I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13694

Protecting Girls from Droughts with Social Safety Nets

Jagori Chatterjee
Joshua D. Merfeld

SEPTEMBER 2020

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13694

Protecting Girls from Droughts with Social Safety Nets

Jagori Chatterjee
Furman University

Joshua D. Merfeld
KDI School of Public Policy and Management and IZA

SEPTEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

Protecting Girls from Droughts with Social Safety Nets*

This paper revisits the relationship between agricultural productivity shocks and the infant sex ratio in India and investigates how this relationship changes when households have access to government-provided employment opportunities outside of agriculture. When a household's preference for sons coincides with adverse agricultural productivity shocks, previous research shows that households tend to disproportionately reduce investments (prenatal and postnatal) in their female children. This behavior leads to a relatively more balanced sex ratio in good rainfall years and a more skewed sex ratio (in favor of boys) in inadequate rainfall years. In a deviation from past work, we find evidence of this primarily through prenatal channels in modern India. We then show that a workfare program that decouples both wages and consumption from rainfall attenuates the relationship between rainfall and the infant sex ratio. Using a back-of-the-envelope calculation, we find that the program could have saved at least 0.7 million girls – relative to boys – if the government had implemented it in 2001 to 2005. Additional results on postnatal channels show substantial impacts on the long-term health outcomes of surviving girls, as rainfall no longer differentially affects girls' height-for-age, relative to boys', following the program's implementation.

JEL Classification: H53, I15, I38, O12

Keywords: sex ratio, child health, consumption smoothing, workfare program, India

Corresponding author:
Joshua D. Merfeld
KDI School of Public Policy and Management
263 Namsejong-ro
Sejong-si
South Korea
E-mail: merfeld@kdis.ac.kr

* We are grateful to Rachel M. Heath for her encouragement and guidance since ideation. We thank Chris Anderson, Urvashi Jain, James D. Long, Jonathan Morduch, Alejandra Ramos, Xu Tan, Jessie Wang and the participants at numerous conferences and seminars for helpful comments and suggestions. We also thank the University of Washington's South Asia Library for access to the National Sample Surveys. Chatterjee thanks the University of Washington for research support through the Grover and Creta Ensley Graduate Fellowship for Economic Policy.

1 Introduction

Girls and boys are treated differently in countries where households exhibit a strong preference for sons. In extreme cases, this discrimination leads to the sex selection of children at early ages[1] through postnatal neglect or prenatal sex-selective abortions[2]. An undesirable outcome of this discrimination is a highly skewed sex ratio, with boys vastly outnumbering girls. The 2011 Census of India reports 919 girls for every 1000 boys under the age of six. In contrast, in countries where families provide equal care for both daughters and sons, the sex ratio is expected to be approximately even (Sen, 1992).

Economic development might help improve the sex ratio over time. Nevertheless, Figure A1 shows that, in India's case, the child sex ratio has only worsened over the last half-century, despite an average annual GDP growth rate of about 5 percent. Indeed, the decrease in fertility following decades of economic development may worsen the sex ratio, as families have fewer children than before but still desire a boy (Jayachandran, 2017). The continued decline in the sex ratio in the face of sustained economic growth suggests an urgent need for a better understanding of the determinants of sex selection in India if we aim to design policies to combat this decline. We contribute to this literature by documenting the impact of agricultural productivity shocks on sex-selective abortions and showing that government policy may be able to alleviate some aspects of the problem.

A notable driver of sex selection is transitory economic shocks. In particular, income shocks in developing countries may exacerbate the already skewed sex ratio. In India, primarily an agrarian society with strong son preference, Rose (1999) finds that the probability a child born during a given year is a girl is increasing in rainfall. In other words,

[1]Previous work has argued this as one of the leading causes of unbalanced sex ratios in South and Southeast Asian countries. In his seminal work, Sen (1990) estimated that more than 100 million women were "missing" worldwide. More recent estimates suggest that this number has been steadily increasing over time (reaching 126 million in 2010) and that India and China account for most of this deficit (Bongaarts and Guilmoto, 2015).

[2]More so the latter following the introduction of reliable ultrasound technology in the 1980s (Bhalotra and Cochrane, 2010; Chen et al., 2013).

when income is higher, a randomly selected newborn is more likely to be female. Due to the lack of formal mechanisms to cope with adverse agricultural shocks, households may reduce investments in their less desired (female) children or selectively continue with a pregnancy if the child is a boy, and terminate the pregnancy if the child is a girl to help smooth consumption. Moreover, even for fetuses carried to term, there is a gender-gap in prenatal investments in societies with strong son preference, and bad economic times may exacerbate this.[3]

If transitory economic shocks do indeed lead to such outcomes, some sex-selective abortions may be the result of a lack of formal – or even informal – risk mitigation mechanisms. The availability of these mechanisms could enable consumption-smoothing during bad times and potentially attenuate the gender-differentiated mortality effects, as well as underinvestments more generally. Therefore, a general conclusion in this body of literature is that moving forward, an important policy measure to improve women's lives is the provision of consumption-smoothing mechanisms (Rose, 1999; Maccini and Yang, 2009). However, despite the growing number of risk-coping programs implemented in developing countries today, there remains a gap in the literature that empirically tests whether these policies indeed reduce female mortality during adverse income shocks. To our knowledge, this is the first paper that formally examines the effects of one such program on female-specific child health outcomes. Specifically, in the context of infant mortality in India, we provide the first evidence of how the relationship between agricultural productivity shocks and female mortality attenuates when a national workfare program enables households to relieve income risks during bad years.

We develop a simple model of son preference, economic shocks, and consumption smoothing. The model shows that the high net cost of daughters in later life implies

[3]For example, women who are pregnant with a boy are more likely to visit antenatal clinics (Bharadwaj and Lakdawala, 2013). Throughout childhood, unequal human capital investments continue through differences in breastfeeding (Jayachandran and Kuziemko, 2011), food allocation (Chen et al., 1981; Das Gupta, 1987), parental time allocation (Barcellos et al., 2014), vaccination (Borooah, 2004; Ganatra and Hirve, 2001), other health-care practices (Ganatra and Hirve, 1994), and education (Song et al., 2006).

that girls' optimal health investment is lower than boys. Therefore, a positive income shock leads to a greater incentive to invest in a girl than a boy. A guaranteed workfare program introduces a new sector where households can supply their labor at a state-level minimum wage. This new sector causes income to increase and stabilize. If the net cost of daughters is high enough, then this income risk mitigation causes optimal investment in girls to increase and in boys to decrease after the workfare program. This change in optimal investments causes the elasticity of health investments with respect to income shocks to decrease for households with daughters and increase for households with sons. Consequently, the gender gap in the elasticity decreases after the workfare program. Empirically, we proxy economic shock with agricultural shocks and show how deviations of district-level rainfall from its long-run average interact with the roll-out of an extensive national workfare program in India, Mahatma Gandhi National Rural Employment Guarantee Scheme (NREGS).

Consistent with previous literature, we find that rainfall continues to be a significant predictor of the gender of a surviving infant in India in the early 2000s. Before the implementation of NREGS, an increase in annual rainfall by one standard deviation (standardized using the district's ten-year rainfall history) significantly increases the probability that an infant born during that year is a girl. However, in contrast to the findings of Rose (1999), who uses data from the 1970s, we do not find that concurrent rainfall is a significant determinant of the gender of older children in the early 2000s. This result is consistent with more recent findings that a large part of the sex selection of children is at the early stages of pregnancy since the advent of reliable ultrasound technology. We then present evidence that the introduction of NREGS attenuates these relationships. A one standard deviation increase in rainfall increases the probability that a child born in a district before NREGS is 5.4 percentage points more likely to be a girl. Following the introduction of the program, this effect is 4.6 percentage points lower for treated districts. This result suggests that there is almost no relationship between agricultural productivity shocks and the sex of an infant following the implementation of NREGS.

Next, we examine the effect of agricultural productivity shocks at the time of birth on the long-run health of surviving children. If parents are underinvesting in girls relative to boys, we would expect to see evidence of this in non-mortality outcomes. In particular, we explore the relationship between rainfall during the year of birth and child anthropometrics by gender. We first confirm that rainfall during the year of birth is a significant predictor of height-for-age for surviving girls in India relative to boys, similar to recent results from Indonesia (Maccini and Yang, 2009). Previous literature suggests that those girls who manage to survive sex selection at birth still receive gender-biased early-life investments. This discrimination in care during the year of birth and subsequent years is likely to have long-run gender-gaps in the health of the surviving children. Consistent with this, we find that before NREGS, an increase in annual rainfall by one standard deviation increases the height-for-age of female children by 0.06 standard deviations, relative to male children. Post NREGS, this differential relationship is significantly attenuated, to the point that rainfall has identical effects on height-for-age for both girls and boys.

Heterogeneity by birth order is consistent with previous literature on sex selection. There is no effect of rainfall on the gender of first-born children, consistent with sex selection being less pronounced for first-born children (Bhalotra and Cochrane, 2010). As expected, there is similarly no effect of NREGS on this relationship. Rainfall is strongly associated with the gender for non-first-born children, however, and NREGS appears to attenuate this relationship, though the estimates are imprecisely estimated. It remains to be seen whether the program affects gender and fertility preferences more broadly, as the three-year implementation period makes long-term outcomes more difficult to study.

We calculate the number of girls that NREGS might have saved if the government had implemented it in the years 2001 to 2005. Assuming that the number of boys is unaffected by rainfall and that the number of girls never exceeds the number of boys[4] and using a

[4]The empirical evidence presented in this paper largely supports both of these assumptions.

back-of-the-envelope calculation, we estimate that approximately 283 additional girls per district per year would have survived to one year of age, relative to boys[5]. Over the five years, the policy could have potentially saved 0.7 million girls.

This paper primarily contributes to three strands of existing literature. First, this work fits into the research on how sex selection is affected by changing economic conditions. Previous research finds that improving economic conditions may increase the number of girls born in countries with a strong preference for sons: Bhalotra et al. (2016) show that increases in gold prices increase neonatal mortality of girls relative to boys in India. Qian (2008) finds that increasing female income increases survival rates for girls in China. Almond et al. (2019) show that increases in household income, from land reforms, increases sex-selection in China. Like Rose (1999) and Maccini and Yang (2009), we confirm that transitory shocks are a significant predictor of both the infant sex-ratio and surviving girls' health outcomes in modern India, at least before the implementation of NREGS.

Second, the study contributes to the literature on the effectiveness of different government policies for girls' well-being. Greater political participation of women increases female infant mortality (Kalsi, 2017), improves female children's health (Bhalotra and Clots-Figueras, 2014), and increases their education and aspiration(Beaman et al., 2012). There is mixed evidence on the success of financial incentives offered to families to have daughters (Anukriti, 2018; Balakrishnan, 2017). Importantly, we find that policymakers need not explicitly direct programs towards girls' health to have significant impacts. In this case, attenuating the correlation between local economic conditions and household incomes improves early human capital investments in children, especially girls.

Finally, we add to the growing literature on the effects of NREGS on development outcomes[6], including wages (Imbert and Papp, 2015; Merfeld, 2019), consumption (Jha

[5]This calculation quantifies gains based only on the sex ratio and rainfall, so it is an admittedly crude measure of the program's exact effects on female welfare.

[6]Although there are some increasing concerns about inefficacy and corruption of such programs (Niehaus and Sukhtankar, 2013), though these are not specific to NREGS in India (Kochar et al., 2009).

et al., 2011; Ravi and Engler, 2015), and time allocation decisions (Shah and Steinberg, 2015). Particularly relevant to our work is the findings of Chari et al. (2019) that NREGS increases overall infant mortality among women who were eligible for the program. Our results more specifically contribute to the developing literature on the agricultural risk-mitigation effects of rural workfare programs and the subsequent impact on socio-economic factors. The most relevant to this paper is the work by Santangelo (2019), who shows that NREGS attenuates the pro-cyclical response of local wages, income, and consumption to agricultural productivity shocks in rural India. Therefore, NREGS is an ideal risk-coping policy to study how a disruption in the positive relationship between agricultural productivity shocks and household consumption can affect excess female infant mortality and the gender gap in other health outcomes. Specifically, after NREGS, adverse agricultural shocks have a weaker effect on domestic terrorism (Fetzer, 2014), a lesser unfavorable impact on child education and health (Foster and Gehrke, 2017; Dasgupta, 2013), and decreases dowry deaths (Sarma, 2019)[7]. Overall, the program appears to play a risk-mitigating role, protecting households against unfavorable shocks. Along similar lines, some research also shows that the program allows families to reallocate resources away from low-risk, low-return employment and towards riskier, higher return opportunities (Gehrke, 2017; Merfeld, 2020). Our work contributes to this literature by showing that NRGES decreases the dependence of female infant survival on favorable agricultural shocks–primarily through prenatal channels.

2 Data

In this section, we describe the primary data sources, and sample and variable construction used in this study to test our two predictions derived in Appendix B.

[7]Complementarily, there is mixed evidence on NREGS's impact on attenuating the effects of heat on human capital. Garg et al. (2018) finds that NREGS decreases heat's adverse effects on child education. However, Banerjee and Maharaj (2019) do not find any attenuating effects of NREGS on the adverse effects of heat on infant mortality.

2.1 Agricultural Productivity Shocks

We proxy agricultural productivity shocks using rainfall. India is primarily an agrarian society, and rainfall has documented adverse effects on agricultural yields, labor demand, and wages in the country (Jayachandran, 2006; Kaur, 2019). We use the 0.5-degree by 0.5-degree grid monthly precipitation data from the University of East Anglia's Climate Research Unit (CRU) to construct agricultural productivity shocks. We aggregate the CRU data to annual precipitation and then match the 2001 Census district-centroids[8] to the closest grid in the CRU data to create yearly district-level rainfall. Our primary measure of productivity shock is a rainfall z-score: the deviation of annual district-level rainfall from its long-run mean (using the previous ten years) and scaled by the long-run standard deviation.

