provides the summary statistics and the result of the baseline balance test. We conducted balance tests to compare the characteristics of the control and treated grids. In addition, we examine the differences between the individually randomized box recipients and travel subsidy recipients within the treatment grids. We used independent samples t-test to test balance between the groups. Almost of our covariates are balanced except the religion, both among the cluster and individual randomization. The study participants have an average age of 36.74 years and an average education level of 5.74 years. In addition, 92 percent of the participants in the study are currently married, and participants have an average of 2.37 children. Average monthly family income is 12,794 bdt.

In addition, we collected data on the participants' medical history and their medication-related behavior, such as the frequency of their visits to different healthcare providers. On average, respondents reported a frequency of 2.76 visits to government hospitals within the past 12 months, and a frequency of 1.64 visits within the past 6 months. They consulted quacks 4.64 times over the last year and 3.32 times in the past 6 months. Besides, respondents reported purchasing medicines by asking drugstore salespersons 2.76 times in the past 6 months, and they bought medicines without any prescription 1.31 times during the same period. Only 25 percent respondent reported they take vitamins or iron tablet. 14 percent women said that they can identify antibiotic. 29 percent respondent reported that antibiotic drugs work differently than other drugs.

We gathered responses regarding the choice of available medical service providers among the participants. The findings show that quacks are the most preferred healthcare services among all, expressed by 24 percent of respondents. In pairwise comparisons of providers, 53 percent of respondents expressed a preference for quacks over government hospitals, underscoring a significant dependence on informal healthcare providers. Conversely, 77 percent of respondents preferred government hospitals over private healthcare facilities, suggesting a higher dependence in public health services compared to private ones.

Table 3.1: Baseline balance table

	All sample			Experiment sample		Treatment Groups	
	All	Control	Treatment	Diff (Treatment-Control)	Box Treatment	Subsidy Treatment	Diff (Subsidy-Box)
Demography							
Age of respondent	36.74	36.91	36.55	0.37	36.05	37.03	-0.98
	(12.04)	(12.18)	(11.90)	(0.51)	(12.06)	(11.74)	(-0.96)
Years of education	5.67	5.77	5.56	0.21	5.78	5.34	0.44
	(3.97)	(4.05)	(3.89)	(0.88)	(3.84)	(3.93)	(1.33)
Marital status(currently married=1)	0.92	0.91	0.94	-0.02	0.93	0.94	-0.01
	(0.27)	(0.28)	(0.25)	(-1.41)	(0.25)	(0.24)	(-0.30)
Religion(Islam=1)	0.81	0.72	0.90	-0.19***	0.92	0.89	0.03
	(0.40)	(0.45)	(0.30)	(-8.36)	(0.28)	(0.31)	(1.01)
Total children	2.37	2.36	2.38	-0.02	2.42	2.34	0.08
	(1.24)	(1.25)	(1.23)	(-0.25)	(1.13)	(1.31)	(0.64)
Medical History							
Govt. hospital visit(6 Months)	1.64	1.57	1.71	-0.15	1.76	1.67	0.09
	(2.52)	(2.45)	(2.58)	(-0.95)	(2.65)	(2.52)	(0.39)
Govt. hospital visit(12 Months)	2.76	2.91	2.60	0.30	2.64	2.57	0.07
	(4.48)	(4.62)	(4.33)	(0.95)	(4.11)	(4.54)	(0.15)
Quack visit(6 Months)	3.32	3.38	3.26	0.12	3.47	3.05	0.43
	(3.74)	(3.82)	(3.65)	(0.53)	(4.10)	(3.15)	(1.32)
Quack visit(12 Months)	4.64	4.69	4.60	0.09	4.90	4.32	0.59
	(6.31)	(6.42)	(6.20)	(0.20)	(7.09)	(5.26)	(0.90)
Bought from medicine shop (6 Months)	2.76	2.76	2.75	0.01	2.80	2.72	0.08
	(4.33)	(3.98)	(4.67)	(0.03)	(5.46)	(3.83)	(0.16)
W/O prescription (6 Months)	1.31	1.26	1.37	-0.12	1.36	1.38	-0.02
	(2.71)	(2.55)	(2.87)	(-0.58)	(2.49)	(3.18)	(-0.07)
Takes Vitamin or Iron(Yes=1)	0.25	0.23	0.27	-0.03	0.27	0.26	0.00
	(0.43)	(0.42)	(0.44)	(-1.27)	(0.44)	(0.44)	(0.09)
Can identify antibiotic(Yes=1)	0.14	0.14	0.14	-0.00	0.14	0.15	-0.01
	(0.35)	(0.35)	(0.35)	(-0.04)	(0.34)	(0.36)	(-0.42)
Knows Antibiotic works differently (Yes=1)	0.29	0.28	0.29	-0.01	0.31	0.28	0.03
	(0.45)	(0.45)	(0.46)	(-0.28)	(0.46)	(0.45)	(0.87)
Medicine Knowledge and Choice							
Most preferable service(Quacks=1)	0.24	0.22	0.26	-0.04	0.25	0.28	-0.04
	(0.43)	(0.42)	(0.44)	(-1.61)	(0.43)	(0.45)	(-0.98)
Choice: Govt. or Quack(Quack=1)	0.53	0.53	0.54	-0.01	0.52	0.55	-0.04
	(0.50)	(0.50)	(0.50)	(-0.14)	(0.50)	(0.50)	(-0.68)
Choice: Govt. vs pvt hospital(Govt.=1)	0.77	0.78	0.76	0.02	0.75	0.77	-0.02
	(0.42)	(0.41)	(0.43)	(0.52)	(0.43)	(0.42)	(-0.48)
Main Health Info Source(Quacks=1)	0.18	0.17	0.19	-0.02	0.19	0.18	0.00
	(0.38)	(0.37)	(0.39)	(-0.94)	(0.39)	(0.39)	(0.12)
Income and Health Expenditure							
Monthly Family Income	12794.90	12563.33	13037.75	-474.42	12594.08	13438.13	-844.05
	(8322.52)	(8420.40)	(8222.42)	(-0.81)	(7002.16)	(9184.58)	(-1.03)
Transport Cost Govt. HC	75.35	76.80	73.81	2.99	71.23	76.05	-4.82
	(74.15)	(75.32)	(72.97)	(0.57)	(72.73)	(73.27)	(-0.65)
Observations	1,128	587	541	1,128	265	276	541

Notes: This table presents descriptive statistics and balance test for the full sample and experimental subgroups for baseline data. Columns 1-3 show statistics for the entire sample, control group, and treatment group, respectively. The treatment group refers to the person who received information about AMR along with either boxes or transport cost subsidies. Column 4 displays the difference between treatment and control groups. Columns 5-6 break down the treatment group into Box and Subsidy subgroups, while column 7 shows their differences. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Most importantly, our findings reveal that quacks serve as the primary health information source for 18 percent of respondents, exceeding all other sources. These statistics highlight the significant impact of unqualified medical professionals on health-seeking behavior among rural communities. The mean transport expense linked to utilizing the closest government health facility is 75 BDT, a factor that may impact the selection of healthcare services among the public.

3.4.6 Distance to healthcare facility and visits

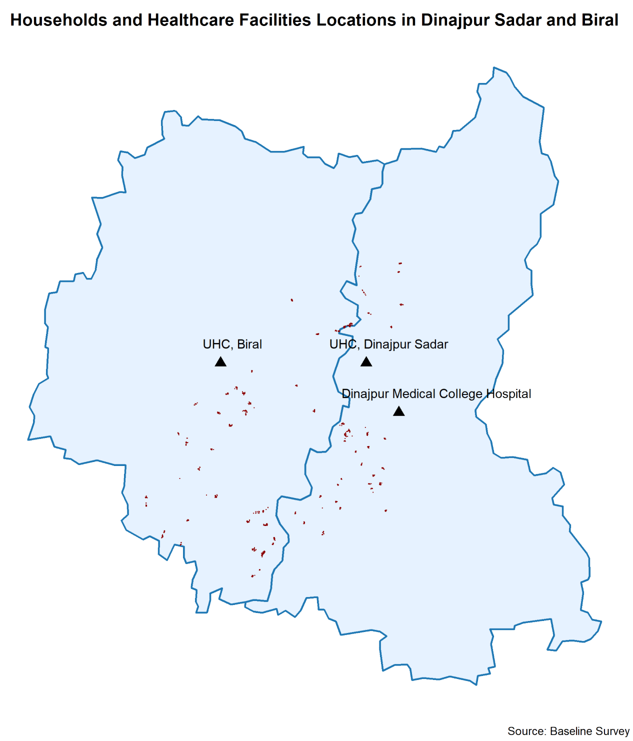


Figure 3.2: Location of government healthcare facilities and respondent residence

Notes: This plot shows both the geographic locations of our respondents and the sub-district (upazila) level government healthcare facilities. The red dots represent the locations of respondents' residences. The triangle indicate the locations of Upazila Health Complexes (UHCs) and medical college hospitals in the study areas.

Distance to healthcare providers is one of the main determinate factors in the choice of healthcare facility. It might take a considerable amount of transport cost and time to reach qualified doctors in government and private facilities if people live in remote areas. while people can walk in unqualified healthcare providers and ask for medications at any time as Unqualified practitioners, such as village doctors reside in almost every village. Using baseline data, we first analyze the impact of distance to medical facility on the medical service utilization behavior of the respondents. Specifically, we investigate whether residing a longer distance from the nearest government upazila-level health complex influences the likelihood of respondents visiting government and unqualified practitioners (e.g., quack) facilities.[12] As unqualified practitioners and self-medication are the main sources of unnecessary antibiotic consumption, spatial analysis provides substantial insights into how geographic distances contribute to these behaviors, thus underscoring the importance of developing focused strategies such as transport subsidy to qualified healthcare providers. We calculated the direct distance form the coordinates of the households to the nearest government upazila level healthcare facility. We used following haversine formula to calculate great-circle distance, which is the shortest distance between two points on the surface of a sphere (Bullock, 2007; Pednekar and Peterson, 2018).

$$d = 2r \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right) \tag{3.2}$$

where ϕ_1, ϕ_2 are the latitudes of the two points(household and healthcare facility) in radians, λ_1, λ_2 are the longitudes of the two points(household and healthcare facility) in radians, r is the radius of the sphere (for Earth, approximately 6,371 kilometers) and d is the great-circle distance. We calculated the distance from the respondent's household to all three government facilities (Dinapur sadar upazila health complex, Biral upazila health complex, and Dinajpur medical college hospital) and created a distance variable taking the minimum distance from those facilities. This variable indicates the minimum distance to nearest government health center from where the respondent lives. Figure 3.2 displays the points where respondent resides and government healthcare facilities are located in our study area.

First, we implemented the OLS regression specification to analyze the impact of the distance to the nearest government hospitals on the choice of healthcare providers. We conducted Moran's I test, which checks for spatial autocorrelation in residuals (Li et al., 2007). In our test, we get the Moran I statistic of 0.222 (expected = -0.00095) with a $p < 0.001$, indicates significant spatial autocorrelation in the residuals of the OLS model. This spatial autocorrelation is shown in the Moran scatterplot, depicted in Figure 3.B.7 in Appendix. The presence of spatial autocorrelation drives

[12]An upazila health complex is primary government healthcare centers at sub-district level that provides outdoor expert doctor consultation opportunity and admission facilities.

Table 3.2: Distance and healthcare facility visits

	Last six month visit					
	Quack			Govt. facility		
	OLS	SAR	SER	OLS	SAR	SER
Distance	0.377***	0.236***	0.370***	−0.039	−0.034	−0.043
	(0.045)	(0.037)	(0.055)	(0.029)	(0.025)	(0.031)
Age	0.041***	0.031***	0.028**	0.020**	0.019**	0.020**
	(0.012)	(0.011)	(0.011)	(0.009)	(0.008)	(0.008)
Education	0.034	0.012	0.006	0.041	0.040*	0.042*
	(0.033)	(0.032)	(0.033)	(0.029)	(0.023)	(0.024)
Married	−0.029	−0.220	−0.299	−0.199	−0.210	−0.200
	(0.527)	(0.423)	(0.426)	(0.313)	(0.306)	(0.309)
Muslim	1.764***	1.175***	1.917***	1.192***	0.958***	1.165***
	(0.243)	(0.269)	(0.399)	(0.137)	(0.197)	(0.231)
Constant	−1.810**	−1.145	−0.999	0.118	0.019	0.159
	(0.830)	(0.758)	(0.868)	(0.604)	(0.549)	(0.581)
Observations	1,057	1,057	1,057	1,071	1,071	1,071
R^2	0.119			0.050		
Adjusted R^2	0.115			0.046		
σ^2		10.983	10.960		5.909	5.906
Akaike Inf. Crit.		5,576.664	5,575.551		4,964.192	4,963.903
Wald Test (df = 1)		106.385***	110.304***		18.161***	18.585***
LR Test (df = 1)		96.650***	97.764***		18.026***	18.315***

Notes: Standard errors are robust for ols. SAR stands for spatial autoregressive model, SER stands for spatial error models. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

us to applly spatial econometric techniques for our analysis of impact of distance on medical facility visits. We used both the spatial error model (SEM) and the spatial autoregressive model (SAR) in our analysis. The SAR model incorporates a spatially lagged dependent variable to directly consider the effects of neighboring observations. On the other hand, SEM addresses spatial dependencies in the error terms instead of the dependent variable.

Table 3.2 describes the results of the ordinary least squares (OLS), spatial error (SEM), and spatial autoregressive (SAR) model that assesses the impact of geographic distance from the place of residence to the nearest government hospital on the utilization of available healthcare services. It has been previously mentioned that visits to informal healthcare providers (quacks) and government facilities directly influence unnecessary antibiotic consumption. The regression results indicate a significant impact of the shortest distance to the nearest government healthcare facility on the number of visits to quacks. Specifically, each additional mile from the residence to the closest Upazila Health Complex or Medical College Hospital is associated with an increase in visits to quacks over the past six months by 0.38 in the OLS model, 0.24 in the SAR model, and 0.37 in the SEM model, with all $p < 0.01$. Conversely, the distance to the nearest government facility exhibits a negative, but non-significant, effect on the visits at these establishments.

3.5 Results

Immediate treatment effect

We started our treatment implementation in January 2024 and evaluated the immediate impact of our designed intervention after two months of roll-out. In March'2024, we collected data on the healthcare service choice of respondents. Figure 3.3 compares the average visit to government healthcare centers, private qualified practitioners and unqualified quack visit by treatment status. The bar graph depicts the impact of the risk information intervention combined with visual cues and travel subsidies, on the frequency of visits to various medical facilities in February.

The average number of visits to registered doctors in government facilities was 0.47, while at private facilities it was 0.09 among visual cue receivers. In comparison, the control group had an average of 0.18 visits to government facilities and 0.05 visits to private facilities at the same time. The average visit to an unqualified medical professional or quack was 1.10, while the average visit among the control group was 1.04. similarly, the average number of visits to registered doctors at government facilities was 0.37 and 0.07 for private facilities for travel subsidy receivers, compared to 0.21 and 0.05 visits respectively for the control group. The average number of visits to unqualified medical professionals or quacks was 0.89 for those receiving travel subsidies, compared to 1.10 for the control group. For both intervention, recipients show increased visits to registered facilities but mixed evidence on quack visit. While cue receivers showed a marginal increase in quack visit, subsidy receivers had shown a reduction in quack visit.

The regression results presented in Table 3.3 show the effects of our intervention

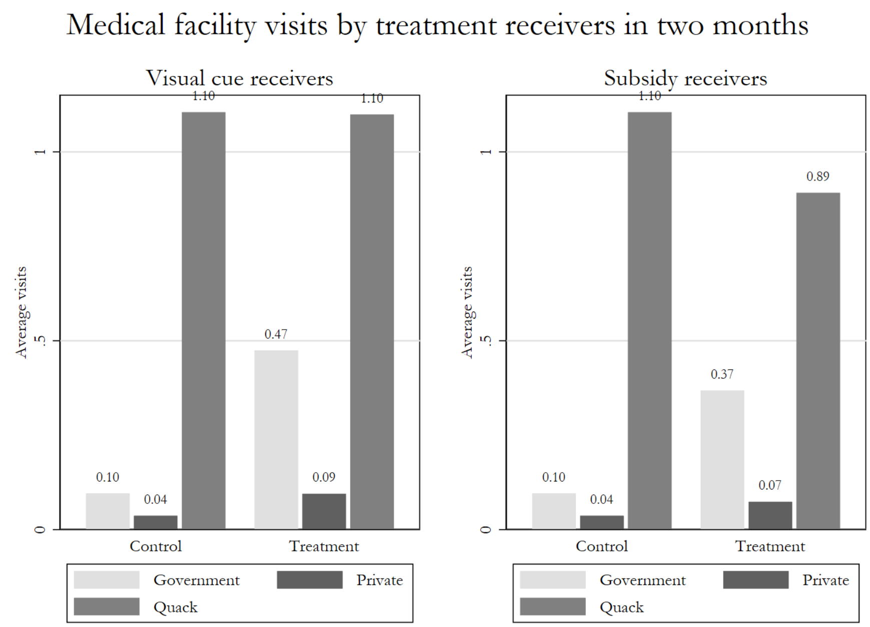


Figure 3.3: Visit to healthcare facilities by treatment recipient

Notes: This plot displays average number of visit to qualified and unqualified healthcare providers after second month of intervention. We sub divided qualified healthcare providers into government and private group. The left panel reports average visits to these facilities for visual cue receivers and right panel reports average visits of travel subsidy receivers and compares with respective control group.

after two months. Columns 1 and 2 display the effect of risk information regarding AMR with any additional interventions- travel subsidy or visual cues. After two months of treatment roll-out, treated individuals had a 0.37 ($p < 0.01$) higher visit to qualified healthcare professionals. When we separately analyze the impact of visual cue and travel subsidy deployed with risk information, we found differential impact. The visual cue recipients had a higher impact on visit to qualified physicians from government facilities, with 0.39 ($p < 0.01$) more visit relative to controls. Those who received travel subsidy showed a 0.27 higher visit ($p < 0.01$) to qualified professionals from government facilities relative to control. We find no significant increase in private hospital visit from either subsidy or cue recipient in two months combined treatment effect. This group does not have significant increase in visit to qualified healthcare providers from private facilities. Besides the combined two month effect, we also have analyzed treatment effects for January and February, that is in first and second month of intervention. Table 3.B.1 shows the treatment effect in first month, which shows

Table 3.3: Treatment evaluation for January–February 2024

	Healthcare facility visit			
	All visit		visit type	
	(1) two-month registered	(2) two-month quack	(3) two-month govt	(4) two-month pvt
Information + Subsidy/Box	0.369*** (0.066)	-0.196 (0.170)		
Information + Subsidy			0.272*** (0.055)	0.030 (0.026)
Information + Box			0.386*** (0.086)	0.052 (0.031)
N	1068	1067	1068	1068

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after two month treatment roll-out. Standard errors are robust and clustered by square grid area, reported in parentheses. Two-month registered means number of visit in last two months to any registered physician. Two-month quack means number of quack visits in two months. Two-month govt means any government hospital visit in two months. and two-month pvt means any private clinic or chamber visit in last two months. All regression are controlled for member age, education, religion and marital status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

that overall 0.23 higher visit (p < 0.01) by treatment recipients. The subsidy reiverer and cue receivers had a 0.22 and 0.21 higher visit (p < 0.01) in registered physicians of government facility. We did not find a significant increase in visit to private facilities in either group. Table 3.B.2 shows the effect of the treatment in the second month, where the overall treatment recipients had a 0.23 higher visit to registered doctors. Subsidy receivers do not have a significantly higher visit to government facilities, but show a significantly higher visit to private registered physicians by 0.04 visits (p < 0.01). Cue receivers have a 0.17 higher visits (p < 0.01) to government hospitals while no significant impact on private facility visits.

As robustness analysis, we added double lasso regression analysis in section 3.C.1 of Appendix, where section 3.C.2 we provide regression models without covariates. Both analyses provide consistent results.

3.6 Conclusion and discussion

Addressing antimicrobial resistance is necessary as it is a threat to several sustainable development goals that range from good health and well-being to combating poverty and hunger. In our study, we examine several approaches to prevent AMR in the least developed country setting, where unqualified health professionals are major healthcare provider in rural areas and over-the-counter antibiotic sales are a common scenario. We conducted a randomized intervention of risk information along with nudges and subsidies to modify healthcare seeking behavior in rural Bangladesh. Our interventions try to revert people towards registered health care providers instead of frequent visits to unqualified quacks, drugstore sales person and self medication practices. This paper reports the results of the early stage interventions aimed at obstructing the paths of unnecessary antibiotic consumption. After two months of intervention, we found a substantial impact on the choice of healthcare provider among the treatment recipents compared to those who did not receive any intervention.

Our baseline survey reveals that rural women have low level of health literacy, particularly concerning their understanding of antibiotics and it's role roles. We found that they visit unqualified health professionals at least twice as often as qualified doctors. Besides frequent self-medication practice, they also frequently buy medicine by taking advice of the drugstore salespersons. The rural population largely relies on quacks as a source of health information. At the same time, quacks are most preferable healthcare provider and they choose quacks over government healthcare centers.

This paper documents two important results that affect AMR. Firstly, our analysis reveals that the distance from the place of residence to the nearest government healthcare center plays a vital role in the choice of healthcare providers. If the distance to the nearest healthcare center increases, people significantly increases their visit to quacks. This finding is consequential, as an increase in the number of visits to unqualified healthcare providers poses higher risks of antimicrobial resistance. Secondly, our randomized interventions of providing risk information along with travel subsidies and color-coded boxes as nudge effectively increased visit to healthcare facilities with qualified healthcare professionals. This increased visit to qualified healthcare provider is expected to reduce self medication and quack visit, leading to reduced risk of unnecessary antibiotic consumption.