2.2 National Rural Employment Guarantee Scheme

The National Rural Employment Guarantee Scheme was implemented in 200 districts since April 2006 (Phase 1), 130 districts starting in June 2007 (Phase 2), and the remaining districts received the program beginning July 2008 (Phase 3)[9]. We use publicly available information on the roll-out of the NREGS at the district-level to create an indicator variable that is equal to one if a district received the program during a year, and zero otherwise[10].

[8]We reclassify all districts to the 2001 Census boundary definitions for this study. In case a district was formed from multiple parent districts, we drop the parents and the newly formed district from the analysis. This methodology leaves 523 in the NSS sample and 313 districts in the IHDS sample for the analysis out of the 593 districts in the 2001 Census.

[9]Some urban districts such as Hyderabad, Kolkata, Chennai, and others were never eligible for NREGS. Therefore, we exclude 16 such 2001 Census urban districts from the analysis.

[10]We assign NREGS treatment to Phase 1 districts from 2007, Phase 2 districts from 2008, and Phase 3 districts from 2009 as the initial implementation of each phase of NREGS missed the dry season at the start of that calendar year, which ends in March, and during which time consumption smoothing becomes more pertinent for the households.

2.3 Child Gender

To measure the effects on the sex-ratio of children, we create an indicator variable that is equal to one if a child born in a given district, during a given year is female, and zero otherwise. The data on the gender of children is from the Employment and Unemployment Rounds of the National Sample Surveys (NSS). NSS is a nationally-representative household labor survey. We use the NSS waves of 2004-05, 2007-08, and 2011-12 for this study as they interviewed a larger sample and are considered the "thick rounds"[11] of the NSS. The NSS survey records the age, sex, and current district of residence of every resident member of the interviewed households at the time of the survey. We use this to create a panel of the sex and year of birth of surviving children born between 2001 and 2011. We take the data on the children born between these waves from the immediately succeeding wave[12]. This process minimizes the missing children in the dataset as a result of migration[13]. Our final sample includes 89,264 infants (less than one at the time of the survey), 98,980 one-year-olds, and 107,764 two-year-olds. Due to the nature of the NSS data, we use a sample of surviving children in our analysis. This methodology is similar to previous studies that investigate female infant mortality in India (Rose, 1999; Kalsi, 2017), as it is advantageous for a couple of reasons. First, Rose (1999) describes that in survey data forgetting dead children is common while recalling birth histories. Such misreporting is plausibly systematically higher for female children than male children in regions where son-preference is strong. Therefore, using a sample of all children will likely underestimate the estimates on sex-selection. Second, the effects on sex-selection in

[11]During this period, the NSS collected similar but smaller labor data during 2005-06 and 2009-10. We use the "thick rounds" of the survey for a larger sample size.

[12]Therefore, we use the 2004-05 survey to construct the panel for all surviving children born between 2001 and 2004. We use the 2007-08 survey to construct the panel for all surviving children born between 2005 and 2007. Lastly, we use the 2011-12 survey to construct the panel for all surviving children born between 2008 and 2011.

[13]The maximum age of a child for whom we create the birth data retrospectively is three years. The 64^{th} Round of the National Sample Surveys (2007-08) records that only 0.14% of the children who are less than three years old migrated outside of their district. Understandably, over 90% of this migration is due to the migration of the parent or the earning member of the family. The low level of migration at such young ages ensures that any attenuation bias in our estimates due to treatment misassignment is minimal.

a sample of surviving children capture both prenatal and postnatal types of selection.

2.4 Long-run Child Health

To capture the effects on surviving children's long-term health, we construct gender-specific height-for-age z-scores using the Center for Disease Control's (CDC) growth charts as the reference[14]. Existing literature documents that children stunted during the first few years of their life can rarely make-up for it later (Martorell et al., 1994). Therefore, height-for-age is an ideal indicator to show variation in long-run heath caused by variation in early-life exposure to economic shocks (Maccini and Yang, 2009; Banerjee et al., 2010)[15]. For this purpose, we use the children's anthropometric data from the 2011-12 wave of the India Human Development Survey (IHDS II). The IHDS II collects height, age, and the district of current residence of the children who were present at the time their households' survey. Our final sample is 18,141 children between 0 and 11 years of age in 2011-2012.

3 Empirical Strategy

In this section, we describe the estimation strategy used in this study to test our two predictions derived in Appendix B. We first describe the identification strategies used on the outcome of child gender and then height-for-age.

3.1 Child Gender

Using the following specification, we first test Prediction 1 described in Appendix B. That is, we test whether agricultural productivity shocks during a year predict the gender of

[14]Specifically, we use the *zanthro* command in Stata, developed by Vidmar et al. (2004), for this purpose.

[15]The extensive literature showing a strong association between an individual's height and their long-run outcomes, such as cognitive abilities, education, and earnings (Lundborg et al., 2014; Vogl, 2014; Case and Paxson, 2008) emphasizes the importance of this measure.

an infant born during that year in India. In other words, we replicate the analysis in Rose (1999) using more recent data from the 21^{st} century.

$$Girl_{idt} = \alpha + \beta Rain_{dt} + \delta_d District'_d + \tau_t Year'_t + \epsilon_{idt} \tag{1}$$

where $Girl_{idt}$ is an indicator variable that is equal to one if a surviving child i born in district d and year t is a girl and is equal to zero otherwise. $Year_t$ is a vector of the year of birth dummies for the sample of infants[16] and captures year-specific shocks common to all districts. $District_d$ is a vector of the district of residence dummies and controls for time-invariant differences across districts. We estimate this model for the sample of infants (children less than one-year-old), the sample of one-year-olds, and two-year-olds to isolate when agricultural productivity shock's are most relevant for the sex-selection of children.

$Rain_{dt}$ is our proxy for agricultural productivity shock during the year t. As described in subsection 2.1, we primarily measure it as a z-score, the deviation of district d's rainfall in year t from its previous 10-year mean and scaled by its previous 10-year standard deviation. Therefore, to test Prediction 1, our primary coefficient of interest is β in Equation 1. Given the theoretical model presented in Appendix B and the findings in previous literature (Rose, 1999), we expect β to be positive, that is, the gender of a surviving child observed in our sample is more likely to be female if they were born in a relatively high-rainfall year.[17]

The identification of β relies on the assumption that, after controlling for district fixed effects, the deviation of a district's annual rainfall from its long-run mean is plausibly

[16] $Year_t$ is a vector of indicators for the year when the child is one-year-old for the sample of one-year-olds and indicators for the year when the child is two-year-old for the sample of two-year-olds.

[17] Apart from a linear rainfall z-score, we also estimate Equation 1 with a more flexible definition of rainfall shock to explore the possibility of a non-linear relationship between agricultural productivity shocks and infant sex-ratio. This specification includes six indicator variables that capture one standard deviation wide bins of the rainfall z-score measure defined above. We also use an ordinal definition of rainfall defined in Jayachandran (2006).

exogenous to the surviving child's gender. In stricter specifications, we add a vector of interactions of the district characteristics from the 2001 Census[18] with the year of birth dummies[19]. By including some measures of initial district conditions and allowing them to vary by the year of birth, we in-part capture the changing unobserved district conditions over time that may be correlated with changes in rainfall shock and child sex-ratio. In addition to the above covariates, for both the NSS and IHDS samples, we also add controls for household characteristics such as household size, an indicator for whether household head is male, household head's age, and household head's education to improve the precision of our estimates.

Next, we test Prediction 2, that is, the impact of NREGS on the relationship between agricultural productivity shocks and child gender. Specifically, we use the following specification to estimate NREGS's effect on the relationship between rainfall in a year and the gender of an infant born in that year:

$$
\begin{aligned}
Girl_{idt} =& \alpha + \mu_1 Rain_{dt} \times NREGS_{dt} + \mu_2 Rain_{dt} + \mu_3 NREGS_{dt} \\
& + \delta_d District'_d + \tau_t Year'_t + \epsilon_{idt},
\end{aligned} \tag{2}
$$

where $NREGS_{dt}$ is an indicator that is equal to one if district d received the guaranteed workfare program in year t[20]. All other variable definitions are consistent with Equation 1. Therefore, to test Prediction 2, μ_1 is our main coefficient of interest in Equation 2. μ_1 is the difference in the relationship between rainfall shocks and infant sex-ratio between the districts that receive NREGS and the districts that do not. In Appendix B, we hypothesize that NREGS helps households as an instrument of smoothing consumption and that in turn reduces the use of sex-selection as such an instrument. Therefore, we expect μ_1 to be

[18]These include the sex-ratio, the log of population, the percentage of rural population, the percentage of Scheduled Castes and Scheduled Tribes, the percentage of literates, and the labor force participation rate.

[19]For the sample of one-year-olds and two-year-olds, the trends use the year indicators for when a child is one-year-old or two-year-old, respectively.

[20]We create $NREGS_{dt}$ equal to one starting 2007 for Phase 1 districts, from 2008 for Phase 2 districts, and 2009 for Phase 3 districts as the implementation of each phase of NREGS missed the dry season at the start of that calendar year, which ends in March and .

negative; that is, NREGS attenuates the positive effect of rainfall on the probability that a surviving infant is a girl.

The phased implementation of NREGS enables us to exploit its temporal and spatial variation and identify its effects on the relationship between infant sex-ratio and agricultural productivity shocks using a framework similar to triple differences. We are comparing the infant sex-ratio before and after NREGS, between districts in different phases of NREGS implementation, and between high and low rainfall. The fundamental assumption in the identification of μ_1 is that the sensitivity of child sex-ratio to our measure of rainfall shock does not trend differently in districts with and without exposure to NREGS. However, the phase-in of the NREGS program was not random, and the assignment was according to a development index. Therefore, districts that receive the program during different phases display different characteristics (Imbert and Papp, 2015), and therefore, time trends may differ by district or phase. We describe this non-random assignment of the program and the corresponding robustness checks in subsection 3.3.

3.2 Long-run Child Health

Next, we will test Prediction 1 again by looking at the long-term consequences of agricultural productivity shock on the gender differences in the surviving children's health. Specifically, we use the following specification:

$$\begin{aligned} HeightForAge_{idt} =& \alpha + \lambda_1 Rain_{dt} \times Girl_{idt} + \lambda_2 Rain_{dt} + \lambda_3 Girl_{idt} \\ & + \delta_d District_d + \tau_t Year_t + \epsilon_{idt}, \end{aligned} \tag{3}$$

where $HeightForAge_{idt}$ is the height-for-age z-score of child i residing in district d and born in year t. $Girl_{idt}$ is an indicator variable equal to one if the surviving child is a girl and zero otherwise. All other variable definitions are consistent with Equation 1. λ_1 in Equation 3 is our primary coefficient of interest. In line with Prediction 1, we expect it to be positive if rainfall shock during the child's year of birth has long-lasting effects, and girls receive differentially more investment than boys in birth years with good rainfall.

Then, to explore Prediction 2, that is, how the guaranteed workfare program changed the relationship between early-life agricultural productivity shock and children's long-run health by gender, we use the following specification:

$$\begin{aligned} HeiightForAge_{idt} =& \alpha + \kappa_1 NREGS_{dt} \times Rain_{dt} \times Girl_{idt} + \kappa_2 Girl_{idt} \times Rain_{dt} \\ &+ \kappa_3 NREGS_{dt} \times Girl_{idt} + \kappa_4 Rain_{dt} \times NREGS_{dt} + \kappa_5 NREGS_{dt} \\ &+ \kappa_6 Girl_{idt} + \kappa_7 Rain_{dt} + \delta_d District_d + \tau_t Year_t + \epsilon_{idt}. \end{aligned} \tag{4}$$

All variable definitions in Equation 4 are consistent with Equation 1, Equation 2, and Equation 3.In Equation 4, κ_1 is our primary coefficient of interest. In line with Prediction 2, we hypothesize that it will be negative. That is, NREGS will attenuate the positive effect of good agricultural shocks during a birth year on the long-run health of girls relative to boys.

As the implementation of NREGS was at the district-level, and we also measure rainfall at the district-level, we cluster standard errors at the district-level in all specifications. This method allows for serial correlation in the error terms across the observations from the same district.[21]

3.3 Identification Assumptions and Robustness Checks

We mention in subsection 3.1 that our identification strategy is similar to a triple difference. Therefore, the identification of μ_1 in Equation 2 relies on the assumption of parallel trends. In other words, for our estimate of μ_1 in Equation 2 to be causal, it must be that the effect of rainfall shocks on infant sex-ratio in districts with and without NREGS would have trended

[21] However, both rainfall and NREGS's phased implementation may be spatially correlated. Therefore we show the robustness of our main results to two-way clustering at district and state-year levels in the Appendix.

identically in the absence of the program. This assumption of parallel trends is of particular concern in our context, as NREGS was implemented in the most backward districts first and in more developed districts later. The 2003 report of the Planning Commission of India bases this development ranking on each district's agricultural wages, agricultural productivity and the population of individuals of backward classes (Scheduled Castes and Scheduled Tribes).[22] To address this concern, we execute several tests.

First, we perform a placebo test, commonly used in differences-in-differences pre-trend tests. We assign placebo NREGS treatment to every district three years before its implementation and drop all observations after the NREGS treatment. So, the placebo test does not have any districts treated in reality[23]. By assigning NREGS to districts before the real implementation, we are implicitly testing whether districts were trending similarly to our main results in the years before implementation. Except for the placebo treatment assignment, the specification is the same as in Equation 2. Specifically, we estimate the following:

$$\begin{aligned} Girl_{idt} =& \alpha + \mu_1 Rain_{dt} \times proxyNREGS_{dt} + \mu_2 Rain_{dt} + \mu_3 proxyNREGS_{dt} \\ &+ \delta_d District'_d + \tau_t Year'_t + \epsilon_{idt} \text{ and,} \end{aligned} \tag{5}$$

where $proxyNREGS_{dt}$ is a dummy equal to one three years before actual NREGS implementation and all subsequent years. All other variables are the same as in Equation 2. Using the same specification as in the main model makes the coefficients directly comparable. This specification ensures that we use the same spatial variation as NREGS's implementation and similar temporal spacing. If pre-program trends are responsible for our results, we expect to see similar results in the placebo test: the estimate of μ_1 in the main model and the placebo model will be in the same direction. We perform a similar

[22]Zimmermann (2018) and Khanna and Zimmermann (2017) discuss the phased implementation of NREGS in detail.