We identified several potential channels that explain the effectiveness of these interventions. Primarily, prior to information intervention, rural women had limited knowledge about AMR. However, understanding of AMR was significantly influenced by the information they received regarding AMR associated risks and transmission pathways. Consequently, their beliefs about various healthcare providers and the potential consequences of the medications those provider were chaned by this provided information, which increased visits to qualified healthcare providers. Secondly, the travel subsidy reduced the loss from out-of-pocket healthcare expenses. This reduced loss drives individual drives respondents towards better healthcare provider choice. Furthermore, the presence of blue and red boxes enhances the the cue recipient's salience of the provided risk information. The boxes also influence the prospective memory on the adoption of healthy medication practices that emphasizes avoid the possible harm of consuming antibiotics without a prescription from qualified physicians.

Our findings have several advantages and significant policy implications in combating AMR. Interventions not only help to reduce the risk of AMR, but also promote the adoption of better health-seeking behaviors and practices, which consequently results in improved overall health outcomes. Both interventions are cost-effective, specially the one-time cost of the visual cue being less than 1 euro. Policy makers may also consider the implementation of complementary transportation services from remote rural areas to government healthcare centers. This approach can enhance the accessibility of high-quality healthcare services and reduce frequent visits to unqualified healthcare providers, drugstore salespeople, and self-medication practices. In addition, it does not require expensive medically trained health professionals in monthly door-to-door visit, health workers with relevant training can supervise people and deliver the risk information regarding AMR. Ultimately, our approaches have the potential to effectively address the hazards associated with antimicrobial resistance and reduce the overall burden on healthcare infrastructure, rendering it a judicious option for policy deliberation.

Appendix

3.A Study Design

3.A.1 Healthcare system hierarchy in Bangladesh

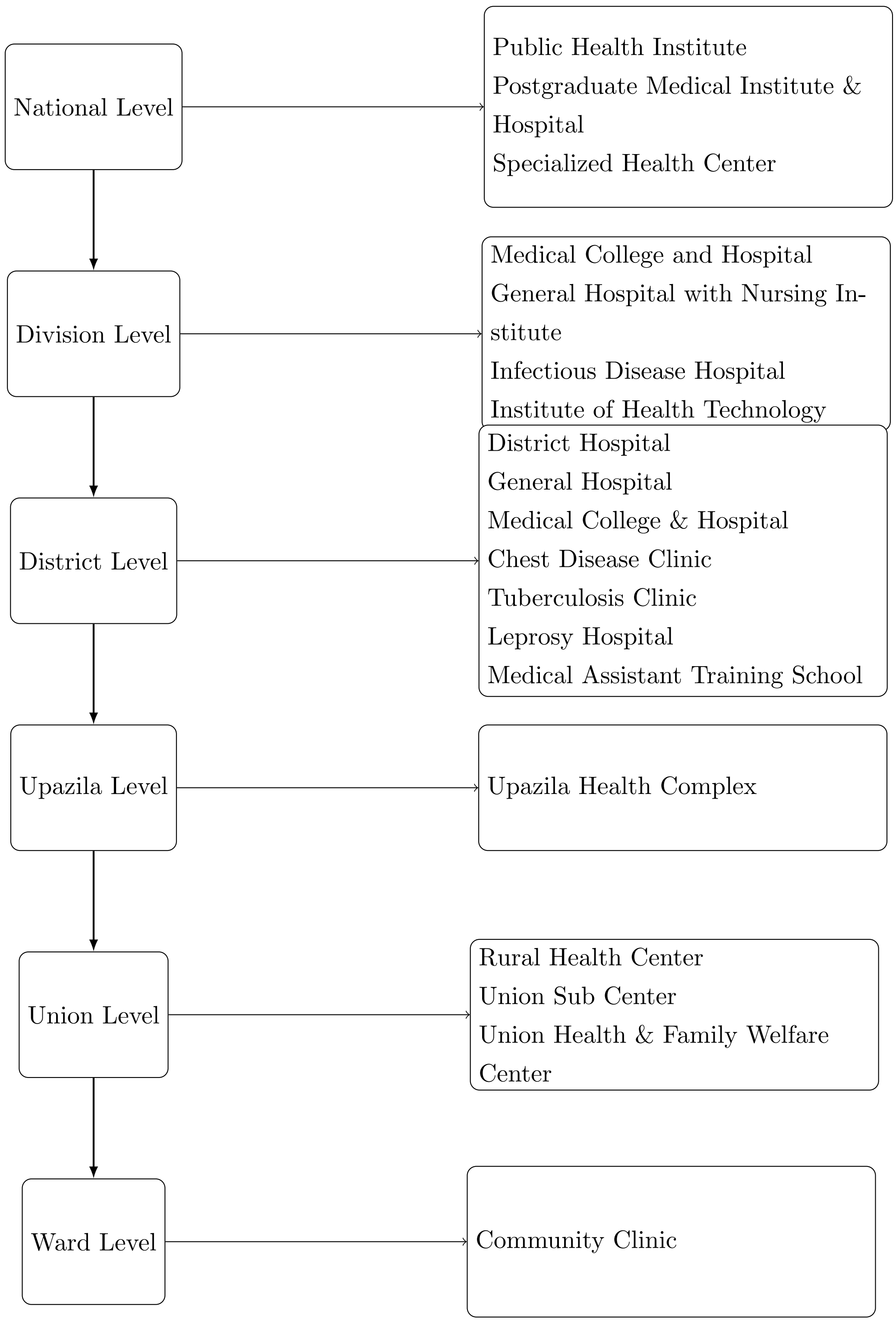
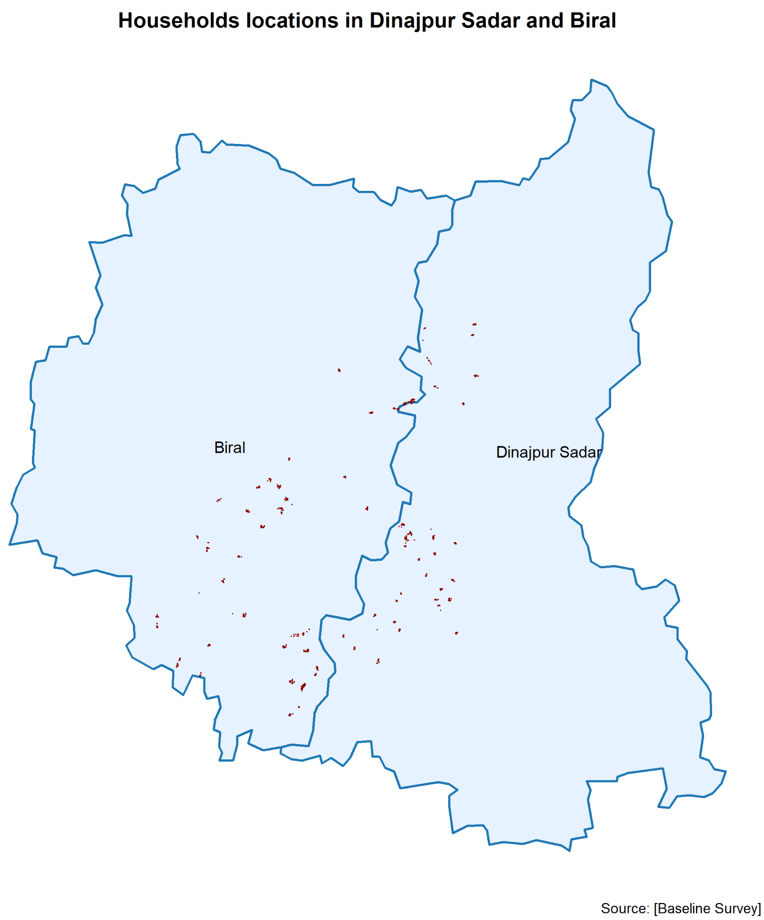


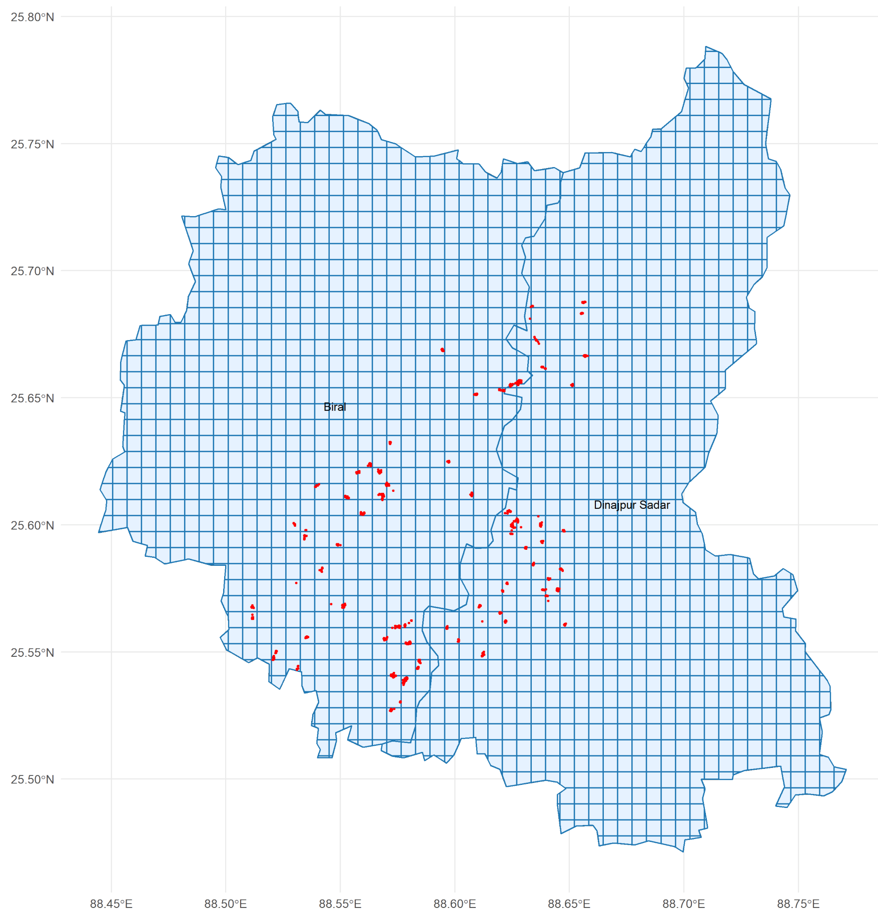
Figure 3.A.1: Study household location



Notes: The red dots in the map shows exact locations of the member households in Dinajpur Sadar and Biral Upazila.

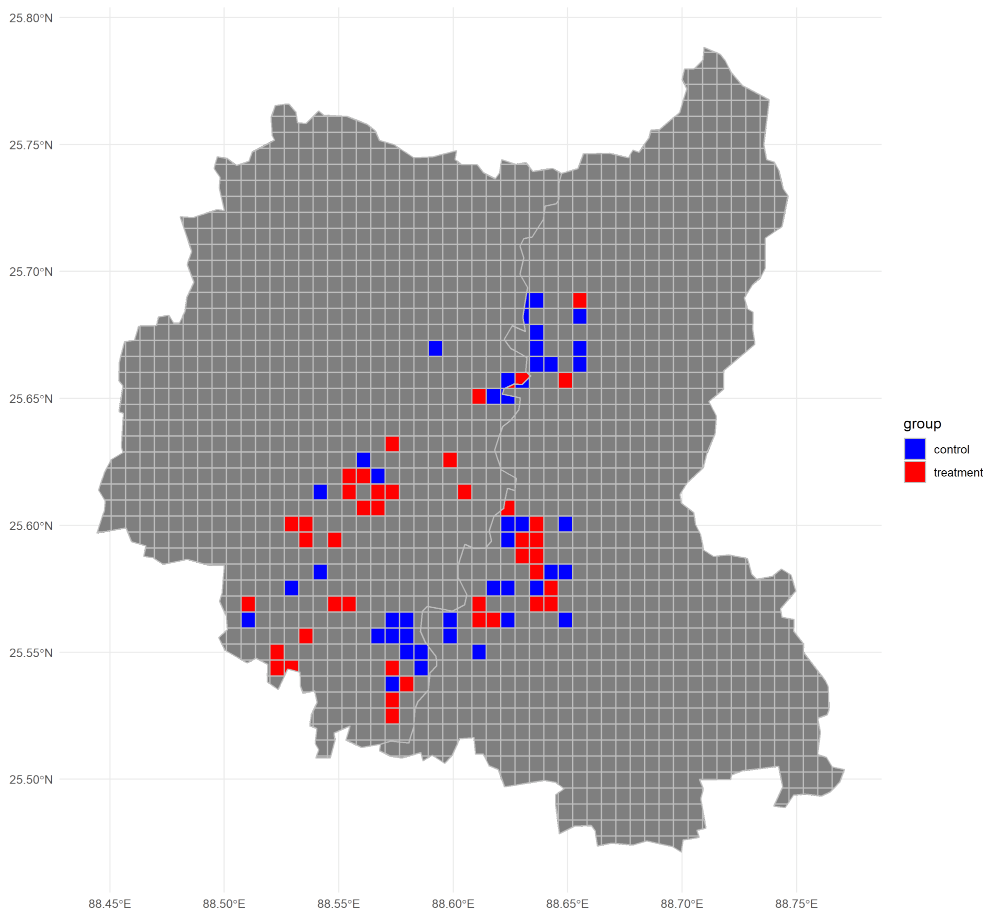
3.A.2 Grid based randomization

Figure 3.A.2: Study household location within square grids



Notes: This map illustrates the study area in Dinajpur Sadar and Biral subdistricts of Bangladesh. The area is divided into a grid system, with each square measuring 700 meters on each side. These grid squares serve as the basis for our spatial analysis and sampling strategy. The dots represent individual households that participated in the study.

Figure 3.A.3: Randomization within square grids



Notes: This map shows the study area in Dinajpur Sadar and Biral subdistricts, divided into grids. These grids serve as clusters for the study. Each grid (cluster) measures 700 meters on each side, covering an area of 0.49 square kilometers. Blue grids represent control clusters, while red grids indicate treatment clusters. The clusters were randomly allocated to either control or treatment status for the study.

3.A.3 Treatment I: Information + box

The guides convey following information from a printed sheet to the box recipients:

Information sheet

box group

We will provide new nutrition card services from this round as well as make you aware about the usage of drugs. We all need to be aware of antibiotic drugs, because-

• If the human body is over-applied with antibiotics or not properly applied, the bacteria in the body develop resistance against antibiotics. This is called antibiotic resistance.

• Antibiotic resistance refers to loss of effectiveness of antibiotics. Taking antibiotics without doctor's advice and due to lack of awareness, antibiotics lose their effectiveness.

• Antibiotic resistance has emerged in many people treated for common diseases in various hospitals in Bangladesh.

• As a result, the antibiotics that used to cure the patient are no longer working. Otherwise very high doses of antibiotics have to be given.

- Many times we take medicines without the advice of a registered doctor, many times we buy medicines ourselves. In this, without our knowledge, the efficiency of antibiotic medicine is getting lost in our body. As a result, antibiotics will no longer work when our body needs them in the future.

So you are requested to always take the medicine as per the advice of the registered doctor.

For your convenience, we will provide you with two medicine boxes. In the box with blue lid, you should keep the medicine and prescription prescribed by the registered doctor (Government/Private Hospital/Chamber/Clinic/Private MBBS Doctor).

-Blue color will be the symbol of safe drug.

- Red box will keep the medicines and all other medicines. Red color to remind of potentially unsafe drugs/antibiotics.

Do not use any box for any other purpose. You will not get any benefits from us in the future.

– Consult MBBS doctor before taking medicine. Try going to a government hospital.

Figure 3.A.4: Visual cues: The color-coded boxes

Notes: Images show two boxes with blue and red lids. Box recipient household received two boxes with blue and red lids. Blue lid boxes contain medication provided by qualified registered doctors and red lid box contain drugs form all other unreliable sources.

Figure 3.A.5: Box usage during study

(a) (b)

Notes: Panel (a) demonstrates how a control group participant keeps the drugs. Photo is taken during baseline Panel (b) show hose box is used by box recipients. The image is collected in mobile app during the household visit.

3.A.4 Treatment II: Information + subsidy

Travel subsidy Group
We will provide new nutrition card services from this round as well as make you aware about the usage of drugs.
We all need to be aware of antibiotic drugs, because
• If the human body is over-applied with antibiotics or not properly applied, the bacteria in the body develop resistance against antibiotics. This is called antibiotic resistance.
• Antibiotic resistance refers to loss of effectiveness of antibiotics. Taking antibiotics without doctor's advice and due to lack of awareness, antibiotics lose their effectiveness.
• Antibiotic resistance has emerged in many people treated for common diseases in various hospitals in Bangladesh.
• As a result, the antibiotics that used to cure the patient are no longer working. Otherwise very high doses of antibiotics have to be given.
- Many times we take medicines without consulting a registered doctor, many times we buy medicines ourselves. In this, without our knowledge, the efficiency of antibiotic medicine is getting lost in our body. As a result, antibiotics will no longer work when our body needs them in the future.
So you are requested to always take the medicine as per the advice of the registered doctor.
To encourage you to seek treatment from your registered doctor, we will cover your (female member's, no other household's) home to government clinic once a month. If you show the hospital ticket/prescription you will be eligible to get this fare. We have no other motive behind this than to ensure your safe use of medicines. If the travel cost of going to the nearest upazila health complex or private hospital is less than (100 taka), you will get this benefit. For this, we will take a photo of the prescription/ticket with the date of that month (once between 1st and 31st) and send it to our office. I will pay this month's travel expenses on our next visit. We cannot offer the same benefits to everyone, but through the lottery each of our members will receive different benefits. You will get travel expenses. If you do not go in any month, that month will not be given in the next month. 1 month, 1 member, 1 time.

– Consult MBBS doctor before taking medicine. Try going to a government hospital.–

Figure 3.A.6: Prescriptions: as evidence of subsidy

Notes: Images show prescription of last month collected in mobile app during the household visit.

3.B Results

Figure 3.B.7: Moran Scatterplot

Notes: The scatterplot represents the standardized residuals obtained from the ordinary least squares (OLS) regression, plotted against their corresponding spatially lagged counterparts. Each dot illustrates a spatial entity, denoted by its residual value. The positive slope of the fitted blue line indicates moderate positive spatial autocorrelation. This autocorrelation suggests that nearby locations exhibit similar residual patterns, requiring spatially adjusted modeling.

Table 3.B.1: Treatment evaluation for Jan 2024

	Healthcare facility visit			
	All visit		visit type	
	(1) Jan registered	(2) january_quack	(3) January govt	(4) January pvt
Information + Subsidy/Box	0.230***	-0.135		
	(0.045)	(0.093)		
Information + Subsidy			0.220***	-0.010
			(0.047)	(0.016)
Information + Box			0.205***	0.030
			(0.044)	(0.021)
N	1068	1068	1068	1068

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after first month treatment roll-out. Standard errors are robust and clustered by square grid area in map, reported in parentheses. Jan registered means number of visit in January to any registered physician. jan_quack means number of quack visits in January. January govt means any government hospital visit in January. and January pvt means any private clinic or chamber visit in January. All include covariates of member age, education, religion and marital status * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.B.2: Treatment evaluation for February 2024

	Healthcare facility visit			
	All visit		visit type	
	(1) Feb registered	(2) February quack	(3) February govt	(4) February pvt
Information + Subsidy/Box	0.135***	-0.063		
	(0.036)	(0.092)		
Information + Subsidy			0.041	0.040**
			(0.030)	(0.018)
Information + Box			0.169***	0.023
			(0.051)	(0.016)
N	1077	1076	1077	1077

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after first month treatment roll-out. Standard errors are robust and clustered by square grid area in map, reported in parentheses. Jan registered means number of visit in January to any registered physician. jan_quack means number of quack visits in January. January govt means any government hospital visit in January. and January pvt means any private clinic or chamber visit in January. All include covariates of member age, education, religion and marital status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.C Robustness

3.C.1 Regression without covariates

Table 3.C.3: Treatment evaluation for two months combined

	Healthcare facility visit			
	All visit		visit type	
	(1) two-month registered	(2) two-month quack	(3) two-month govt	(4) two-month pvt
Information + Subsidy/Box	0.372***	-0.111		
	(0.066)	(0.173)		
Information + Subsidy			0.272***	0.037
			(0.054)	(0.026)
Information + Box			0.378***	0.058*
			(0.088)	(0.030)
N	1084	1083	1084	1084

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after two month treatment roll-out. Standard errors are robust and clustered by square grid area, reported in parentheses. Two-month registered means number of visit in last two months to any registered physician. Two-month quack means number of quack visits in two month. Two-month govt means any government hospital visit in two months. and two-month pvt means any private clinic or chamber visit in last two months. No controls are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.4: Treatment evaluation for January

	Healthcare facility visit			
	All visit		visit type	
	(1) Feb registered	(2) February quack	(3) February govt	(4) February pvt
Information + Subsidy/Box	0.146***	-0.013		
	(0.037)	(0.095)		
Information + Subsidy			0.050	0.045**
			(0.031)	(0.018)
Information + Box			0.172***	0.027*
			(0.053)	(0.015)
N	1093	1092	1093	1093

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after first month treatment roll-out. Standard errors are robust and clustered by square grid area in map, reported in parentheses. Jan registered means number of visit in January to any registered physician. jan_quack means number of quack visits in January. January govt means any government hospital visit in January. and January pvt means any private clinic or chamber visit in January. No controls are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.5: Treatment evaluation for February

	Healthcare facility visit			
	All visit		visit type	
	(1)	(2)	(3)	(4)
	Jan registered	january_quack	January govt	January pvt
Information + Subsidy/Box	0.223***	-0.101		
	(0.042)	(0.094)		
Information + Subsidy			0.220***	-0.009
			(0.047)	(0.015)
Information + Box			0.205***	0.030
			(0.044)	(0.021)
N	1084	1084	1084	1084

Notes: This table shows the results of linear regression models estimating treatment effects on different health care provider visit after second month treatment roll-out. Standard errors are robust and clustered by square grid area in map, reported in parentheses. Feb registered means number of visit in February to any registered physician. February quack means number of quack visits in February. February govt means any government hospital visit in February. and February pvt means any private clinic or chamber visit in February. No controls are included. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.C.2 Double-selection lasso regression