[23]In other words; we code Phase 1 districts as if they received NREGS in 2004, Phase 2 districts in 2005, and Phase 3 districts in 2006.

placebo test for the height-for-age variable.

The placebo model above explicitly tests for the presence of different trends in how our outcome variable responds to rainfall by treatment status, which can confound our main estimates of interest. Next, we add stricter specifications of the main model that partially controls for such differential pre-treatment trends. Specifically, we present results from a more stringent specification of Equation 2, where we include phase-specific rainfall trends (or district-specific rainfall trends for an even stricter specification)[24]. These flexible specifications allow for the response of infant sex-ratio and height-for-age by gender to rainfall shocks to trend differently for each phase of NREGS implementation (or district). This method enables us to partly control for possible differential pre-treatment trends by treatment status.

Another identification concern for our empirical strategy is that other events may have occurred, or the government may have implemented other policies at the same time as NREGS. For example, using a rural bank expansion policy from 1970s, Rosenblum (2016) shows that greater access to bank branches affects the gender gap in child mortality outcomes. The Reserve Bank of India implemented another program to expand the growth of bank branches in financially backward districts, the Branch Authorization Policy in 2005 (Young, 2017). Therefore, this banking program could confound our results. Additionally, access to health centers can also impact the gender gap in early life investments (Chakravarty, 2010; Ravindran, 2018). Though India's most extensive child development program, Integrated Child Development Services (ICDS) program, has been expanding since 1975, the construction of Anganwadi centers[25] is a permissible activity under NREGS. Therefore, it may have accelerated with NREGS's implementation and can confound our estimates. To explicitly test whether these policies confound our results, we add the interaction of our measure of rainfall shock and an indicator for highly banked districts[26]

[24]We include phase and gender-specific trends or district and gender-specific trends for the robustness check of Equation 4.

[25]Courtyard shelters in rural villages that provide basic health facilities under ICDS.

[26]We define highly banked districts as those that have a population to bank ratio equal to or more than the

in Table A5 and Table A6. We also add the interaction of our measure of rainfall shock and a dummy for whether a village has a health center for the height-for-age outcome in Table A6 as this information is only available for the IHDS sample.

3.4 Summary Statistics

We present summary statistics for the main variables used in the analyses in Table 1. The top panel uses the NSS data. Across all three NREGS phases, girls make up less than half of newborns, one-year-olds, and two-year-olds. Consistent with the phased rollout of NREGS – in which the most underdeveloped districts were the first to receive the program – the household head is slightly younger and has less education, on average, in phase-one districts.

The second panel presents the anthropometric measures of height-for-age using the IHDS data. Phase three districts appear to have somewhat “healthier” children, with the highest height-for-age Z-scores for both girls and boys. However, it is essential to note that the Z-score is still well below the international standard (mean) (jay), even in phase-three districts. Interestingly, both boys and girls appear to be slightly shorter in phase-two districts than in phase-one districts, despite phase-one districts generally being poorer.

Finally, the third panel presents summary statistics at the district level, using the 2001 Census data. The data confirms what we see in the previous two panels and the stated variables used to determine the order of the program’s rollout. Phase-one districts have a higher SC/ST population, lower literacy, and are more rural than phase-two or phase-three districts. Interestingly, the sex-ratio is relatively more even in phase-one districts than in phase-two, and phase-three districts, which may suggest that parents in poorer districts are less likely to sex-select.

We begin with a simple graphical representation of our primary motivation in Figure 1.

national median during a year. We use bank branch data from the Reserve Bank of India’s Directory of Commercial Banks downloaded on June 26, 2016.

The figure shows kernel-weighted polynomial regressions of infant gender (left panel) and child's height-for-age by gender (middle and right panel) on our measure of rainfall shock during the year of birth. These figures strongly suggest that agricultural productivity shocks are positively related to girls' survival and health relative to boys. In other words, female mortality is higher, and female health is worse when the year of birth is a bad agricultural year. However, these are simply raw correlations and, as such, should be interpreted with caution.

4 Results

4.1 Main Results

4.1.1 Child Gender

We move to a more robust empirical examination of Prediction 1 in Table 2, that is, the effect of agricultural productivity shocks during a year on the gender of a randomly selected surviving child born that year (Rose, 1999). In the first five columns, we restrict estimation to children under the age of one, while columns 6 and 7 show the estimates on a sample of one-year-olds and two-year-olds, respectively.

In columns 1 through 3 in Table 2, we examine the relationship between rainfall and infant gender before the implementation of NREGS in each district. Column 1 presents the results from the simple specification in Equation 1. We find that a one-standard-deviation increase in rainfall (relative to the district's ten-year mean and standardized using the district's ten-year standard deviation) during a year increases the probability that a randomly chosen surviving infant born during that year is female by 1.4 percentage points. This result is consistent with Prediction 1 in our theoretical model. In other words, a positive agricultural shock increases investments in girls relative to boys. Adding more control variables for district characteristics and household characteristics in columns 2 and

3 increases the estimated effect size slightly; the coefficient in both columns now is 0.019. In all three cases, the coefficient is statistically significantly different from zero ($p < 0.01$).

To put these numbers in context, the interquartile range for the rainfall z-score is approximately 2.08 standard deviations. Our results indicate that a randomly-selected infant born during a year is a girl could increase by almost four percentage points if rainfall during that year increased from the 25th percentile of rainfall to the 75th percentile. This result shows that consistent with the findings in Rose (1999) during the 1970s, adverse rainfall shocks continue to exacerbate the "missing" women in India during the 21^{st} century.

Column 4 in Table 2 removes the sample restriction of pre-NREGS years and estimates the relationship in Equation 1 for the years 2001-2011. After removing the sample restriction, the coefficient on rainfall decreases by more than 40 percent, from 0.019 to 0.011. This result is suggestive evidence that NREGS may be an important channel for this attenuation.

In column 5 of Table 2, we add the previous year's rainfall z-score and the following year's rainfall z-score to the regression. Both coefficients are small and statistically insignificant. The evidence suggests that agricultural productivity shock right around birth or during pregnancy is the most critical determinant of female infant survival relative to males. We explore this possibility further in columns 6 and 7. Rose (1999) found that rainfall up to school going age had a significant impact on the survival of female children relative to male children. While the coefficient on the following year's rainfall z-score in column 5 suggests that this is unlikely to be the case in our sample, we now test this explicitly. Column 6 investigates the effect of rainfall during the year that a child is one year old on the probability that the child is female. Column 7 repeats this for the sample of two-year-olds. In neither column 6 nor 7 is the coefficient on rainfall z-score significant. The coefficient in column 6 is just 0.003, and the coefficient in column seven is negative but insignificant. These results again suggest that only rain right around birth is a significant predictor of child gender in modern India. These results also support the argument in Bharadwaj and Lakdawala (2013) that families are more likely to sex-select during pregnancy, relative to previous decades, and may partially explain why our results diverge from Rose (1999) on this point.

We next move to the analysis of the effects of NREGS on the relationship between agricultural productivity shocks and infant gender in Table 3. Consistent with the estimates in Table 2, the coefficient on "Year of birth rainfall (Z)" in Table 3 is always positive, suggesting that the effect of favorable rainfall on the probability of being female is positive before NREGS's implementation. Recall that Prediction 2 in the theoretical model posits that NREGS could attenuate this relationship between rainfall and child gender through risk mitigation. In other words, the model predicts that the elasticity of health investments with respect to productivity shocks is lower in girls than boys following the implementation of NREGS. Therefore, our coefficient of interest in Table 3 is the interaction term between rainfall and NREGS. This coefficient represents the change in the effect of rainfall on infant sex ratio following the implementation of the program[27]. This interaction term is negative, suggesting that the relationship between agricultural productivity shocks and infant sex-ratio decreases markedly following the program roll-out. In all four columns, the interaction term is more than 80 percent as large as the coefficient on rainfall, and the linear combination is never significant. This finding suggests that NREGS almost wholly reduced the relationship between rainfall and the sex ratio.

Additionally, the coefficients are notably stable across specifications. Column 2 adds year-of-birth fixed effects, column 3 adds household variables, and column 4 adds phase linear trends to insulate the estimates from possible differences in infant sex-ratio trends by phase before the program's implementation. Column 5 in Table 3 adds state-year fixed effects to the specification. Therefore, we estimate the interaction effect within state-year and across districts. Since now we are using much less variation in the rainfall z-score and NREGS roll-out, the coefficient on the interaction and also on rainfall is slightly smaller and less precisely estimated in this specification rendering them statistically insignificant.

[27] Alternatively, we might interpret this as changes in the effect of NREGS based on deviations in rainfall.

However, notably, consistent with all other specifications, even in this strict specification, the coefficient on the rainfall z-score is positive, the interaction is negative and attenuates more than 80 percent of the rainfall z-score coefficient.

Columns 1 through 5 utilize the entire panel we have constructed, from 2001-2011. While we include district and year-of-birth fixed effects, there may still be concerns that we are isolating variations in years far removed from the NREGS implementation. To test this possibility, in column 6, we restrict estimation only to the years 2005-2009, one year before the first phase of NREGS to one year following the final phase of NREGS. Though the results are slightly more imprecise, conclusions are unchanged. The interaction term is now slightly larger than the rainfall coefficient, though the linear combination is not significantly different from zero, similar to columns 1 through 5. If the change in the effect of rainfall is indeed due to the implementation of NREGS, we would expect to see these changes manifest themselves in the years of NREGS implementation. The result in column 6 is reassuring in that respect.

The coefficient on NREGS captures the program's effect on sex ratio during the years when rainfall is approximately equal to the district's ten-year mean. This coefficient is never close to statistically significant in any of the specifications in Table 3. The coefficient is negative in the first three columns. However, it becomes positive after controlling for phase-specific linear trends in columns 4 and 5. Moreover, the positive coefficient becomes much smaller once we restrict the sample to years closest to the program's implementation in column 6. Though very imprecisely estimated, these may indicate possible income effects or increased women's bargaining effect of NREGS that could positively impact girls' survival rates relative to boys during years with average rainfall. Alternatively, there may be an opposing level effect of NREGS on female infant survival during average rainfall years through an increase in overall infant mortality (Chari et al., 2019). An overall decline in children's survival might make it more pertinent to sex-select to increase the probability of having a son (Jayachandran, 2017).

4.1.2 Long-run Child Health

The previous results focused on the sex ratio of surviving children. It may be that the effect of rainfall at the time of birth also extends to long-run indicators of human capital investments for the surviving girls relative to boys, like anthropometrics. Figure 1 already presented suggestive evidence that this is indeed the case. Table 4 explores this possibility more formally. The dependent variable in all columns is height-for-age, defined using CDC growth charts. Column 1 estimates the effect of year-of-birth rainfall on height-for-age in 2011-12 (the year of the survey). The coefficient on rainfall is positive but small and insignificant. The female dummy coefficient is negative and significant, confirming the existence of differential investment in favor of boys relative to girls in India, independent of rainfall.

The female dummy coefficient, in conjunction with our previous results of the effect of rainfall on infant gender, raises the possibility that rainfall may differentially affect height-for-age for boys and girls. To explore this possibility, the specification in column 2 adds an interaction between female dummy and rainfall. The coefficient on rainfall – which now represents the effect of rainfall on boys' height-for-age – is minimal in magnitude. Moreover, the coefficient on the interaction term between rainfall and female is positive and statistically significant. The linear combination of this coefficient with the coefficient on rainfall is also significant (results not shown; p=0.049). Therefore, rainfall during the year of birth seemingly affects height-for-age for girls but not for boys. An important caveat is that this relationship is only identified by surviving children. It seems plausible that poorer households may be more affected by rainfall shocks, such that children who do not survive would come from the lower end of the height-for-age distribution (Barcellos et al., 2014). If so, then we can interpret our estimates in Table 4 as a lower bound of the actual effect.

Column 3 in Table 4 removes the pre-NREGS restriction and estimates the relationship over the years 1998-2012. Similar to Table 2, the favorable impact of rainfall on girls' height-for-age disappears and is no longer statistically significant. This result again supports the contention that something changed between 2006 and 2012. Column 4 explores the effects of NREGS on the relationship between rainfall and height-for-age, restricting the effect to be the same for both boys and girls. The effect of rainfall on height-for-age before NREGS implementation is positive and now marginally significant. However, NREGS does not statistically significantly attenuate this effect for all children on average.

Column 5 in Table 4 allows the effects of NREGS to vary by gender, a possibility that previous results on infant gender in this paper suggests. We find further evidence that NREGS impacts human capital investments differently for girls and boys. In particular, the triple interaction of *NREGS* $\times$ *Female* $\times$ *Rainfall* is negative and statistically significant, indicating NREGS attenuated the relationship between rainfall and height-for-age more for girls than for boys. This result is consistent with Prediction 2 of our theoretical model, which hypothesizes that the elasticity of health investments in girls relative to boys will decrease following the implementation of NREGS.[28]

The effect of NREGS on boys' height-for-age during average rainfall years (the coefficient on NREGS) is negative, small, and statistically insignificant. While the effect of NREGS on girls' height-for-age relative to boys' (the coefficient on Female times NREGS) is positive and statistically significant. Therefore, there is evidence that NREGS increased postnatal health investments in girls relative to boys even during average rainfall years. This result could be because NREGS has increased income effects beyond the income stabilization effect that differentially benefits girls. If the investment in girls were too small, to begin with, the increase in the household's utility would be more considerable from investing in them with the increase in income. This argument is precisely in line with our theoretical model.