Table 3.C.6: Treatment evaluation for January–February 2024: double-selection lasso model

	Healthcare facility visit			
	All visit		visit type	
	(1) two-month registered	(2) two-month quack	(3) two-month govt	(4) two-month pvt
Information + Subsidy/Box	0.371***	-0.127		
	(0.066)	(0.167)		
Information + Subsidy			0.267***	0.038
			(0.054)	(0.026)
Information + Box			0.378***	0.059**
			(0.088)	(0.030)
N	1068	1067	1068	1068

Notes: This table shows result of double-selection lasso linear regression specification. Standard errors are robust and clustered by square grid area, reported in parentheses. Two-month registered means number of visit in last two months to any registered physician. Two-month quack means number of quack visits in two month. Two-month govt means any government hospital visit in two months. and two-month pvt means any private clinic or chamber visit in last two months. All include covariates of member age, education, religion and marital status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.7: Treatment evaluation for January 2024: double-selection lasso model

	Healthcare facility visit			
	All visit		visit type	
	(1)	(2)	(3)	(4)
	two-month registered	two-month quack	two-month govt	two-month pvt
Information + Subsidy/Box	0.371***	-0.127		
	(0.066)	(0.167)		
Information + Subsidy			0.267***	0.038
			(0.054)	(0.026)
Information + Box			0.378***	0.059**
			(0.088)	(0.030)
N	1068	1067	1068	1068

Notes: This table shows result of double-selection lasso linear regression specification. Standard errors are robust and clustered by square grid area, reported in parentheses. Jan registered means number of visit in January to any registered physician. January quack means number of quack visits in January. January govt means any government hospital visit in two months. and January pvt means any private clinic or chamber visit in January. All include covariates of member age, education, religion and marital status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.C.8: Treatment evaluation for February 2024: double-selection lasso model

	Healthcare facility visit			
	All visit		visit type	
	(1)	(2)	(3)	(4)
	Feb registered	February quack	February govt	February pvt
Information + Subsidy/Box	0.145***	-0.020		
	(0.037)	(0.091)		
Information + Subsidy			0.045	0.040**
			(0.030)	(0.018)
Information + Box			0.171***	0.022
			(0.053)	(0.016)
N	1077	1076	1077	1077

Notes: This table shows result of double-selection lasso linear regression specification. Standard errors are robust and clustered by square grid area, reported in parentheses. February registered means number of visit in February to any registered physician. February quack means number of quack visits in two month. February govt means any government hospital visit in February. and February pvt means any private clinic or chamber visit in February. All include covariates of member age, education, religion and marital status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.D Filed operation

Figure 3.D.8: BMI measurement usage during study

(a) (b)

Notes: Both control and treatment households receive free height, weight and BMI measurement.

Figure 3.D.9: Basic health-care service- BP check

Notes: Both control and treatment households receive free blood pressure measurement, and BMI measurement.

Figure 3.D.10: Basic health-care service- diabetic check

Notes: Free Blood glucose measurement has been done among the respondents.

Bibliography

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*, 72(1):1–19.

Abarca-Gómez, L., Abdeen, Z. A., Hamid, Z. A., Abu-Rmeileh, N. M., Acosta-Cazares, B., Acuin, C., Adams, R. J., Aekplakorn, W., Afsana, K., Aguilar-Salinas, C. A., , et al. (2017). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128· 9 million children, adolescents, and adults. *The lancet*, 390(10113):2627–2642.

Abdullah, M. and Wheeler, E. F. (1985). Seasonal variations, and the intra-household distribution of food in a Bangladeshi village. *The American Journal of Clinical Nutrition*, 41(6):1305–1313.

Adams, A. and Andrew, A. (2019). Preferences and beliefs in the marriage market for young brides. Working Paper W19/05, IFS Working Papers.

Adams, A. M., Ahmed, T., Arifeen, S. E., Evans, T. G., Huda, T., and Reichenbach, L. (2013). Innovation for universal health coverage in Bangladesh: a call to action. *The Lancet*, 382(9910):2104–2111.

Agarwal, B. (1997). "Bargaining" and Gender Relations: Within and Beyond the Household. *Feminist Economics*, 3(1):1–51.

Ahmed, I., Rabbi, M. B., and Sultana, S. (2019). Antibiotic resistance in Bangladesh: A systematic review. *International Journal of Infectious Diseases*, 80:54–61.

Ahmed, S. M., Evans, T. G., Standing, H., and Mahmud, S. (2013). Harnessing pluralism for better health in Bangladesh. *The Lancet*, 382(9906):1746–1755.

Ahmed, S. M. and Hossain, M. A. (2007). Knowledge and practice of unqualified and semi-qualified allopathic providers in rural Bangladesh: Implications for the HRH problem. *Health Policy*, 84(2):332–343.

Ahmed, S. M., Hossain, M. A., RajaChowdhury, A. M., and Bhuiya, A. U. (2011). The health workforce crisis in Bangladesh: shortage, inappropriate skill-mix and inequitable distribution. *Human Resources for Health*, 9(1):3.

Ahmed, S. M. and Islam, Q. S. (2012). Availability and Rational Use of Drugs in Primary Healthcare Facilities Following the National Drug Policy of 1982: Is Bangladesh on Right Track? *Journal of Health, Population, and Nutrition*, 30(1):99–108.

Ahmed, S. M., Naher, N., Hossain, T., and Rawal, L. B. (2017). Exploring the status of retail private drug shops in Bangladesh and action points for developing an accredited drug shop model: a facility based cross-sectional study. *Journal of Pharmaceutical Policy and Practice*, 10(1):21.

Ahmed, T., Mahfuz, M., Ireen, S., Ahmed, A. S., Rahman, S., Islam, M. M., Alam, N., Hossain, M. I., Rahman, S. M., Ali, M. M., Choudhury, F. P., and Cravioto, A. (2012). Nutrition of Children and Women in Bangladesh: Trends and Directions for the Future. *Journal of Health, Population, and Nutrition*, 30(1):1–11.

Ahuja, A. and Ostermann, S. L. (2016). Crossing Caste Boundaries in the Modern Indian Marriage Market. *Studies in Comparative International Development*, 51(3):365–387.

Akhtar, Z., Mah-E-Muneer, S., Rashid, M. M., Ahmed, M. S., Islam, M. A., Chowdhury, S., Khan, Z., Hassan, M. Z., Islam, K., Parveen, S., Debnath, N., Rahman, M., and Chowdhury, F. (2021). Antibiotics Use and Its Knowledge in the Community: A Mobile Phone Survey during the COVID-19 Pandemic in Bangladesh. *Antibiotics*, 10(9):1052.

Alston, M., Whittenbury, K., Haynes, A., and Godden, N. (2014). Are climate challenges reinforcing child and forced marriage and dowry as adaptation strategies in the context of Bangladesh? *Women's Studies International Forum*, 47:137–144.

Amin, R. and Mariam, A. G. (1987). Son preference in Bangladesh: an emerging barrier to fertility regulation. *Journal of Biosocial Science*, 19(2):221–228.

Andaleeb, S. S., Siddiqui, N., and Khandakar, S. (2007). Patient satisfaction with health services in Bangladesh. *Health Policy and Planning*, 22(4):263–273.

Andrabi, T., Daniels, B., and Das, J. (2023). Human Capital Accumulation and Disasters: Evidence from the Pakistan Earthquake of 2005. *Journal of Human Resources*, 58(4):1057–1096.

Angrist, J. D. and Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics*, 114(2):533–575.

Anukriti, S. and Dasgupta, S. (2017). Marriage Markets in Developing Countries. *IZA Discussion Papers*.

Arnold, B. F., Rerolle, F., Tedijanto, C., Njenga, S. M., Rahman, M., Ercumen, A., Mertens, A., Pickering, A. J., Lin, A., Arnold, C. D., Das, K., Stewart, C. P., Null, C., Luby, S. P., Colford, J. M., Hubbard, A. E., and Benjamin-Chung, J. (2024). Geographic pair matching in large-scale cluster randomized trials. *Nature Communications*, 15(1):1069.

Arnold, F., Choe, M. K., and Roy, T. K. (1998). Son Preference, the Family-Building Process and Child Mortality in India. *Population Studies*, 52(3):301–315.

Arnup, S. J., McKenzie, J. E., Hemming, K., Pilcher, D., and Forbes, A. B. (2017). Understanding the cluster randomised crossover design: a graphical illustration of the components of variation and a sample size tutorial. *Trials*, 18(1):381.

Asadullah, M. N., Islam, K. M. M., and Wahhaj, Z. (2021a). Child marriage, climate vulnerability and natural disasters in coastal Bangladesh. *Journal of Biosocial Science*, 53(6):948–967.

Asadullah, M. N., Mansoor, N., Randazzo, T., and Wahhaj, Z. (2021b). Is son preference disappearing from Bangladesh? *World Development*, 140:105353.

Asian Development Bank (2021). Key Indicators for Asia and the Pacific 2021. Technical report, Asian Development Bank.

Athey, S. and Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, 74(2):431–497.

Attané, I. (2006). Where Have All the Women Gone? *Index on Censorship*, 35(4):116–122.

Azim, M. R., Ifteakhar, K. N., Rahman, M. M., and Sakib, Q. N. (2023). Public knowledge, attitudes, and practices (KAP) regarding antibiotics use and antimicrobial resistance (AMR) in Bangladesh. *Heliyon*, 9(10):e21166.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labor Markets and Poverty in Village Economies*. *The Quarterly Journal of Economics*, 132(2):811–870.

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.

Banerjee, K. (1999). Gender Stratification and the Contemporary Marriage Market in India. *Journal of Family Issues*, 20(5):648–676.

Bangladesh Bureau of Statistics (2022). Population and Housing Census 2022: Preliminary Report. Technical report, Dhaka.

Barari, S., Caria, S., Davola, A., Falco, P., Fetzer, T., Fiorin, S., Hensel, L., Ivchenko, A., Jachimowicz, J., King, G., Kraft-Todd, G., Ledda, A., MacLennan, M., Mutoi, L., Pagani, C., Reutskaja, E., Roth, C., and Slepoi, F. R. (2020). Evaluating COVID-19 Public Health Messaging in Italy: Self-Reported Compliance and Growing Mental Health Concerns.

BBS (2019). Progotir Pathey, Bangladesh Multiple Indicator Cluster Survey 2019, Survey Findings Report. Technical report, Bangladesh Bureau of Statistics (BBS), Dhaka.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–846.

Becker, G. S. (1974). A Theory of Marriage. In *Economics of the Family: Marriage, Children, and Human Capital*, pages 299–351. University of Chicago Press.

Becker, G. S. (1993). *A Treatise on the Family: Enlarged Edition*. Harvard University Press.

Becker, G. S. (2007). Health as human capital: synthesis and extensions. *Oxford Economic Papers*, 59(3):379–410.

Becker, G. S. and Becker, G. S. (2009). *A Treatise on the Family*. Harvard university press.