[28]Since these results use the IHDS, we can control for village fixed effects, which we do in column five. Medical facility availability, such as antenatal clinics and Angadwadi centers at smaller geographical units, determines many health outcomes, so the inclusion of village fixed effects might be expected to alter the estimated impact of the program. Reassuringly, the inclusion of village fixed effects in Column 6 does not affect our substantive conclusions and increases precision.

4.1.3 Robustness Checks

One possible explanation for our findings is that the difference-in-differences assumption of parallel trends does not hold. We move to a standard test of pre-trends in Table 5. As described in subsection 3.3, we test the plausibility of our identification assumptions by constructing a "proxy" NREGS variable, with implementation moved up by three years and dropping districts when they receive treatment (to avoid any data from the years of implementation). If pre-trends are responsible for our results, we expect to see similar results in Table 5, as we did in the previous results. Columns 1 of Table 5 presents the corresponding results for the infant-gender sample. The first coefficient in the column is of interest: the coefficient is in the opposite sign of our results. It is positive and statistically significant. This finding violates the parallel trend assumption. However, it says that the trend was in the opposite direction of the result. In other words, the effect of rainfall on the sex ratio increased in NREGS treated districts relative to untreated districts in the three years just before NREGS implementation. This result is suggestive evidence that pre-trends are not responsible for our results. Instead, our estimates on infant-gender are an underestimate of the true effect. In column 2, we add district-specific rainfall trends added to partly control for the possible pre-trends. Expectedly, the coefficient on $ProxyNREGS \times Rainfall$ decreases and becomes statistically insignificant after this inclusion; however, it is still positive and substantial.

Columns 3 and 4 present the corresponding estimates for the IHDS sample and height-for-age. Column 4 includes district-gender-specific rainfall trends. The triple-interaction specification is again identical to that in Table 4, but with implementation moved up three years. In column 3, the coefficient is negative but much smaller in magnitude compared to the results in Table 4. In column 4, the coefficient is large and positive. This finding again suggests that our results may be underestimating the true effect of NREGS. However, the coefficient in column 4 is imprecisely estimated, making inference difficult.

Although the parallel trend is a disputable assumption in the triple-differences strategy, the pre-trends also fail to explain the results that we observe in Table 3 and Table 4. To reduce the possible downward bias in our estimates due to pre-trends, we include rainfall trends by treatment status in our main specifications. We do this in Table A5 and Table A6 of the appendix. This inclusion allows for the effect of rainfall on the dependent variables of interest to trend differently in each phase of NREGS (or district in the stricter specification)[29].

Table A5 presents results for the sex ratio. Column 1 includes phase-specific rainfall trends, while column two includes more flexible district-specific rainfall trends. The coefficients on the interaction term remain significant and are slightly larger than our previous results. This increase is consistent with the argument that we underestimate the true attenuation effect of NREGS due to opposing pre-trends. Table A6 presents the results for height-for-age. We modify the specifications here to allow the rainfall time trends to differ by gender. Column 1 again includes the phase-specific trends, while column two is the more flexible specification with district-specific trends. The coefficient of interest, the triple interaction between NREGS, rainfall, and female, remains negative in both columns but is attenuated. However, the standard errors are twice as large as standard errors in other columns, making inference difficult.

In conjunction with the results in Table A5 and Table A6, there appears to be little evidence that differential trends are responsible for our results. However, we interpret the height-for-age results with caution due to the slight attenuation observed in Table A6.

Another possible explanation of the results is the occurrence of events and policies along with NREGS. For example, access to banking and health centers expanded significantly around the dates of NREGS implementation. Besides, both access to banks and health centers during varying income shocks could plausibly affect our outcomes of interest. To explicitly test for this, we add controls interactions between a high banking district

[29]In all trend specifications, we include only the coefficient of interest due to the difficulty in interpreting other coefficients when we allow (rainfall) trends to be phase or district-specific.

indicator and the availability of a health center with the rainfall z-score in the last columns of Table A5 and Table A6. Our substantive conclusions are unchanged by these inclusions.

We perform some additional robustness checks with respect to the choice of specification and standard error clustering. These are described in Appendix C

4.2 Heterogeneity

4.2.1 Birth Order

The results above suggest that NREGS attenuates the relationship between rainfall and the gender of children born in India. In this section, we explore heterogeneity in this result by birth order.

The estimates above include all children, regardless of birth order. Previous results have found that sex selection increases in birth order, with relatively little sex selection of firstborn children (Bhalotra and Cochrane, 2010). If this is the case, we would expect to see smaller effects of rainfall on the gender of firstborn children and, correspondingly, not much of the attenuation effect of NREGS on this relationship. Table 6 tests these arguments and restricts our sample to the children of the household head so that we can identify the birth order of the children. First, column 1 explores the effects of rainfall on the gender of all household head's children. We do this to confirm that our results hold when examining the children of the head only. The magnitude of the rainfall coefficient is positive and strongly significant and of a similar magnitude to our rainfall coefficient in column 3 Table 2, suggesting that our results hold when focussing only on the children of the head.

We present the rainfall results using only the firstborn child of the household head in column 2 of Table 6. Consistent with previous literature, the coefficient is negative, small, and statistically insignificant. This result suggests that rainfall is uncorrelated with the gender of firstborn children in India. In column 3, we repeat the analysis using the sample of non-first-born children. The coefficient is now positive, large, and statistically significant. The magnitude is also quite a bit larger than our main results, consistent with the fact that the main effect of rainfall in the entire sample is an average over all the children. In other words, the magnitude must be more substantial for non-first-born children to compensate for the null effect among firstborn children.

Columns four through six in Table 6 examine whether the effects of NREGS are similar across these three subsamples of children. In all three columns, we lose substantial precision due to the interaction term and the smaller sub-samples. Nonetheless, the magnitude and direction of the coefficients can shed some light on the sex selection process, rainfall, and the effects of NREGS. Reassuringly, conclusions regarding the coefficients on rainfall shock and its interaction with NREGS in column four are similar to the main results reported in Table 3. Consistent with the null effect of rainfall shock on firstborn children in column 2, column 5 finds a very small coefficient estimate and statistically insignificant impact of NREGS on the relationship between rainfall shock and the gender of the firstborn child. Column 6 repeats the analysis using non-first-born children. Again consistent with previous results, we find larger (relative to all children of the head) effects of rainfall in non-NREGS districts and a larger effect of NREGS on this relationship. The results in columns 5 and 6 arguably strengthen the identification assumptions as well. The results suggest that any spurious trends that drive our results would have to be specific to non-first-born children only.

4.2.2 Prenatal vs. Postnatal

One key difference between our findings and the findings in Rose (1999) is that we do not find a relationship between rainfall and the sex ratio after the first year of life. This result suggests that there may be different mechanisms driving the results today than in the 1970s, the decade in which the data used in that paper were collected. One possibility for this deviation is sex-selective abortions. The use of ultrasound for sex determination and abortions for sex selection is illegal in India. However, law enforcement is weak, and the accessibility of these technologies has increased in recent decades (Kurup, 2011). To explore whether sex-selective abortions, rather than postnatal neglect between birth and one year of age, may drive the results that we observe, we use the data on state-level abortion rates in 2000 compiled by Robert Johnston. [30] [31] If sex-selective abortions explain much of the finding, we expect to see the effect of rainfall on infant sex ratio to be higher in states with existing higher abortion rates. We test this hypothesis in Table 7. The first column confirms this is indeed the case: The positive effect on rainfall on female infant survival (relative to male infants) is higher in states with previously high abortion rates.[32]

However, another possible driver of the relationship may be the correlation between abortion access and infant mortality. That is, areas with higher abortion access may also have higher overall infant mortality, which may lead to higher sex selection due to the chances of fewer surviving children. In column 2 of Table 7, we add an interaction between a state's infant mortality rate in 2000[33] and rainfall z-score. With this inclusion, the coefficient on the rainfall z-score and abortion rate interaction changes only slightly, and the interaction between rainfall z-score and the infant mortality rate, while imprecisely estimated, is quite small in magnitude. This result is consistent with sex-selective abortions being the main driving force behind the relationship, as well as with the recent literature that points to the role that ultrasound technology has played in sex selection (Bhalotra and Cochrane, 2010; Chen et al., 2013). The fact that infant mortality and the relationship are uncorrelated and that we explore gender-specific outcomes rather than aggregate infant outcomes makes our results complementary to Chari et al. (2019).

Columns 4 and 5 in Table 7 test whether NREGS had more substantial effects on the rainfall-sex ratio relationship when abortion rates are higher. Unfortunately, the results are

[30] johnstonsarchive.net/policy/abortion/india/ab-indias2.html

[31] We use data from 2000 to exploit pre-determined variation in access to and tendencies in sex-selection.

[32] Column 3 in Table 7 presents the results using a dummy variable for the abortion rate, which is equal to one if the state's abortion rate in 2000 was higher than the median. The results are consistent with column 1.

[33] http://niti.gov.in/content/infant-mortality-rate-imr-1000-live-births

very imprecisely estimated, but the signs and magnitudes of the coefficients are consistent with the general findings of this paper.

4.3 NREGS, Rainfall, and "Missing" Girls

All the results presented in this paper suggest that adverse rainfall shocks profoundly impacted the survival of girls in the early 2000s relative to boys. Consequently, it appears that adverse income shocks, as proxied by rainfall, may be responsible for a substantial number of "missing" women (Sen, 1990, 1992; Bongaarts and Guilmoto, 2015). Additionally, if this is the case, our findings suggest that NREGS may have "saved" a significant number of these girls. In this section, we estimate the number of missing girls caused by rainfall and that NREGS could have possibly saved since the program appears to have substantially attenuated the relationship between rainfall and the probability of a girl surviving relative to a boy.

We estimate these numbers by assuming that the birth of boys is unaffected by rainfall. We analyze the plausibility of this assumption in Table A7. We collapse the NSS data to the district-year level, summing the number of boys and the number of girls born in each cell (weighted by the NSS national sampling weights). We then regress the (log of) number of girls and boys in each district and year on district-level annual rainfall shock. We restrict estimation to pre-NREGS years (from 2001 to 2005) and include district fixed effects, year fixed effects, and the census variables interacting with year dummies.

The results in Table A7 show that an increase in rainfall increases the number of surviving girls born but does not statistically significantly affect the number of surviving boys born during a year. A one-standard-deviation increase in a year's rainfall is associated with a 4.5 percent increase in the number of girls born during that year. However, column 2 shows no significant correlation between rainfall and the number of boys born in a year: The coefficient in column two is only slightly more than 0.01, and the t-statistic is well less than one. This result suggests that the assumption that rainfall does not affect the number of newborn boys is plausible. This result is also noteworthy as it suggests that households are not decreasing concurrent fertility in response to rainfall shocks. If this were the case, we would expect to see a significantly positive coefficient on rainfall for the sample of boys as well. These results are consistent with the results presented in Table 2. It supports the hypothesis that decreasing early-life investment in female children is a primary risk mitigation strategy but not reducing early-life investment in male children.

Next, we assume that the number of boys and the number of girls in a district is equal when rainfall is approximately one standard deviation above its ten-year mean. We choose +1 standard deviations since, at rainfall between 0 and 1 standard deviation, the relationship between rainfall and sex-ratio becomes statistically insignificant (Table A2)[34]. Next, we take the effect of rainfall on the probability of being a girl from the coefficient in column 3 of Table 2: 0.019. Using this, we predict the number of girls born for the years 2001 to 2005 in each district, relying on the assumption that the number of boys and girls are equal at $Z = 1$. Finally, we estimate the effect of rainfall on the probability of being a girl after the implementation of NREGS. We calculate this from column 3 in Table 3: $(0.053 - 0.045) = 0.008$. Taking this point estimate, we predict the number of girls for the years 2001 to 2005 if NREGS was hypothetically present.

Over the years 2001 to 2005, the difference in the two predicted values is approximately 0.7 million girls. This number translates to around 14.7 thousand girls per year in India or around 283 girls per district per year. In other words, if NREGS had been available in the years 2001 to 2005, we estimate that at least 0.7 million more girls relative to boys would have been alive in 2006. Though large, this calculation does not take into account any improvement in the lives of the surviving girls who receive more investments around birth and are thus healthier. However, this calculation also does not consider the possibility that NREGS has an overall detrimental impact on the infant mortality rate for

[34]Though the coefficient estimate for this range is still more than a quarter of the coefficient estimate on the most extreme case of low rainfall bin. So, we may be undervaluing the number of girls that may have survived to some extent.

mothers who are eligible for NREGS (Chari et al., 2019). As such, we derive a relative estimate (of girls relative to boys), not an absolute estimate.

5 Discussion

In this section, we discuss some possible explanations for our findings: risk mitigation and an increase in women's value and bargaining power in the household.

One possible channel through which NREGS affects these gender-differentiated health outcomes is its ability to help households mitigate income risks. We emphasized this channel in our theoretical model, and we draw from the empirical contribution of Santangelo (2019). She shows that NREGS decouples the comovement of wages and consumption with rainfall using the NSS data: A 1% increase in monsoon rainfall increases wages by 6.2% and consumption by 5.6% in the absence of NREGS. The implementation of NREGS decreases these coefficients by 5.3% and 5.2%, respectively. These results suggest that an increased ability to alleviate income risks during bad rainfall years following the implementation of NREGS is a possible mechanism that explains our results.

The original act intended women to make up a large proportion of NREGS beneficiaries. The original legislation mandated that: 1) women make up at least one-third of beneficiaries; 2) worksites provide a crèche for the care of children; and 3) men and women are paid equal wages (Ministry of Law and Justice, 2005). According to one government advisor, the mandated minimum program wage often double the prevailing wage rate for women at the time.[35] While enforcement of these requirements varied by state – and even district – this suggests another mechanism through which NREGS could affect the outcomes studied in this paper: women's bargaining power and the resulting increase in human capital investments in their female children (Qian, 2008; Duflo, 2003). Besides, if the future (expected) earnings of girls increase, parents may respond by increasing investments in

[35]http://www.levyinstitute.org/pubs/EFFE/Mehrotra_Rio_May9_08.pdf

girls' human capital (Balakrishnan, 2017; Heath and Mobarak, 2015; Jensen, 2012) – and may even affect the sex ratio (Balakrishnan, 2017; Qian, 2008). If rural households believe the program will persist into the future, then they may adjust their investments in girls.