Bepari, A. K., Rabbi, G., Shaon, H. R., Khan, S. I., Zahid, Z. I., Dalal, K., and Reza, H. M. (2023). Factors Driving Antimicrobial Resistance in Rural Bangladesh: A Cross-Sectional Study on Antibiotic Use-Related Knowledge, Attitude, and Practice Among Unqualified Village Medical Practitioners and Pharmacy Shopkeepers. *Advances in Therapy*, 40(8):3478–3494.

Bharadwaj, P. (2015). Impact of Changes in Marriage Law: Implications for Fertility and School Enrollment. *The Journal of Human Resources*, 50(3):614–654.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2017). Choose to Lose: Health Plan Choices from a Menu with Dominated Option*. *The Quarterly Journal of Economics*, 132(3):1319–1372.

Binka, F. N., Kubaje, A., Adjuik, M., Williams, L. A., Lengeler, C., Maude, G. H., Armah, G. E., Kajihara, B., Adiamah, J. H., and Smith, P. G. (1996). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical medicine & international health: TM & IH*, 1(2):147–154.

Birech, J. (2013). Child marriage: a cultural health phenomenon. *International Journal of Humanities and Social Science*, 3(17):97–103.

Blackwell, C., Goya-Tocchetto, D., and Sturman, Z. (2017). Nudges in the Restroom: How Hand-Washing Can Be Impacted by Environmental Cues.

Blanchflower, D. G. and Oswald, A. J. (2004). Money, sex and happiness: An empirical study. *Scandinavian Journal of Economics*, 106(3):393–415.

Blommaert, A., Marais, C., Hens, N., Coenen, S., Muller, A., Goossens, H., and Beutels, P. (2014). Determinants of between-country differences in ambulatory antibiotic use and antibiotic resistance in Europe: a longitudinal observational study. *Journal of Antimicrobial Chemotherapy*, 69(2):535–547.

Bonna, A. S., Pavel, S. R., Ferdous, J., Khan, S. A., and Ali, M. (2022). Antibiotic resistance: An increasingly threatening but neglected public health challenge in Bangladesh. *International Journal of Surgery Open*, 49:100581.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14(Volume 14, 2022):521–544.

Boserup, E. (1989). *Woman's Role in Economic Development.* Earthscan.

Bousema, T., Stresman, G., Baidjoe, A. Y., Bradley, J., Knight, P., Stone, W., Osoti, V., Makori, E., Owaga, C., Odongo, W., China, P., Shagari, S., Doumbo, O. K., Sauerwein, R. W., Kariuki, S., Drakeley, C., Stevenson, J., and Cox, J. (2016). The Impact of Hotspot-Targeted Interventions on Malaria Transmission in Rachuonyo South District in the Western Kenyan Highlands: A Cluster-Randomized Controlled Trial. *PLOS Medicine*, 13(4).

Brandimonte, M., Einstein, G. O., and McDaniel, M. A., editors (1996). *Prospective memory: Theory and applications.* Prospective memory: Theory and applications. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US.

Brenøe, A. A. (2021). Brothers increase women's gender conformity. *Journal of Population Economics.*

Bullock, R. (2007). Great circle distances and bearings between two locations. *MDT, June*, 5:1–3.

Bundorf, M. K., Polyakova, M., and Tai-Seale, M. (2019). How do Humans Interact with Algorithms? Experimental Evidence from Health Insurance.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Campbell, M. J. and Walters, S. J. (2014). *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research.* John Wiley & Sons.

Card, D., Dobkin, C., and Maestas, N. (2008). The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *American Economic Review*, 98(5):2242–2258.

Card, D. and Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review*, 84(4):772–93.

Carlson, D. L. (2012). Deviations From Desired Age at Marriage: Mental Health Differences Across Marital Status. *Journal of Marriage and Family*, 74(4):743–758.

Cattaneo, M. D., Frandsen, B. R., and Titiunik, R. (2015). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. *Journal of Causal Inference*, 3(1):1–24.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). A Practical Introduction to Regression Discontinuity Designs: Foundations. *Elements in Quantitative and Computational Methods for the Social Sciences*.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2023). A Practical Introduction to Regression Discontinuity Designs: Extensions.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation Testing Based on Density Discontinuity. *The Stata Journal: Promoting communications on statistics and Stata*, 18(1):234–261.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016). Inference in Regression Discontinuity Designs under Local Randomization. *The Stata Journal: Promoting communications on statistics and Stata*, 16(2):331–367.

Chanda, P. and Joshi, S. R. (2022). Understanding the Small World: The Microbes. In Verma, P., editor, *Industrial Microbiology and Biotechnology*, pages 1–61. Springer, Singapore.

Chao, F., Masquelier, B., You, D., Hug, L., Liu, Y., Sharrow, D., Rue, H., Ombao, H., Alkema, L., Cao, B., Gaigbe-Togbe, V., Spoorenberg, T., Strong, K. L., Suzuki, E., and You, D. (2023). Sex differences in mortality among children, adolescents, and young people aged 0–24 years: a systematic assessment of national, regional, and global trends from 1990 to 2021. *The Lancet Global Health*, 11(10):e1519–e1530.

Chari, A., Heath, R., Maertens, A., and Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal of Development Economics*, 127:42–55.

Chen, L. C., Huq, E., and D'Souza, S. (1981). Sex Bias in the Family Allocation of Food and Health Care in Rural Bangladesh. *Population and Development Review*, 7(1):55–70.

Chetty, R., Friedman, J. N., and Saez, E. (2013). Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, 103(7):2683–2721.

Chinkhumba, J., Godlonton, S., and Thornton, R. (2014). The Demand for Medical Male Circumcision. *American Economic Journal: Applied Economics*, 6(2):152–177.

Chowdhury, F. D. (2004). The socio-cultural context of child marriage in a Bangladeshi village. *International Journal of Social Welfare*, 13(3):244–253.

Ciancio, A., Delavande, A., Kohler, H.-P., and Kohler, I. V. (2024). Mortality Risk Information, Survival Expectations and Sexual Behaviours. *The Economic Journal*, 134(660):1431–1464.

Cole, T. J. (2012). The development of growth references and growth charts. *Annals of human biology*, 39(5):382–394.

Collin, M. and Talbot, T. (2023). Are age-of-marriage laws enforced? Evidence from developing countries. *Journal of Development Economics*, 160:102950.

Cools, A. and Patacchini, E. (2019). The brother earnings penalty. *Labour Economics*, 58:37–51.

Corwin, L. A. (1977). Caste, Class and the Love-Marriage: Social Change in India. *Journal of Marriage and Family*, 39(4):823–831.

Dadgostar, P. (2019). Antimicrobial Resistance: Implications and Costs. *Infection and Drug Resistance*, 12:3903–3910.

Dahl, G. B. and Moretti, E. (2008). The Demand for Sons. *The Review of Economic Studies*, 75(4):1085–1120.

Das, S., Mia, M. N., Hanifi, S. M. A., Hoque, S., and Bhuiya, A. (2017). Health literacy in a community with low levels of education: findings from Chakaria, a rural area of Bangladesh. *BMC Public Health*, 17(1):203.

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(2):153–187.

de Onis, M., Onyango, A. W., Borghi, E., Siyam, A., Nishida, C., and Siekmann, J. (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization*, 85(9):660–667.

Delisle, H. F. (2008). Poverty: the double burden of malnutrition in mothers and the intergenerational impact. *Annals of the New York Academy of Sciences*, 1136:172–184.

Diamond-Smith, N., Luke, N., and Mcgarvey, S. (2008). 'Too many girls, too much dowry': son preference and daughter aversion in rural Tamil Nadu, India. *Culture, health & sexuality*, 10(7):697–708.

Donner, A. (1998). Some aspects of the design and analysis of cluster randomization trials. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47(1):95–113.

Dube, L. (1997). *Women and kinship: perspectives on gender in South and South-East Asia*. United Nations University Press, Tokyo ; New York.

Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic Literature*, 50(4):1051–1079.

Duflo, E., Dupas, P., and Kremer, M. (2015). Education, HIV, and Early Fertility: Experimental Evidence from Kenya. *American Economic Review*, 105(9):2757–2797.

Dupas, P. (2011). Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics*, 3(1):1–34.

Dupas, P. and Miguel, E. (2017). Impacts and Determinants of Health Levels in Low-Income Countries. In *Handbook of Economic Field Experiments*, volume 2, pages 3–93. Elsevier.

D'Souza, A. and Tandon, S. (2019). Intrahousehold Nutritional Inequities in Rural Bangladesh. *Economic Development and Cultural Change*, 67(3):625–657.

Eilat-Adar, S., Periquito, I., Mo, A., Zeev, A., Golan, R., and Gunaratnam, N. (2021). Nutritional Intervention at a Girl's Orphanage in Sri Lanka Decreased Stunting after One Year. *Health*, 13(1):60–67.

Eldridge, S. and Kerry, S. (2012). *A Practical Guide to Cluster Randomised Trials in Health Services Research*. John Wiley & Sons.

Emanuel, I. (1986). Maternal health during childhood and later reproductive performance. *Annals of the New York Academy of Sciences*, 477:27–39.

Field, E. and Ambrus, A. (2008). Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh. *Journal of Political Economy*, 116(5):881–930.

Filmer, D., Friedman, J., and Schady, N. (2008). Development, Modernization, and Son Preference in Fertility Decisions. SSRN Scholarly Paper ID 1267077, Social Science Research Network, Rochester, NY.

Fong, G. T., Hammond, D., and Hitchman, S. C. (2009). The impact of pictures on the effectiveness of tobacco warnings. *Bulletin of the World Health Organization*, 87(8):640–643.

Frandsen, B. R. (2017). Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete. In *Regression Discontinuity Designs*, volume 38 of *Advances in Econometrics*, pages 281–315. Emerald Publishing Limited.

Gaffney-Rhys, R. (2011). International law as an instrument to combat child marriage. *The International Journal of Human Rights*, 15(3):359–373.

Gerend, M. A. and Sias, T. (2009). Message framing and color priming: How subtle threat cues affect persuasion. *Journal of Experimental Social Psychology*, 45(4):999–1002.

Gill, K., Pande, R., and Malhotra, A. (2007). Women deliver for development. *The Lancet*, 370(9595):1347–1357.

Gina Kennedy, T. B. a. M. D. (2011). *Guidelines for measuring household and individual dietary diversity*. FAO, Rome, Italy.

Giné, X., Karlan, D., and Zinman, J. (2010). Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. *American Economic Journal: Applied Economics*, 2(4):213–235.

Gove, W. R., Hughes, M., and Style, C. B. (1983). Does marriage have positive effects on the psychological well-being of the individual? *Journal of health and social behavior*, pages 122–131.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):223–255.

Group, W. B. (2019). World development indicators 2019. Technical report, World Bank.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American economic review*, pages 622–641.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Designing Information Provision Experiments. *Journal of Economic Literature*, 61(1):3–40.

Hafeez, N. and Quintana-Domeque, C. (2018). Son Preference and Gender-Biased Breastfeeding in Pakistan. *Economic Development and Cultural Change*, 66(2):179–215.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, 69(1):201–209.

Hallsworth, M., Chadborn, T., Sallis, A., Sanders, M., Berry, D., Greaves, F., Clements, L., and Davies, S. C. (2016). Provision of social norm feedback to high prescribers of antibiotics in general practice: a pragmatic national randomised controlled trial. *The Lancet*, 387(10029):1743–1752.

Hamid, S. A., Ahsan, S. M., and Begum, A. (2014). Disease-Specific Impoverishment Impact of Out-of-Pocket Payments for Health Care: Evidence from Rural Bangladesh. *Applied Health Economics and Health Policy*, 12(4):421–433.

Hayes, R. J. and Moulton, L. H. (2017). *Cluster Randomised Trials*. Chapman and Hall/CRC, New York, 2 edition.

Heckman, J., Ichimura, H., Smith, J., and Todd, P. (1998). Characterizing Selection Bias Using Experimental Data. *Econometrica*, 66(5):1017–1098.

Helen Keller International, J. P. G. S. o. P. H. (2016). State of food security and nutrition in Bangladesh: 2014.

Hemming, K. and Taljaard, M. (2023). Key considerations for designing, conducting and analysing a cluster randomized trial. *International Journal of Epidemiology*, 52(5):1648–1658.

Herrera, H., Rebato, E., Arechabaleta, G., Lagrange, H., Salces, I., and Susanne, C. (2003). Body mass index and energy intake in Venezuelan University students. *Nutrition Research*, 23(3):389–400.

Hoque, R., Ahmed, S. M., Naher, N., Islam, M. A., Rousham, E. K., Islam, B. Z., and Hassan, S. (2020). Tackling antimicrobial resistance in Bangladesh: A scoping review of policy and practice in human, animal and environment sectors. *PLoS ONE*, 15(1):e0227947.

Hosegood, V. and Campbell, O. M. (2003). Body mass index, height, weight, arm circumference, and mortality in rural Bangladeshi women: a 19-y longitudinal study. *The American Journal of Clinical Nutrition*, 77(2):341–347.

Hossain, M., Asadullah, M. N., and Kambhampati, U. (2021). Women's empowerment and gender-differentiated food allocation in Bangladesh. *Review of Economics of the Household*, 19(3):739–767.

Hunersen, K., Attal, B., Jeffery, A., Metzler, J., Alkibsi, T., Elnakib, S., and Robinson, W. C. (2021). Child Marriage in Yemen: A Mixed Methods Study in Ongoing Conflict and Displacement. *Journal of Refugee Studies*, 34(4):4551–4571.

Imai, K., King, G., and Nall, C. (2009). The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation. *Statistical Science*, 24(1).

Imbens, G. and Kalyanaraman, K. (2012). Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.

International Food Policy Research Institute (IFPRI) (2020). Bangladesh Integrated Household Survey (BIHS) 2018-2019.

Islam, A. and Biswas, T. (2014). Health System in Bangladesh: Challenges and Opportunities. *American Journal of Health Research*, 2:366.

Jain, P., Bepari, A. K., Sen, P. K., Rafe, T., Imtiaz, R., Hossain, M., and Reza, H. M. (2021). High prevalence of multiple antibiotic resistance in clinical E. coli isolates from Bangladesh and prediction of molecular resistance determinants using WGS of an XDR isolate. *Scientific Reports*, 11(1):22859.

Jakiela, P., Ozier, O., Fernald, L., and Knauer, H. (2020). *Big Sisters*. Policy Research Working Papers. The World Bank.

James, W. P., Ferro-Luzzi, A., and Waterlow, J. C. (1988). Definition of chronic energy deficiency in adults. Report of a working party of the International Dietary Energy Consultative Group. *European Journal of Clinical Nutrition*, 42(12):969–981.

Jayachandran, S. (2015). The Roots of Gender Inequality in Developing Countries. *Annual Review of Economics*, 7(1):63–88.

Jayachandran, S. and Pande, R. (2017). Why Are Indian Children So Short? The Role of Birth Order and Son Preference. *American Economic Review*, 107(9):2600–2629.

Joarder, T., Chaudhury, T. Z., and Mannan, I. (2019). Universal Health Coverage in Bangladesh: Activities, Challenges, and Suggestions. *Advances in Public Health*, 2019:e4954095.

John, A. and Orkin, K. (2022). Can Simple Psychological Interventions Increase Preventive Health Investment? *Journal of the European Economic Association*, 20(3):1001–1047.

Kabeer, N., Huq, L., and Mahmud, S. (2014). Diverging Stories of "Missing Women" in South Asia: Is Son Preference Weakening in Bangladesh? *Feminist Economics*, 20(4):138–163.

Kennedy, E., Binder, G., Humphries-Waa, K., Tidhar, T., Cini, K., Comrie-Thomson, L., Vaughan, C., Francis, K., Scott, N., Wulan, N., Patton, G., and Azzopardi, P. (2020). Gender inequalities in health and wellbeing across the first two decades of life: an analysis of 40 low-income and middle-income countries in the Asia-Pacific region. *The Lancet Global Health*, 8(12):e1473–e1488.

Kennedy, G., Ballard, T., and Dop, M. C. (2011). *Guidelines for Measuring Household and Individual Dietary Diversity*. Food and Agriculture Organization of the United Nations.

Khan, J. A. M., Ahmed, S., and Evans, T. G. (2017). Catastrophic healthcare expenditure and poverty related to out-of-pocket payments for healthcare in Bangladesh—an estimation of financial risk protection of universal health coverage. *Health Policy and Planning*, 32(8):1102–1110.

Khan, M. A., Khan, N., Rahman, O., Mustagir, G., Hossain, K., Islam, R., and Khan, H. T. A. (2021). Trends and projections of under-5 mortality in Bangladesh including the effects of maternal high-risk fertility behaviours and use of healthcare services. *PLOS ONE*, 16(2):e0246210.

Khatun, W., Alam, A., Rasheed, S., Huda, T. M., and Dibley, M. J. (2018). Exploring the intergenerational effects of undernutrition: association of maternal height with neonatal, infant and under-five mortality in Bangladesh. *BMJ global health*, 3(6):e000881.

Kiecolt-Glaser, J. K. and Newton, T. L. (2001). Marriage and health: his and hers. *Psychological bulletin*, 127(4):472.

Klein, E. Y., Van Boeckel, T. P., Martinez, E. M., Pant, S., Gandra, S., Levin, S. A., Goossens, H., and Laxminarayan, R. (2018). Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. *Proceedings of the National Academy of Sciences*, 115(15):E3463–E3470.

Koenig, M. A. and D'Souza, S. (1986). Sex differences in childhood mortality in rural Bangladesh. *Social Science & Medicine (1982)*, 22(1):15–22.

Koya, S. F., Ganesh, S., Selvaraj, S., Wirtz, V. J., Galea, S., and Rockers, P. C. (2024). Determinants of private-sector antibiotic consumption in India: findings from a quasi-experimental fixed-effects regression analysis using cross-sectional time-series data, 2011–2019. *Scientific Reports*, 14(1):5052.

Kremer, M., Rao, G., and Schilbach, F. (2019). Behavioral development economics. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2 of *Handbook of Behavioral Economics - Foundations and Applications 2*, pages 345–458. North-Holland.

Kurpad, A. V., Muthayya, S., and Vaz, M. (2005). Consequences of inadequate food energy and negative energy balance in humans. *Public Health Nutrition*, 8(7A):1053–1076.

Lawson, M. L., Schnepf, R., and Cook, N. (2016). The Obama Administration's Feed the Future Initiative. Technical report.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674.

Lee, D. S. and Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2):281–355.

Li, H., Calder, C. A., and Cressie, N. (2007). Beyond Moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model. *Geographical Analysis*, 39(4):357–375.

Loh, E. S. (1996). Productivity differences and the marriage wage premium for white males. *Journal of Human Resources*, pages 566–589.

Lovell, M. R., Forder, P. M., Stockler, M. R., Butow, P., Briganti, E. M., Chye, R., Goldstein, D., and Boyle, F. M. (2010). A Randomized Controlled Trial of a Standardized Educational Intervention for Patients with Cancer Pain. *Journal of Pain and Symptom Management*, 40(1):49–59.

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

Mannan, M. A. (2013). Access to Public Health Facilities in Bangladesh: A Study on Facility Utilisation and Burden of Treatment. *The Bangladesh Development Studies*, 36(4):25–80.

Mansour, H. and McKinnish, T. (2018). Marriage-Market Search and Sorting: Explanations and Evidence. In Averett, S. L., Argys, L. M., and Hoffman, S. D., editors, *The Oxford Handbook of Women and the Economy*. Oxford University Press.

Marphatia, A. A., Cole, T. J., Grijalva-Eternod, C., and Wells, J. C. K. (2016). Associations of gender inequality with child malnutrition and mortality across 96 countries. *Global Health, Epidemiology and Genomics*, 1:e6.

Maswikwa, B., Richter, L., Kaufman, J., and Nandi, A. (2015). Minimum Marriage Age Laws and the Prevalence Of Child Marriage and Adolescent Birth: Evidence from Sub-Saharan Africa. *International Perspectives on Sexual and Reproductive Health*, 41(2):58–68.

Mathew, T. K., Zubair, M., and Tadi, P. (2024). Blood Glucose Monitoring. In *StatPearls*. StatPearls Publishing, Treasure Island (FL).

McDaniel, M. A. and Einstein, G. O. (2007). *Prospective memory: An overview and synthesis of an emerging field*. Prospective memory: An overview and synthesis of an emerging field. Sage Publications, Inc, Thousand Oaks, CA, US.

McGavock, T. (2021). Here waits the bride? The effect of Ethiopia's child marriage law. *Journal of Development Economics*, 149:102580.

Meeker, D., Linder, J. A., Fox, C. R., Friedberg, M. W., Persell, S. D., Goldstein, N. J., Knight, T. K., Hay, J. W., and Doctor, J. N. (2016). Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing Among Primary Care Practices: A Randomized Clinical Trial. *JAMA*, 315(6):562–570.

Miguel, E. and Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1):159–217.

Milazzo, A. (2018). Why are adult women missing? Son preference and maternal survival in India. *Journal of Development Economics*, 134:467–484.

Mitra, A. K., Rahman, M. M., and Fuchs, G. J. (2000). Risk factors and gender differentials for death among children hospitalized with diarrhoea in Bangladesh. *Journal of Health, Population, and Nutrition*, 18(3):151–156.

Molla, A. A. and Chi, C. (2017). Who pays for healthcare in Bangladesh? An analysis of progressivity in health systems financing. *International Journal for Equity in Health*, 16:167.

Morgan, D. J., Okeke, I. N., Laxminarayan, R., Perencevich, E. N., and Weisenberg, S. (2011). Non-prescription antimicrobial use worldwide: a systematic review. *The Lancet Infectious Diseases*, 11(9):692–701.

Mourtada, R., Schlecht, J., and DeJong, J. (2017). A qualitative study exploring child marriage practices among Syrian conflict-affected populations in Lebanon. *Conflict and Health*, 11(1):27.

Murray, C. J. L., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, and and others (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, 399(10325):629–655.

Narayan, J., John, D., and Ramadas, N. (2019). Malnutrition in India: status and government initiatives. *Journal of Public Health Policy*, 40(1):126–141.