Azam (2011) finds that NREGS increases women's participation in public works. However, there is mixed evidence on the effect of NREGS on women's casual wages and labor force participation (Imbert and Papp, 2015; Azam, 2011; Zimmermann, 2018). On children's human capital investment, there some evidence of increases in female children's educational outcomes relative to male children (Mani et al., 2014). Afridi et al. (2016) suggests that a part of this can be the increase in the bargaining power of mothers who worked for the NREGS program. By exploring increases in different types of consumption expenses in areas where NREGS hired more women, Ojha (2018) finds suggestive evidence that NREGS increases female bargaining power. Overall there is some evidence of increase in female autonomy as a result of NREGS.

We look at the main results in light of the indication that female autonomy increases as a result of NREGS. The level effect of NREGS in previous tables can be interpreted as the effect of NREGS on the outcome of interest when the rainfall deviation is zero (that is, when rainfall is equal to the district's ten-year mean). This coefficient is statistically insignificant in all of the child sex-ratio specifications. These estimates are prima facie evidence that an increase in women's bargaining power due to NREGS does not statistically significantly affect child sex-ratio, at least in average rainfall years. However, we find a positive effect of NREGS on female children's height for age relative to male children. This result is consistent with the education results in (Mani et al., 2014; Afridi et al., 2016). That is, for surviving children, an increase in mother's autonomy may be a possible channel for increased human capital investment in female children compared to male children. This result also opens up the possibility that mothers exercise this increased female autonomy more during bad rainfall years – when their daughters are more vulnerable.

For children born during average rainfall years, one reason that could explain why we observe some evidence of increased bargaining power as a channel for increased health investment in female children but not for improving the sex-ratio of infants is a model of limited commitment in household bargaining. In an extreme case, such an intra-household arrangement implies that women have a higher bargaining power only when they work. In NREGS's case of unskilled manual labor, this is likely improbable when women are pregnant. However, it is still likely for mothers with surviving female children to increase postnatal early-life investments in their daughters, leading to their improved height-for-age.

Overall, we cannot rule out that an increase in women's bargaining power is a possible channel through which our results manifest. We find suggestive evidence that this is a more likely channel for postnatal investments than prenatal investments.

6 Conclusion

In this paper, we explore the effects of risk-mitigation through workfare programs in rural India on the relationship between agricultural productivity shocks and sex selection of infants. First, using more recent data, we re-establish that a positive agricultural shock reduces female child mortality. Second, we show that the introduction of Mahatma Gandhi National Rural Employment Guarantee Scheme (NREGS) mitigates the effect of income shocks on the sex selection of infants. Third, we find that before the advent of NREGS, a positive agricultural shock is also more positively related to the health of surviving female children than male children. Lastly, NREGS mitigates this relationship between income shocks and the health of surviving girls.

This paper establishes that policies that successfully provide tools for alleviating income risks to rural households in India can also successfully reduce sex selection of infants and decrease differential child health investments by gender. Though the paper uses one such policy, a rural workfare program, to show that a program which provides households with a social protection during lean agricultural years reduces sex selection among children, the channels explored in this paper more broadly establish that policies that help risk-mitigation can decrease sex selection when son-preference prevails. This result is especially important since the most common policy directed at reducing female child mortality is providing households with financial incentives for having daughters. However, recent literature shows that such policies' success is very sensitive to the design of these policies (Anukriti, 2018; Balakrishnan, 2017). Therefore, risk-mitigation and similar policies that help households smooth consumption over time may be an attractive development intervention, with favorable consequences for the sex ratio and female health investments.

Several questions remain from our analyses. First, our results suggest that NREGS would have saved approximately 283 girls per district per year from 2001 to 2005. Mechanically, this indicates that NREGS would have increased the total number of children in those years since rainfall appears to be uncorrelated with the number of boys born each year. However, a lingering question from this analysis is whether lifetime fertility increases. In other words, would surviving girls take the place of an additional child, or would women have the same number of future children as they would have before the implementation of the program? Given that the roll-out of NREGS happened over just three years, we are unable to explore long-term fertility changes using our data and empirical methodology (Sukhtankar, 2016). Second, this paper explores sex selection in response to an income shock. We show that NREGS decreases sex selection due to fluctuations in income. However, the results do not suggest that son preference diminishes following NREGS, or that sex selection does not occur through other channels. We find that one form of sex selection decreases following the implementation of the program. However, there are many other mechanisms apart from income shocks that may lead to sex selection. Our paper does not address these different possibilities and how NREGS interacts with them.

References

Afridi, F., Mukhopadhyay, A., and Sahoo, S. (2016). Female labor force participation and child education in india: evidence from the national rural employment guarantee scheme. *IZA Journal of Labor & Development*, 5(1):7.

Almond, D., Li, H., and Zhang, S. (2019). Land reform and sex selection in china. *Journal of Political Economy*, 127(2):560–585.

Anukriti, S. (2018). Financial Incentives and the Fertility-Sex Ratio Trade-off. *American Economic Journal: Applied Economics*.

Anukriti, S., Kwon, S., and Prakash, N. (2017). Dowry: Household responses to expected marriage payments. *Working Paper*.

Azam, M. (2011). The impact of indian job guarantee scheme on labor market outcomes: Evidence from a natural experiment. *Available at SSRN 1941959*.

Balakrishnan, U. (2017). Incentives for Girls and Gender Bias in India. *Working Paper*.

Banerjee, A., Duflo, E., Postel-Vinay, G., and Watts, T. (2010). Long-run health impacts of income shocks: Wine and phylloxera in nineteenth-century France. *The Review of Economics and Statistics*, 92:714–728.

Banerjee, R. and Maharaj, R. (2019). Heat, infant mortality and adaptation: Evidence from India. *Journal of Development Economics*, page 102378.

Barcellos, S., Carvalho, L., and Lleras-Muney, A. (2014). Child Gender and Parental Investments In India: Are Boys and Girls Treated Differently? *American Economic Journal: Applied Economics*, 6(1):157–189.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science*, 335(6068):582–586.

Bhalotra, S., Chakravarty, A., and Gulesci, A. (2016). The Price of Gold: Dowry and Death in India. *IZA Discussion Paper 9679*.

Bhalotra, S. and Clots-Figueras, I. (2014). Health and the political agency of women. *American Economic Journal: Economic Policy*, 6(2):164–97.

Bhalotra, S. and Cochrane, T. (2010). Where Have All the Young Girls Gone? Identifying Sex-Selective Abortion in India. *IZA Discussion Paper*, 5381.

Bharadwaj, P. and Lakdawala, L. (2013). Discrimination Begins in the Womb: Evidence of Sex-Selective Prenatal Investments. *Journal of Human Resources*, 48(1):71–113.

Bongaarts, J. and Guilmoto, C. Z. (2015). How many more missing women? Excess female mortality and prenatal sex selection 1970âĂŞ2050. *Ithaca: Cornell University Press*, 42(2):241–269.

Borooah, V. (2004). Gender Bias Among Children in India in their Diet and Immunisation Against Disease. *MPRA Paper from University Library of Munich, Germany*.

Case, A. and Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, 116:499–532.

Chakravarty, A. (2010). Supply Shocks and Gender Bias in Child Health Investments: Evidence from the ICDS Programme in India. *The B.E. Journal of Economic Analysis Policy*, 10(1):1–26.

Chari, A. V., Glick, P., Okeke, E., and Srinivasan, S. V. (2019). Workfare and infant health: Evidence from India's public works program. *Journal of Development Economics*, 138:116–134.

Chen, L., 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.

Chen, Y., Li, H., and Meng, L. (2013). Prenatal Sex Selection and Missing Girls in China: Evidence from the Diffusion of Diagnostic Ultrasound. *Journal of Human Resources*, 48(1):36–70.

Das Gupta, M. (1987). Sex Bias in the Family Allocation of Food and Health Care in Rural Bangladesh. *Population and Development Review*, 13(1):77–100.

Dasgupta, A. (2013). *Can the major public works policy buffer negative shocks in early childhood?: evidence from Andhra Pradesh, India*. Young Lives.

de Chaisemartin, C. and d'Haultfoeuille, X. (2019). Two-way fixed effects estimators with heterogeneous treatment effects. *NBER Working Paper No. 25904*.

Duflo, E. (2003). Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in south africa. *The World Bank Economic Review*, 17(1):1–25.

Eswaran, M. (2002). The empowerment of women, fertility, and child mortality: Towards a theoretical analysis. *Journal of Population Economics*, 15(3):433–454.

Fetzer, T. (2014). Social Insurance and Conflict: Evidence from India. *Working paper*.

Foster, A. D. and Gehrke, E. (2017). Consumption Risk and Human Capital Accumulation in India. *NBER Working Paper No. 24041*.

Ganatra, B. and Hirve, S. (1994). Male bias in health care utilization for under-fives in a rural community in Western India. *Bulletin of the World Health Organization*, 72(1):101–104.

Ganatra, B. and Hirve, S. (2001). Does increased access increase equality? Gender and child health investments in India. *Journal of Development Economics*, 89(1):62–76.

Garg, T., Jagnani, M., and Taraz, V. (2018). Temperature and human capital in india.

Gehrke, E. (2017). An employment guarantee as risk insurance? assessing the effects of the nregs on agricultural production decisions. *The World Bank Economic Review*, 1:23.

Heath, R. and Mobarak, A. M. (2015). Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics*, 115:1–15.

Imbert, C. and Papp, J. (2015). Labor Market Effects of Social Programs: Evidence from India's Employment Guarantee. *American Economic Journal: Applied Economics*, 7(2):233–263.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Jayachandran, S. (2017). Fertility Decline and Missing Women. *American Economic Journal: Applied Economics*, 9(1):118–1392.

Jayachandran, S. and Kuziemko, I. (2011). Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India. *The Quarterly Journal of Economics*, 126(3):1485–1538.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? experimental evidence from india. *The Quarterly Journal of Economics*, 127(2):753–792.

Jha, R., Bhattacharyya, S., and Gaiha, R. (2011). Social safety nets and nutrient deprivation: An analysis of the National Rural Employment Guarantee Program and the Public Distribution System in India. *Journal of Asian Economics*, 22(2):189–201.

Kalsi, P. (2017). Seeing is Believing - Can Increasing the Number of Female Leaders Reduce Sex Selection in Rural India? *Journal of Development Economics*, 126(1).

Kaur, S. (2019). Nominal Wage Rigidity in Village Labor Markets. *American Economic Review*, pages 3585–3616.

Khanna, G. and Zimmermann, L. (2017). Guns and butter? Fighting violence with the promise of development. *Journal of Development Economics*, 124:120–141.

Kochar, A., Singh, K., and Singh, S. (2009). Targeting public goods to the poor in a segregated economy: An empirical analysis of central mandates in rural India. *Journal of Public Economics*, 93(7):917 – 930.

Kurup, S. (2011). The return of ultrasound.

Lundborg, P., Nystedt, P., and Rooth, D.-O. (2014). Height and earnings: The role of cognitive and noncognitive skills. *Journal of Human Resources*, 49:141–166.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–26.

Mani, S., Behrman, J. S., Galab, S., and Reddy, P. P. (2014). *Impact of the NREGS on schooling and intellectual human capital*. Young Lives.

Martorell, R., Khan, L. K., and Schroeder, D. G. (1994). Reversibility of stunting: epidemiological findings in children from developing countries. *European Journal of Clinical Nutrition*, 48:S45–57.

Merfeld, J. D. (2019). Spatially Heterogeneous Effects of a Public Works Program. *Journal of Development Economics*, 136:151–167.

Merfeld, J. D. (2020). Moving up or Just Surviving? Non-Farm Self-Employment in India. *American Journal of Agricultural Economics*, 102(1).

Ministry of Law and Justice (2005). The National Rural Employment Guarantee Act. shttp://nrega.nic.in/rajaswa.pdf.

Niehaus, P. and Sukhtankar, S. (2013). Corruption dynamics: The golden goose effect. *American Economic Journal: Economic Policy*, 5(4):230–69.

Ojha, M. (2018). Essays in development economics.

Qian, N. (2008). Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *The Quarterly Journal of Economics*, 123(3):1251–1285.

Ravi, S. and Engler, M. (2015). Workfare as an effective way to fight poverty: The case of India's NREGS. *World Development*, 67:57–71.

Ravindran, S. (2018). Parental Investments and Early Childhood Development: Short and Long Run Evidence from India. *Working Paper*.

Rose, E. (1999). Consumption Smoothing and Excess Female Mortality in Rural India. *Review of Economics and Statistics*, 81(1):41–49.

Rosenblum, D. (2013). The effect of fertility decisions on excess female mortality in India. *Journal of Population Economics*, 26(1):147–180.

Rosenblum, D. (2016). Are Banks Bad for Boys? Estimating the Effect of Banks on Child Mortality, Education, and fertility in Rural India. *Canadian Center for Healthj Economics Working Paper Series*, Working Paper No: 160003.

Santangelo, G. (2019). Firms and Farms: The Impact of Agricultural Productivity on the Local Indian Economy. *Working paper*.

Sarma, N. (2019). Domestic violence and workfare: An evaluation of india's mgnrega. *Available at SSRN 3341589*.

Sen, A. (1990). More Than 100 Million Women Are Missing. *The New York Review of Books*, 20:61–66.

Sen, A. (1992). How many more missing women? Excess female mortality and prenatal sex selection 1970-2050. *BMJ: British Medical Journal*, 304(6827):587.