Nevo, I., Fitzpatrick, M., Thomas, R.-E., Gluck, P. A., Lenchus, J. D., Arheart, K. L., and Birnbach, D. J. (2010). The efficacy of visual cues to improve hand hygiene compliance. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 5(6):325–331.

Okojie, C. E. (1994). Gender inequalities of health in the Third World. *Social Science & Medicine (1982)*, 39(9):1237–1247.

O'Neill, J. (2016). Tackling drug-resistant infections globally: final report and recommendations. Report, Government of the United Kingdom.

Otoo-Oyortey, N. and Pobi, S. (2003). Early marriage and poverty: exploring links and key policy issues. *Gender & Development*, 11(2):42–51.

Pande, R., Malhotra, A., Mathur, S., Mehta, M., Malhotra, A., Lycette, M. A., Kambou, S. D., Magar, V., Gay, J., and Lary, H. (2006). Son preference and daughter neglect in India.

Paul, P. (2019). Effects of education and poverty on the prevalence of girl child marriage in India: A district–level analysis. *Children and Youth Services Review*, 100:16–21.

Pednekar, P. and Peterson, A. (2018). Mapping pharmacy deserts and determining accessibility to community pharmacy services for elderly enrolled in a State Pharmaceutical Assistance Program. *PLOS ONE*, 13(6):e0198173.

Peter, N., Lundborg, P., Mikkelsen, S., and Webbink, D. (2018). The effect of a sibling's gender on earnings and family formation. *Labour Economics*, 54:61–78.

Prasetyo, D. B., Sofyan, L., Muchtar, P. A., and Dewi, D. F. (2022). Nudging to handwash during the pandemic – The use of visual priming and salience. *Analyses of Social Issues and Public Policy*, 22(3):836–856.

Pravossoudovitch, K., Cury, F., Young, S. G., and Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red–danger association. *Ergonomics*, 57(4):503–510.

Prestinaci, F., Pezzotti, P., and Pantosti, A. (2015). Antimicrobial resistance: a global multifaceted phenomenon. *Pathogens and Global Health*, 109(7):309–318.

Rahman, M. A., Rahman, M. S., Rahman, M. A., Szymlek-Gay, E. A., Uddin, R., and Islam, S. M. S. (2021). Prevalence of and factors associated with anaemia in women of reproductive age in Bangladesh, Maldives and Nepal: Evidence from nationally-representative survey data. *PLOS ONE*, 16(1):e0245335.

Rahman, M. M., Islam, M. R., Rahman, M. S., Hossain, F., Alam, A., Rahman, M. O., Jung, J., and Akter, S. (2022). Forgone healthcare and financial burden due to out-of-pocket payments in Bangladesh: a multilevel analysis. *Health Economics Review*, 12(1):5.

Rahman, T., Gasbarro, D., Alam, K., and Alam, K. (2024). Rural–urban disparities in household catastrophic health expenditure in Bangladesh: a multivariate decomposition analysis. *International Journal for Equity in Health*, 23(1):43.

Rai, P., Paudel, I. S., Ghimire, A., Pokharel, P. K., Rijal, R., and Niraula, S. R. (2014). Effect of gender preference on fertility: cross-sectional study among women of Tharu community from rural area of eastern region of Nepal. *Reproductive Health*, 11(1):15.

Raj, A., Jackson, E., and Dunham, S. (2018). Girl Child Marriage: A Persistent Global Women's Health and Human Rights Violation. In Choudhury, S., Erausquin, J. T., and Withers, M., editors, *Global Perspectives on Women's Sexual and Reproductive Health Across the Lifecourse*, pages 3–19. Springer International Publishing, Cham.

Raj, A., Saggurti, N., Balaiah, D., and Silverman, J. G. (2009). Prevalence of Child Marriage and its Impact on the Fertility and Fertility Control Behaviors of Young Women in India. *Lancet*, 373(9678):1883–1889.

Ravi, S. and Engler, M. (2015). Workfare as an Effective Way to Fight Poverty: The Case of India's NREGS. *World Development*, 67(C):57–71.

Rice, T. (2013). The Behavioral Economics of Health and Health Care. *Annual Review of Public Health*, 34(Volume 34, 2013):431–447.

Rodgers, H., Atkinson, C., Bond, S., Suddes, M., Dobson, R., and Curless, R. (1999). Randomized Controlled Trial of a Comprehensive Stroke Education Program for Patients and Caregivers. *Stroke*, 30(12):2585–2591.

Rush, D. (2000). Nutrition and maternal mortality in the developing world. *The American Journal of Clinical Nutrition*, 72(1):212S–240S.

Russell, W. R., Baka, A., Björck, I., Delzenne, N., Gao, D., Griffiths, H. R., Hadjilucas, E., Juvonen, K., Lahtinen, S., Lansink, M., Loon, L. V., Mykkänen, H., Östman, E., Riccardi, G., Vinoy, S., and Weickert, M. O. (2016). Impact of Diet Composition on Blood Glucose Regulation. *Critical Reviews in Food Science and Nutrition*, 56(4):541–590.

Sachdev, H. S., Porwal, A., Sarna, A., Acharya, R., Ramesh, S., Kapil, U., and Kurpad, A. V. (2021). Intraindividual double-burden of anthropometric undernutrition and "metabolic obesity" in Indian children: a paradox that needs action. *European Journal of Clinical Nutrition*, 75(8):1205–1217.

Santhya, K. G., Ram, U., Acharya, R., Jejeebhoy, S. J., Ram, F., and Singh, A. (2010). Associations between early marriage and young women's marital and reproductive health outcomes: evidence from India. *International Perspectives on Sexual and Reproductive Health*, 36(3):132–139.

Sarker, A. R., Sultana, M., Alam, K., Ali, N., Sheikh, N., Akram, R., and Morton, A. (2021). Households' out-of-pocket expenditure for healthcare in Bangladesh: A health financing incidence analysis. *The International Journal of Health Planning and Management*, 36(6):2106–2117.

Sekhri, S. and Debnath, S. (2014). Intergenerational Consequences of Early Age Marriages of Girls: Effect on Children's Human Capital. *The Journal of Development Studies*, 50(12):1670–1686.

Self, S. and Grabowski, R. (2012). Son Preference, Autonomy and Maternal Health in Rural India. *Oxford Development Studies*, 40(3):305–323.

Sen, A. (2002). More Than 100 Million Women Are Missing. In *Gender and Justice*. Routledge.

Sen, A. and Sengupta, S. (1983). Malnutrition of Rural Children and the Sex Bias. *Economic and Political Weekly*, 18(19/21):855–864.

Sethuraman, K. and Duvvury, N. (2007). The Nexus of Gender Discrimination with Malnutrition: An Introduction. *Economic and Political Weekly*, 42(44):49–53.

Shetty, P. S. and James, W. P. (1994). Body mass index. A measure of chronic energy deficiency in adults. *FAO food and nutrition paper*, 56:1–57.

Subramanian, S. V., Ackerson, L. K., Davey Smith, G., and John, N. A. (2009). Association of Maternal Height With Child Mortality, Anthropometric Failure, and Anemia in India. *JAMA*, 301(16):1691–1701.

Sudhinaraset, M., Ingram, M., Lofthouse, H. K., and Montagu, D. (2013). What Is the Role of Informal Healthcare Providers in Developing Countries? A Systematic Review. *PLOS ONE*, 8(2):e54978.

Sulis, G., Sayood, S., and Gandra, S. (2022). Antimicrobial resistance in low- and middle-income countries: current status and future directions. *Expert Review of Anti-infective Therapy*, 20(2):147–160.

Swindale, A. and Bilinksy, P. (2006). Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide (Version 2). Technical report, Food and Nutrition Technical Assistance III Project (FANTA), Washington, D.C.

Talukder, M., Rob, U., and Noor, F. (2014). Assessment of sex selection in Bangladesh. Technical report, Population Council.

Tangcharoensathien, V., Chanvatik, S., and Sommanustweechai, A. (2018). Complex determinants of inappropriate use of antibiotics. *Bulletin of the World Health Organization*, 96(2):141–144.

Thaler, R. H. and Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.

The World Bank Group (2023). World Development Indicators.

Thistlethwaite, D. L. and Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309–317.

Uecker, J. E. (2012). Marriage and Mental Health among Young Adults. *Journal of Health and Social Behavior*, 53(1):67–83.

Umberson, D. (1992). Gender, marital status and the social control of health behavior. *Social Science & Medicine*, 34(8):907–917.

Umberson, D., Williams, K., Powers, D. A., Liu, H., and Needham, B. (2006). You Make Me Sick: Marital Quality and Health Over the Life Course. *Journal of Health and Social Behavior*, 47(1):1–16.

UNICEF (2005). *Early Marriage A Harmful Traditional Practice A Statistical Exploration 2005*. UNICEF.

UNICEF (2023). Ending child marriage: A profile of progress in India, 2023 update. Technical report, New York.

UNICEF (2024). Child marriage data.

Verbrugge, L. M. (1979). Marital Status and Health. *Journal of Marriage and Family*, 41(2):267–285.

Waite, L. J. and Gallagher, M. (2001). *The case for marriage: Why married people are happier, healthier, and better off financially*. Random House Digital, Inc.

Wells, J. C., Sawaya, A. L., Wibaek, R., Mwangome, M., Poullas, M. S., Yajnik, C. S., and Demaio, A. (2020). The double burden of malnutrition: aetiological pathways and consequences for health. *The Lancet*, 395(10217):75–88.

Wheeler, M. L. and Pi-Sunyer, F. X. (2008). Carbohydrate Issues: Type and Amount. *Journal of the American Dietetic Association*, 108(4, Supplement):S34–S39.

WHO Global InfoBase team (2005). The SuRF Report 2. Surveillance of chronic disease risk factors: country level data and comparable estimates. Technical report, World Health Organization, Geneva.

Williams, K. and Umberson, D. (2004). Marital Status, Marital Transitions, and Health: A Gendered Life Course Perspective. *Journal of health and social behavior*, 45(1):81–98.

Williamson, N. E. (1976). *Sons Or Daughters: Cross Cultural Survey Parent Preferences*. SAGE Publications.

Wodon, Q., Tavares, P., Fiala, O., Le Nestour, A., and Wise, L. (2017). Ending child marriage: legal age for marriage, illegal child marriages, and the need for interventions. *London and Washington, DC: Save the Children and The World Bank*.

World Health Organization (2015). *Bangladesh health system review*, volume 5. WHO Regional Office for the Western Pacific.

Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J., and Murray, C. J. L. (2003). Household catastrophic health expenditure: a multicountry analysis. *Lancet (London, England)*, 362(9378):111–117.

Zhang, Y., Wang, H., Wang, X., Liu, M., Wang, Y., Wang, Y., and Zhou, H. (2019). The association between urbanization and child height: a multilevel study in China. *BMC Public Health*, 19(1):569.

Özaltin, E., Hill, K., and Subramanian, S. V. (2010). Association of Maternal Stature With Offspring Mortality, Underweight, and Stunting in Low- to Middle-Income Countries. *JAMA*, 303(15):1507–1516.