Shah, M. and Steinberg, B. M. (2015). Workfare and human capital investment: Evidence from India. *National Bureau of Economic Research Working Paper*.

Song, L., Appleton, S., and Knight, J. (2006). Does increased access increase equality? Gender and child health investments in India. *World Development*, 34(9):1639–1653.

Sukhtankar, S. (2016). India's National Rural Employment Guarantee Scheme: What Do We Really Know about the World's Largest Workfare Program? In *India Policy Forum*, volume 13, pages 2009–10.

Vidmar, S., Carlin, J., Hesketh, K., Cole, T., et al. (2004). Standardizing anthropometric measures in children and adolescents with new functions for egen. *Stata Journal*, 4(1):50–55.

Vogl, T. S. (2014). Height, skills, and labor market outcomes in Mexico. *Journal of Development Economics*, 107:84–96.

Young, N. (2017). Banking and Growth: Evidence From a Regression Discontinuity Analysis. *EBRD Working Paper No. 207*.

Zimmermann, L. (2018). Why Guarantee Employment? Evidence from a Large Indian Public-Works Program. *Working Paper*.

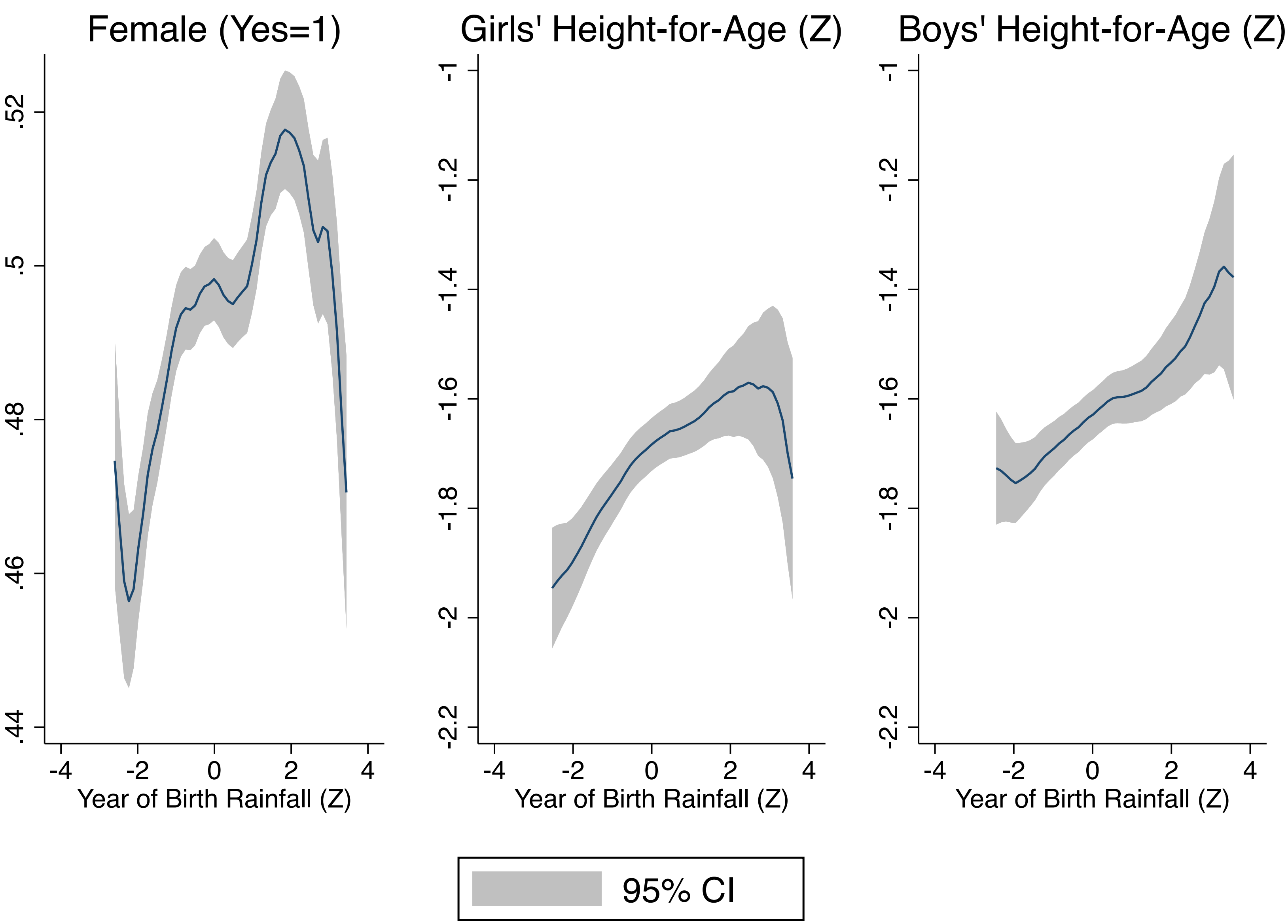
Figures & Tables

Table 1: Summary Statistics

	Phase 1	Phase 2	Phase 3
Panel A: NSS Children			
Girl (if < 1 year old)	0.49	0.48	0.48
	(0.50)	(0.50)	(0.50)
Girl (if one year old)	0.49	0.48	0.49
	(0.50)	(0.50)	(0.50)
Girl (if two years old)	0.48	0.47	0.48
	(0.50)	(0.50)	(0.50)
Household size	6.55	6.51	6.59
	(2.92)	(2.93)	(2.95)
Head is male	0.93	0.93	0.93
	(0.25)	(0.26)	(0.26)
Head age	41.62	42.00	42.63
	(13.64)	(13.80)	(14.37)
Head education	1.90	2.00	2.26
	(1.36)	(1.43)	(1.51)
Observations	51,493	38,055	73,759
Panel B: IHDS Children			
Girls' height for age (Z)	−1.65	−1.73	−1.57
	(1.56)	(1.52)	(1.53)
Boys' height for age (Z)	−1.53	−1.80	−1.33
	(1.45)	(1.45)	(1.46)
Observations	2,385	1,793	4,260
Panel C: Census Districts (NSS Sample)			
Percent SC/ST	0.38	0.31	0.27
	(0.20)	(0.21)	(0.22)
Percent literate	0.47	0.53	0.58
	(0.11)	(0.13)	(0.10)
Labor force participation	0.42	0.40	0.40
	(0.07)	(0.07)	(0.07)
Population (log)	14.06	14.11	13.98
	(0.87)	(0.89)	(1.09)
Percent rural	0.86	0.82	0.72
	(0.09)	(0.13)	(0.17)
Sex ratio	945.83	940.27	926.76
	(45.99)	(46.97)	(64.60)
Observations	171	112	236

Statistics are means. All individual statistics are nationally representative and are estimated using survey weights. The individual statistics for the NSS are for children less than two years old, for the years 2001-2005. The NSS Districts data are from the 2000 census. The IHDS anthropometrics are constructed using CDC charts and the *zanthro* command in Stata (Vidmar et al., 2004).

Figure 1: Rainfall in Year of Birth and Child Outcomes



Graphs are kernel-weighted local polynomial regressions. All observations are before the implementation of NREGS in a district. The top and bottom one percent of rainfall values are trimmed for ease of presentation.

Table 2: Rainfall and Child Gender

	Newborns					One-Year Olds	Two-Year Olds
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Current rainfall (Z)	0.014***	0.019***	0.019***	0.011*	0.019***	0.003	−0.006
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Previous rainfall (Z)					0.001		
					(0.006)		
Next rainfall (Z)					0.006		
					(0.006)		
Pre NREGS	Yes	Yes	Yes	No	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Vars	No	Yes	Yes	Yes	Yes	Yes	Yes
Household Vars	No	No	Yes	Yes	Yes	Yes	Yes
Observations	66,312	65,810	65,791	88,547	65,791	72,315	78,593

Standard errors are in parentheses and are clustered at the district level. Columns one through three and five through seven use the years 2001-2007; column four uses the years 2001-2011. All data are from NSS waves 61, 64, and 68. Newborns are defined as chilren less than one year of age. Current rainfall is standardized using the mean and standard deviation of the previous 10 years. * p<0.1 ** p<0.05 *** p<0.01

Table 3: NREGS, Rainfall, and Child Gender

	Years 2001-2011					Years 2005-2009
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Rainfall (z) times	−0.038**	−0.044**	−0.045**	−0.046**	−0.035	−0.054**
NREGS	(0.018)	(0.020)	(0.020)	(0.020)	(0.025)	(0.025)
Year of birth rainfall (Z)	0.047***	0.053***	0.053***	0.054***	0.041	0.052*
	(0.018)	(0.020)	(0.020)	(0.020)	(0.025)	(0.027)
NREGS	−0.038	−0.025	−0.024	0.017	0.026	0.006
	(0.027)	(0.037)	(0.037)	(0.042)	(0.043)	(0.043)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	No	Yes
Year by State FE	No	No	No	No	Yes	No
District Vars	No	Yes	Yes	Yes	Yes	Yes
Household Vars	No	No	Yes	Yes	Yes	Yes
Phase Linear Trend	No	No	No	Yes	Yes	No
Observations	89264	88570	88547	88547	88547	38451

Standard errors are in parentheses and are clustered at the district level. The dependent variable in all columns is whether a newborn (defined as less than one year of age) is a girl. Columns one through four use the years 2001-2011, while column five restricts estimation to just one year prior to NREGS to one year following implementation of the final phase. Current rainfall is standardized using the mean and standard deviation of the previous 10 years. * p<0.1 ** p<0.05 *** p<0.01

Table 4: NREGS, Rainfall, and Child Height-for-Age

	District FE					Village FE
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year of birth rainfall (Z)	0.026	0.002	0.025	0.035*	0.014	0.057***
	(0.019)	(0.023)	(0.018)	(0.018)	(0.022)	(0.021)
Female	−0.087**	−0.091**	−0.047	−0.047	−0.084**	−0.080*
	(0.036)	(0.036)	(0.029)	(0.029)	(0.036)	(0.042)
Female times Rainfall		0.051*	−0.003		0.044	0.023
		(0.030)	(0.024)		(0.030)	(0.019)
NREGS				0.003	−0.048	−0.037
				(0.088)	(0.090)	(0.080)
Rainfall (z) times NREGS				−0.028	0.028	0.017
				(0.031)	(0.037)	(0.028)
Female times NREGS					0.116*	0.094**
					(0.062)	(0.036)
NREGS times Female times Rainfall					−0.118*	−0.098**
					(0.060)	(0.040)
Years	Pre NREGS	Pre NREGS	1998-2012	1998-2012	1998-2012	1998-2012
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,309	13,309	19,373	19,373	19,373	19,373

Standard errors are in parentheses and are clustered at the district level (columns one through five) or village level (column six). Columns one and two include children born between the years 1998 and 2005, though only 3.38 percent of observations come from prior to 2002. Columns three through six includes children born during the years 1998-2011 (only 1.20 percent of observations are from prior to 2002). The dependent variable in all columns is height-for-age, standardized using the CDC charts and the *zanthro* command in Stata (Vidmar et al., 2004). Rainfall is always defined as rainfall during year of birth. * p<0.1 ** p<0.05 *** p<0.01

Table 5: Testing the Parallel Trends Assumption

	Female		HAZ	
	(1)	(2)	(3)	(4)
Proxy NREGS times Female times Rainfall			−0.009	0.175
			(0.059)	(0.138)
Proxy NREGS times Rainfall	0.023**	0.020	−0.049	
	(0.010)	(0.017)	(0.045)	
Female times Rainfall			0.053	
			(0.042)	
Proxy NREGS times Female			0.054	
			(0.074)	
Proxy NREGS	0.004		0.049	
	(0.016)		(0.079)	
Female			−0.118**	
			(0.048)	
Rainfall (Z)	0.009*		0.027	
	(0.005)		(0.031)	
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes
Rainfall trends	No	Yes	No	Yes
Observations	57529	57529	28216	28216

Standard errors are in parentheses and are clustered at the district level. Columns one and two use the National Sample Survey and the dependent variable is whether a newborn is female. Columns three and four use the IHDS and the dependent variable is height-for-age Z score. The placebo NREGS variable is defined similarly to the NREGS variables in the prior tables, but with implementation date moved up one year (e.g. assuming phase one districts received the program in 2005 instead of 2006).
* p<0.1 ** p<0.05 *** p<0.01

Table 6: First Born and Sex Selection

	Pre-NREGS			Effects of NREGS		
DV in all columns: Female	Child of Head	First Child	Not First Child	Child of Head	First Child	Not First Child
Rainfall (z) times NREGS				−0.034	0.000	−0.048
				(0.027)	(0.052)	(0.033)
Year of birth rainfall (Z)	0.020***	−0.004	0.032***	0.037	−0.006	0.051
	(0.006)	(0.013)	(0.008)	(0.026)	(0.051)	(0.033)
NREGS				−0.002	0.006	0.000
				(0.048)	(0.071)	(0.052)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes	Yes
Observations	36,130	8,732	27,398	48,341	12,139	36,202

Standard errors are in parentheses and are clustered at the district level. The first three columns use observations prior to NREGS. The last three columns use the years 2001-2011. First born is defined as the oldest child currently in the household.
* p<0.1 ** p<0.05 *** p<0.01

Table 7: Abortion and Infant Mortality as Drivers

	Pre-NREGS	Pre-NREGS	Pre-NREGS	All	2005-2009
Year of birth rainfall (Z)	0.008	0.007	0.004	0.043	−0.019
	(0.008)	(0.025)	(0.008)	(0.039)	(0.051)
Rainfall times Abortion rate	0.280**	0.284*			
(per pregnancy)	(0.136)	(0.162)			
Rainfall times Infant mortality		0.011			
rate (per birth)		(0.351)			
Rainfall times Abortion rate			0.019**	0.024	0.082
(above median)			(0.009)	(0.044)	(0.061)
NREGS				0.014	0.022
				(0.049)	(0.059)
Rainfall (z) times NREGS				−0.025	0.037
				(0.042)	(0.058)
NREGS times Abortion (median)				−0.043	−0.020
				(0.042)	(0.058)
NREGS times Rain time Abortion (median)				−0.033	−0.105
				(0.047)	(0.064)
District FE	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes
Observations	63170	63170	63170	85217	37066

Standard errors are in parentheses and are clustered at the district level. The first three columns use observations prior to NREGS. The last three columns use the years 2001-2011. The abortion rate is defined as per 1,000 pregnancies but is multiplied by 1,000, so its interpretation is "per pregnancy." Infant mortality is similarly scaled to be interpreted as "per live birth." The abortion rate is culled from multiple sources and is defined at the state level in 2000.
* $p<0.1$ ** $p<0.05$ *** $p<0.01$

A Results

Figure A1: Sex Ratio in India

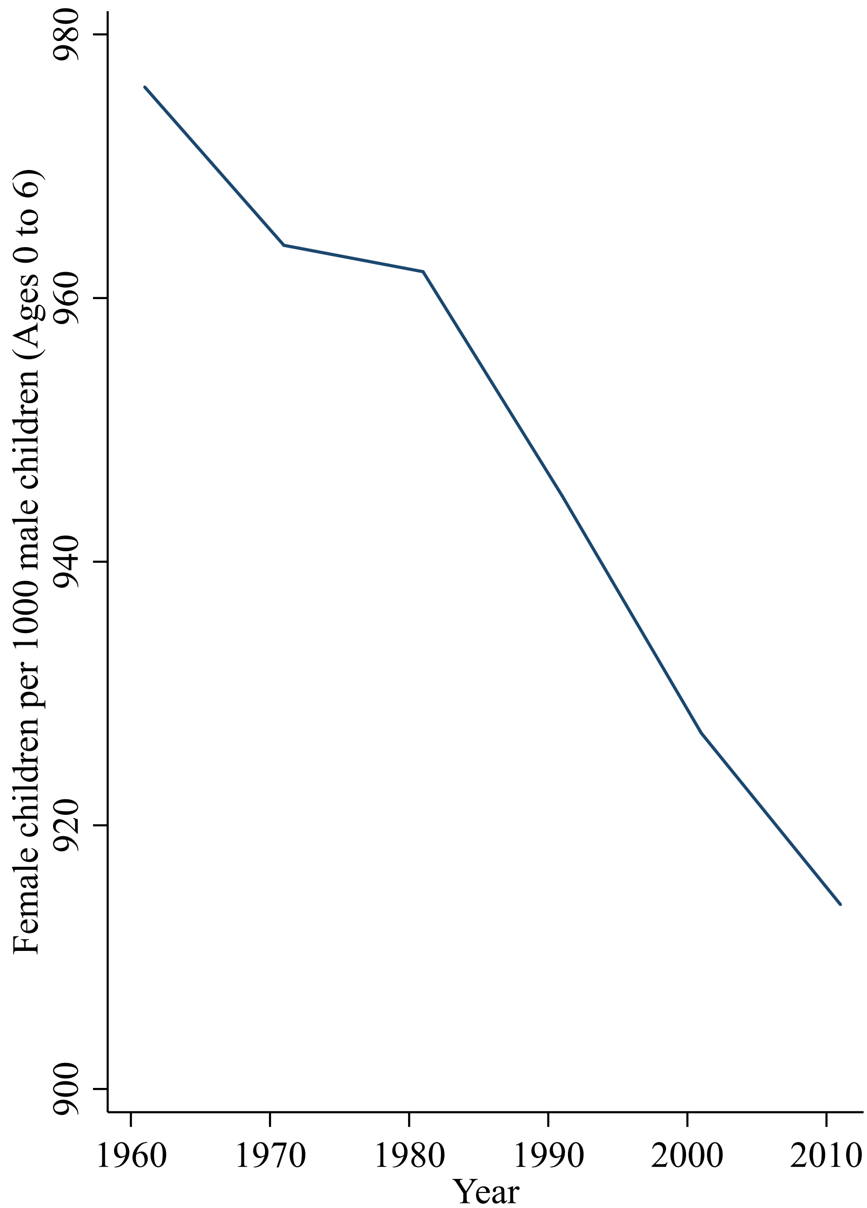


Table A1: Serial Correlation in Rainfall

	(1)	(2)	(3)
Previous year's rainfall	0.015	−0.014	−0.002
	(0.015)	(0.016)	(0.016)
State FE	Yes	No	No
District FE	No	Yes	Yes
Year FE	No	No	Yes
Observations	77227	77227	77227

Standard errors are in parentheses and are clustered at the district level. The dependent variable is the current year's rainfall.
* $p<0.1$ ** $p<0.05$ *** $p<0.01$

Table A2: Rainfall Robustness

	Rainfall Robustness			NREGS	
	Model 1	Model 2	Model 3	All Years	All Years
Year of birth rainfall (Z)	0.026***				
	(0.007)				
Rain <-2		−0.085***		-0.031	
		(0.029)		(0.044)	
Rain between -1 and -2		−0.050**		-0.115**	
		(0.024)		(0.051)	
Rain between 0 and -1		−0.045**		-0.052	
		(0.019)		(0.041)	
Rain between 0 and 1		−0.024		0.010	
		(0.019)		(0.041)	
Rain between 1 and 2		−0.010		-0.012	
		(0.020)		(0.075)	
NREGS times Rain <-2				0.015	
				(0.056)	
NREGS times Rain between -1 and -2 times				0.094	
				(0.062)	
NREGS times Rain between 0 and -1 times				0.037	
				(0.052)	
NREGS times Rain between 0 and 1 times				-0.010	
				(0.049)	
NREGS times Rain between 1 and 2 times				0.056	
				(0.083)	
NREGS				-0.049	−0.012
				(0.054)	(0.037)
Ordinal rainfall (cuts -1 and 1)			0.021**		0.060*
			(0.010)		(0.035)
NREGS times Ordinal rainfall					−0.037
					(0.037)
District FE	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes
State/Year of Birth FE	Yes	No	No	No	No
Observations	65791	65791	65791	88547	88547

Standard errors are in parentheses and are clustered at the district level. The first column repeats results from **??** but adds state by wave fixed effects. Column two creates “bins” of rainfall. Column three uses a simple dummy variable equal to one if rainfall is greater than $Z = 1$. Column four defines an ordinal variable, similar to Jayachandran (2006).

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

Table A3: Rainfall and Female - Two-Way FE Robustness

	TWFE	Dist FE Only	Year FE Only
Year of birth rainfall (Z)	0.053***	0.052***	0.053***
	(0.020)	(0.019)	(0.018)
Rainfall (z) times NREGS	−0.045**	−0.044**	−0.046**
	(0.020)	(0.020)	(0.018)
NREGS	−0.024	−0.028	−0.025
	(0.037)	(0.036)	(0.032)
District Vars	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes
Observations	88547	88547	88547

Standard errors are in parentheses and are clustered at the district level. The first column repeats the main results. The second column removes the year fixed effects while the third column removes district fixed effects but includes year fixed effects.
* p<0.1 ** p<0.05 *** p<0.01

Table A4: Rainfall and Height-for-Age - Two-Way FE Robustness

	TWFE	Dist FE Only	Year FE Only
Year of birth rainfall (Z)	0.014	0.008	0.058***
	(0.022)	(0.022)	(0.021)
Female	−0.084**	−0.082**	−0.079**
	(0.036)	(0.036)	(0.035)
Female times Rainfall	0.044	0.042	0.023
	(0.030)	(0.031)	(0.030)
NREGS	−0.048	−0.426***	−0.035
	(0.090)	(0.077)	(0.098)
Rainfall (z) times NREGS	0.028	0.039	0.021
	(0.037)	(0.039)	(0.034)
Female times NREGS	0.116*	0.089	0.086
	(0.062)	(0.063)	(0.065)
NREGS times Female times Rainfall	−0.118*	−0.116*	−0.098
	(0.060)	(0.063)	(0.062)
District Vars	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes
Observations	19373	19376	19373

Standard errors are in parentheses and are clustered at the district level. The first column repeats the main results. The second column removes the year fixed effects while the third column removes district fixed effects but includes year fixed effects.
* $p<0.1$ ** $p<0.05$ *** $p<0.01$

Table A5: Testing the Parallel Trends Assumption - Sex Ratio

	Phase trends	District trends	Banks
Rainfall times NREGS	−0.068***	−0.058*	−0.054***
	(0.023)	(0.035)	(0.020)
Year of birth rainfall (Z)			0.071***
			(0.020)
Rainfall times Bank			−0.019**
			(0.009)
NREGS	0.026	−0.007	−0.020
	(0.043)	(0.044)	(0.037)
High bank			0.042
			(0.093)
Phase FE	Yes	No	No
District FE	No	Yes	Yes
Year of Birth FE	Yes	Yes	Yes
District Vars	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes
Observations	88547	88547	85217

Standard errors are in parentheses and are clustered at the district level. Columns one and two use the National Sample Survey and the dependent variable is whether a newborn is female. Columns three and four use the IHDS and the dependent variable is height-for-age Z score.
* p<0.1 ** p<0.05 *** p<0.01

Table A6: Testing the Parallel Trends Assumption - Height

	Phase Trends	District Trends	Banks	Health Centers
Year of birth rainfall (Z)			0.032	0.040
			(0.031)	(0.031)
Rainfall (z) × Female			0.026	0.009
			(0.044)	(0.038)
Rainfall (z) × NREGS × Female	−0.076	−0.052	−0.126*	−0.114*
	(0.100)	(0.129)	(0.065)	(0.060)
Rainfall (z) × High bank × Female			0.019	
			(0.049)	
Rainfall (z) × Health Centers × Female				0.065
				(0.043)
District FE	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes
Observations	19,373	19,373	18,404	19,373

Standard errors are in parentheses and are clustered at the district level. Columns one and two use the National Sample Survey and the dependent variable is whether a newborn is female. Columns three and four use the IHDS and the dependent variable is height-for-age Z score.
* p<0.1 ** p<0.05 *** p<0.01

Table A7: Rainfall and Number of Newborns at District/Year Level

	Girls	Boys
Current rainfall (Z)	0.044**	0.012
	(0.020)	(0.019)
District FE	Yes	Yes
Year FE	Yes	Yes
District Vars	Yes	Yes

Standard errors are in parentheses and are clustered at the district level. All regressions use the National Sample Survey and are at the district/year level. The dependent variable in the first column is (log of) number of newborn girls, defined as girls under one year of age. The dependent variable in the second column is (log of) number of newborn boys.
* p<0.1 ** p<0.05 *** p<0.01

Table A8: NREGS, Rainfall, and Child Gender – Two-Way Clustering

	Years 2001-2011			Years 2005-2009
	Model 1	Model 2	Model 3	Model 4
Year of birth rainfall (Z)	0.019***	0.053***	0.054***	0.052**
	(0.006)	(0.018)	(0.018)	(0.023)
Rainfall times NREGS		−0.045**	−0.046**	−0.054**
		(0.018)	(0.018)	(0.022)
NREGS		−0.024	0.017	0.006
		(0.037)	(0.041)	(0.038)
Year of Birth FE	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes
Phase Linear Trend	No	No	Yes	No
Observations	65790	88547	88547	38451

Standard errors are in parentheses and are clustered at both the district level and the state-year level to account for spatial correlation in rainfall within a year.
* p<0.1 ** p<0.05 *** p<0.01

Table A9: NREGS, Rainfall, and Child Height-for-Age - Two-Way Clustering

	District FE					Village FE
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Year of birth rainfall (Z)	0.026	0.002	0.025	0.035*	0.014	0.057*
	(0.020)	(0.027)	(0.026)	(0.020)	(0.026)	(0.031)
Female	−0.087**	−0.091**	−0.047	−0.047	−0.084**	−0.080
	(0.041)	(0.042)	(0.031)	(0.030)	(0.037)	(0.052)
Female times Rainfall		0.051	−0.003		0.044	0.023
		(0.037)	(0.033)		(0.037)	(0.038)
NREGS				0.003	−0.048	−0.037
				(0.074)	(0.076)	(0.109)
Rainfall (z) times NREGS				−0.028	0.028	0.017
				(0.039)	(0.050)	(0.053)
Female times NREGS					0.116*	0.094
					(0.060)	(0.062)
NREGS times Female times Rainfall					−0.118	−0.098
					(0.074)	(0.073)
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13307	13307	19373	19373	19373	19373

Standard errors are in parentheses and are clustered at the district level (columns one through five) or village level (column six). Columns one and two include children born between the years 1998 and 2005, though only 3.38 percent of observations come from prior to 2002. Columns three through six includes children born during the years 1998-2011 (only 1.20 percent of observations are from prior to 2002). The dependent variable in all columns is height-for-age, standardized using the CDC charts and the *zanthro* command in Stata (Vidmar et al., 2004). Rainfall is always defined as rainfall during year of birth. * p<0.1 ** p<0.05 *** p<0.01

B Conceptual Framework

Building on the framework developed in Eswaran (2002), Rosenblum (2013) and Balakrishnan (2017), we present a simple theoretical model to demonstrate that, in a society where the birth of a girl is associated with very high future costs (relative to a boy), a favorable agricultural productivity shock can increase the survival of female children relative to male children. Using this model, we then show that access to employment opportunities outside the agricultural sector can alleviate this effect.

B.1 Baseline Set-up

We model a representative household's decision of how much to invest in their child conditional on whether they have a girl or a boy child. We do this by comparing two identical households except than one birthed a boy and the other a girl. Both households live for two periods (for the second period, we denote the discount factor by β). Every period, both households have one unit of total labor hours that they can engage in productive enterprises to generate income. In the baseline model, agriculture is the only productive enterprise available to households. Therefore, for both households, labor supply to the agricultural sector is inelastic and equal to one during every period[1]. The agricultural production function that both households face is given by $\alpha_t F(L_t)$, where $t = 1, 2$ and notes the time period, α_t is the productivity parameter, and $F(.)$ follows $F(.) > 0$ and $F(.) < 0$.

In the first period, the households' derive utility ($u(.)$) from consumption alone and choose the health investments in their child (k_b for the household with a male child and k_g for the household with a female child). The instantaneous utility function is increasing and concave. The probability of a child's survival into the second period ($p(.)$) is increasing

[1]The literature on missing labor markets in developing countries motivates this assumption. While this is a realistic assumption in India's context, allowing for a lower degree of incomplete labor markets should not change the fundamental mechanism that increases in productivity shocks will increase the survival rates of girls compared to boys.

in the investments made during the first period, k_j (where $j = b, g$) and concave[2]. In the second period, the household derives utility from consumption and also from its surviving child (v) (same for both male and female child). If the child is male and survives into the second period, then the household receives a net benefit (D) in the second period. If the child is female and survives into the second period, then the household incurs a net cost (D) in the second period.

So, we assume that the household has no intrinsic preferences over male and female children. However, their preference for a son stems from the future benefits he brings and the future costs associated with a daughter. The net benefit from an alive son can be their labor income, the dowry receipts, and the labor income of their spouses. The net cost of an alive daughter may be the loss in labor incomes and dowry payments upon their marriage (Bhalotra et al., 2016; Anukriti et al., 2017)[3].

Therefore, the household's optimization problem is as follows if it has a female child in the first period:

$$\underset{k_g}{\text{maximize}} \quad U^g = u(\alpha_1 F(1) - k_g) + p(k_g)\beta u(\alpha_2 F(1) - D) + (1 - p(k_g))\beta u(\alpha_2 F(1)) + p(k_g)\beta v \tag{6}$$

Similarly, the household's optimization problem is as follows if it has a male child in the first period:

$$\underset{k_b}{\text{maximize}} \quad U^b = u(\alpha_1 F(1) - k_b) + p(k_b)\beta u(\alpha_2 F(1) + D) + (1 - p(k_b))\beta u(\alpha_2 F(1)) + p(k_b)\beta v \tag{7}$$

The above baseline model gives the following testable prediction:

Testable Prediction 1 *If the net cost of having a girl is large enough relative to a boy, positive*

[2]Therefore, we assume that the function of survival is not different between female and male children.

[3]Assuming intrinsic lower preference for a female child compared to a male child will not alter the propositions of this model but will instead magnify them.

agricultural shock during a child's year of birth leads to more health investments in the child if the child is female relative to if the child is male. Consequently, in response to a positive agricultural shock, the increase in infant girls' survival rates is more substantial than the increase in infant boys' survival rate. The proof is as follows:

The optimal health investments a female child is given by[4]:

$$\frac{\partial U^g(k_g)}{\partial k_g} = -\frac{\partial u(\alpha_1 F(1) - k_g^*)}{\partial k_g} + \beta \frac{\partial p(k_g^*)}{\partial k_g}[u(\alpha_2 F(1) - D) - u(\alpha_2 F(1)) + v] = 0 \tag{8}$$

The optimal health investments a male child is given by:

$$\frac{\partial U^b(k_b)}{\partial k_b} = -\frac{\partial u(\alpha_1 F(1) - k_b^*)}{\partial k_b} + \beta \frac{\partial p(k_b^*)}{\partial k_b}[u(\alpha_2 F(1) + D) - u(\alpha_2 F(1)) + v] = 0 \tag{9}$$

After the total differentiation of equation Equation 8 with respect to α_1, we:

$$\frac{\partial k_g^*}{\partial \alpha_1} = \frac{F(1)}{1 + \frac{\beta \frac{\partial^2 p(k_g^*)}{\partial k_g^2}[u(\alpha_2 F(1) - D) - u(\alpha_2 F(1)) + v]}{\frac{\partial^2 U(\alpha_1 F(1) - k_g^*)}{\partial k_g^2}}} > 0 \tag{10}$$

After the total differentiation of equation Equation 9 with respect to α_1, we:

$$\frac{\partial k_b^*}{\partial \alpha_1} = \frac{F(1)}{1 + \frac{\beta \frac{\partial^2 p(k_b^*)}{\partial k_b^2}[u(\alpha_2 F(1) + D) - u(\alpha_2 F(1)) + v]}{\frac{\partial^2 U(\alpha_1 F(1) - k_b^*)}{\partial k_b^2}}} > 0 \tag{11}$$

As $p''(.) < 0$ and $u''(.) < 0$, if D is large enough then:

$$\frac{\partial k_g^*}{\partial \alpha_1} > \frac{\partial k_b^*}{\partial \alpha_1} \tag{12}$$

[4]We focus on interior solutions. Specifically, we look at cases where the agricultural productivity parameter (α_1) is not large enough such that $p(k_g^*) < 1$. This also implies that $p(k_b^*) < 1$. This follows from rearranging Equation 8 and Equation 9, that gives us $k_g^* < k_b^*$).

This proves **Prediction 1**.

The intuition for Testable Prediction 1 is as follows: A household chooses a child's optimal health investment by investing up to the point that the decrease in the first period's utility that the health investments cause (by decreasing consumption) is equal to the increase in the second period's expected utility from having that child survive. The net cost of having a daughter and the net benefit of having a son implies that the marginal utility from having a female child survive into the second period is lower than the marginal utility from having a male child survive into the second period. Therefore, all other things equal, the optimal health investment in a female child is less than the optimal health investment in a male child.

Due to the lower optimal health investment in a female child compared to a male child, a positive agricultural shock leads to a larger marginal utility for a household with a female child than one with a male child (due to the concave utility function). This difference provides households with a female child a greater incentive to invest in their child than those with a male child in response to a favorable agricultural shock.

B.2 Introducing NREGS

The Mahatma Gandhi Rural Employment Guarantee Act entitles every rural household in India to 100 days of employment in public works at the state-level minimum wage. Therefore, we assume that the introduction of NREGS introduces a non-agricultural sector in the economy where the household can supply labor hours, N_t, for a fixed wage, w_N.

Therefore, the household's optimization problem in Equation 6 changes as follows if it has a female child in the first period:

$$\begin{aligned} \underset{k_g, N_1, N_2}{\text{maximize}} \quad U^g_{NREGS} =& u(\alpha_1 F(1-N_1) - k_g) + p(k_g)\beta u(\alpha_2 F(1-N_2) - D) \\ &+ (1-p(k_g))\beta u(\alpha_2 F(1-N_2)) + p(k_g)\beta v \end{aligned} \tag{13}$$

The household's optimization problem in Equation 7 changes as follows if it has a male child in the first period:

$$\begin{aligned} \underset{k_b, N_1, N_2}{\text{maximize}} \quad U^b_{NREGS} =& u(\alpha_1 F(1-N_1) - k_b) + p(k_b)\beta u(\alpha_2 F(1-N_2) + D) \\ &+ (1-p(k_b))\beta u(\alpha_2 F(1-N_2)) + p(k_b)\beta v \end{aligned} \tag{14}$$

The introduction of NREGS to the baseline model gives the following testable prediction:

Testable Prediction 2 *If the net cost of having a girl is large enough relative to a boy, the difference in the elasticity of health investments with respect to agricultural productivity between a female and male child is lower in the presence of NREGS.* The proof is as follows:

After the introduction of NREGS, the optimal health investments a female child is given by:

$$\begin{aligned} \left.\frac{\partial U^g(k_g)}{\partial k_g}\right|_{NREGS} =& -\frac{\partial u(\alpha_1 F(1-N_1^*) + w_N N_1^* - k_g^*|_{NREGS})}{\partial k_g} \\ &+ \beta \frac{\partial p(k_g^*|_{NREGS})}{\partial k_g}[u(\alpha_2 F(1-N_2^*) + w_N N_2^* - D) - u(\alpha_2 F(1-N_2^*) + w_N N_2^*) + v] \\ &= 0 \end{aligned} \tag{15}$$

After the introduction of NREGS, the optimal health investments a female child is given by:

$$\begin{aligned} \left.\frac{\partial U^g(k_b)}{\partial k_b}\right|_{NREGS} =& -\frac{\partial u(\alpha_1 F(1-N_1^*) + w_N N_1^* - k_b^*|_{NREGS})}{\partial k_b} \\ &+ \beta \frac{\partial p(k_b^*|_{NREGS})}{\partial k_b}[u(\alpha_2 F(1-N_2^*) + w_N N_2^* + D) - u(\alpha_2 F(1-N_2^*) + w_N N_2^*) + v] \\ &= 0 \end{aligned} \tag{16}$$

Also, the first order condition with respect to N_t in both periods, yields:

$$\alpha_t F'(1 - N_t^*) = w_N \tag{17}$$

That is, the household supplies labor to NREGS up to the point where incremental earning from agriculture equals the NREGS wages.

Using Cramer's Rule, we get:

$$\frac{\partial k_g^*|_{NREGS}}{\partial \alpha_1} = \frac{F(1-N_1^*)}{1 + \frac{\beta \frac{\partial^2 p(k_g^*|_{NREGS})}{\partial k_g^2}[u(\alpha_2 F(1-N_2^*)+w_N N_2^* - D) - u(\alpha_2 F(1-N_2^*)+w_N N_2^*)+v]}{\frac{\partial^2 U(\alpha_1 F(1-N_1^*)+w_N N_1^* - k_g^*|_{NREGS})}{\partial k_g^2}}} > 0 \tag{18}$$

$$\left.\frac{\partial k_b^*}{\partial \alpha_1}\right|_{NREGS} = \frac{F(1-N_1^*)}{1 + \frac{\beta \frac{\partial^2 p(k_b^*|_{NREGS})}{\partial k_b^2}[u(\alpha_2 F(1-N_2^*)+w_N N_2^* + D) - u(\alpha_2 F(1-N_2^*)+w_N N_2^*)+v]}{\frac{\partial^2 U(\alpha_1 F(1-N_1^*)+w_N N_1^* - k_b^*|_{NREGS})}{\partial k_b^2}}} > 0 \tag{19}$$

Note that $N_1^* \geq 0$, $u'(.) > 0$, $u''(.) > 0$, and the sum of agricultural and NREGS earnings is never less than only agricultural earnings for the same realization of the agricultural productivity parameter. Then for a large enough D,

$$\frac{\partial k_b^*|_{NREGS}}{\partial \alpha_1} > \frac{\partial k_b^*}{\partial \alpha_1}, \text{ and} \tag{20}$$

$$\frac{\partial k_g^*}{\partial \alpha_1} > \frac{\partial k_g^*|_{NREGS}}{\partial \alpha_1} \tag{21}$$

Subtracting Equation 20 from Equation 21, we have:

$$\frac{\partial k_g^*}{\partial \alpha_1} - \frac{\partial k_b^*}{\partial \alpha_1} > \frac{\partial k_g^*|_{NREGS}}{\partial \alpha_1} - \frac{\partial k_b^*|_{NREGS}}{\partial \alpha_1}. \tag{22}$$

This proves **Prediction 2**.

The intuition for Testable Prediction 2 is as follows: The introduction of NREGS, causes household income to be at least as substantial as it would have been without NREGS (for the same realization of the productivity parameter). The high net cost of having a daughter implies that the daughter's optimal health investment was low enough without NREGS to justify an increase in this investment after the household's income increases from NREGS. Due to the lower optimal health investment in a female child without NREGS compared to with NREGS, the household that does not have access to NREGS will have a more substantial marginal utility from investing in the daughter during a positive agricultural shock than the household with access to NREGS (due to the concave utility function). This marginal utility difference provides households with no access to NREGS, a greater incentive to invest in their daughter in response to a favorable agricultural shock compared to a household with access to NREGS.

Contrarily, households with a son derive a large net benefit if the child survives. This substantial gain implies that the son's optimal health investment was high enough without NREGS to justify a small decrease in this investment and a corresponding increase in consumption expenses after the household's income increases from NREGS. The lower optimal health investment in a male child without NREGS compared to with NREGS implies that the former household without NREGS will have a more substantial marginal utility from investing in the son during a positive agricultural shock than the household with access to NREGS (due to the concave utility function). This marginal utility difference provides households with access to NREGS, a greater incentive to invest in their son in response to a favorable agricultural shock compared to a household without access to NREGS.

The above opposite behavior of households with female and male children implies a decline in the gender gap in the elasticity of health investments with respect to agricultural productivity observed in Prediction 1 after an increase in and stabilization of income through NREGS.

There may be other possible channels for this result, such as an increase in women's bargaining power. We discuss this in more detail in section 5.

C Additional Robustness Checks

Table A2 presents several additional robustness checks, mostly related to specification choices. We restrict the samples in columns one through four to years before NREGS in a district. We first show that the inclusion of state-by-year fixed effects does not affect our conclusions. Their addition increases the coefficient to 0.026, suggesting that even within-state variation in yearly rainfall is a significant predictor of child gender. We also explore different definitions of our rainfall variables, including bins, and an ordinal variable described in Jayachandran (2006). In all cases, the empirical evidence supports that, in the early 2000s, boys were more likely to survive – or be born – than girls when households faced anything other than a positive income shock. This table also shows the interaction between NREGS and these definitions of rainfall. In these specifications, the interaction's coefficient is in the expected direction and consistent with the main results; however, the estimates are imprecise.

Three tables in the appendix present several tests for standard error-related concerns. We first test for district-level autocorrelation for rainfall in Table A1. We find no evidence of such autocorrelation after accounting for state fixed effects, district fixed effects, or district and year fixed effects. Since rainfall may exhibit spatial correlation within a year, Table A8 and Table A9 show the main results by adding two-way clustered standard errors at the district level and the state-year level. The results are qualitatively similar, though, in the height-for-age models, there is a loss of some precision.

Lastly, recent work argues that two-way-fixed-effect models can bias the average treatment effects (de Chaisemartin and d'Haultfoeuille, 2019) Table A3 and Table A4 presents results for infant gender and height-for-age, respectively, by removing either the year or district fixed effects. The results are qualitatively unchanged, regardless of specification.