Protecting Girls from Droughts with Social Safety Nets*

Jagori Chatterjee[†] **Joshua D. Merfeld**[‡]

September 18, 2019

Abstract

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 has shown 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 bad rainfall years. We find evidence of both prenatal and postnatal channels in India and show that a workfare program, which decouples both wages and consumption from rainfall, attenuates the relationship between rainfall and the infant sex ratio. Using a back-of-the-envelope calculation and the assumption that females should never significantly outnumber males, we find that the program could have saved around 550 girls per district per year – relative to boys – if the government had implemented it in the years 2001 to 2005. Additional results on postnatal channels show substantial impacts on long-run health outcomes of surviving girls, as rainfall no longer differentially affects girls' height-for-age, relative to boys', following the implementation of the program. In an important deviation from previous research, the entirety of the relationship between the sex ratio and rainfall is apparently driven by sex-selective abortions, not infant mortality.

Keywords: sex ratio, gender-gap, infant mortality, child health, income shock, consumption smoothing, workfare program, National Rural Employment Guarantee, India
JEL Codes: H53, I15, I38, O12

*We are grateful to Rachel M. Heath for her encouragement and guidance since ideation. We thank Chris Anderson, James D. Long, Jonathan Morduch, Xu Tan, and the participants at the 2019 Royal Economic Society Annual Conference, the 2019 Midwest Economics Association Annual Meeting, and the 2018 Pacific Northwest Labor Day Workshop 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 dissertation support through the Grover and Creta Ensley Graduate Fellowship for Economic Policy for this work.

[†]Department of Economics, Furman University; jagori.chatterjee@furman.edu

[‡]Robert F. Wagner Graduate School of Public Service, New York University; merfeld@nyu.edu

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, whereas in countries where families provide equal care for both daughters and sons, the sex ratio is expected to be approximately even (Sen, 1992).

It has been argued that economic development might help improve the sex ratio over time. Yet, 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 actually worsen the sex ratio, as families have fewer children 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, especially if we aim to design policies to combat this decline. We contribute to this literature by documenting the importance of agricultural productivity shocks and sex-selective abortions and by showing that government policy may be able to alleviate at least some aspects of the problem.

A notable driver of sex selection is transient economic shocks. In particular, income shocks in developing countries may exacerbate the already skewed sex ratio. In India, a primarily agrarian society with a strong son preference, Rose (1999) finds that the prob-

[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).

ability a child born during a given year is a girl is increasing in rainfall. In other words, 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 decide to continue with a pregnancy only if the child is a boy, and terminate the pregnancy if the child is a girl to help smooth consumption during this adverse time. Moreover, even for fetuses carried to term, there is a gender-gap in prenatal investments in societies with a strong son preference and this gap may be exacerbated during bad economic times. 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).

If transient economic shocks do indeed lead to such outcomes, some sex selective abortions may be the result of a lack of formal – or even informal – insurance 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 smooth consumption during bad years.

We show how deviations of district-level rainfall from its long-run average – a proxy for agricultural productivity shocks – interact with the roll-out of a large national workfare program in India, the Mahatma Gandhi National Rural Employment Guarantee Scheme (NREGS).[3] A growing body of literature on the program sheds light on the guaranteed employment program's effect on various development outcomes, including wages (Imbert and Papp, 2015; Merfeld, 2019b), consumption (Jha et al., 2011; Ravi and Engler, 2015), risk (Foster and Gehrke, 2017; Gehrke, 2017; Fetzer, 2014; Merfeld, 2019a), and time allocation decisions (Shah and Steinberg, 2015). 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.

Consistent with previous literature, we find that rainfall continues to be a significant predictor of the gender of an infant and girls' height-for-age 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. Instead, we show that the effect of rainfall is increasing in the state-level abortion rate but is unaffected by the state-level infant mortality rate, suggesting prenatal,

[3]NREGS guarantees up to 100 days of wage employment at the state-level minimum wage in a financial year to every household in India whose adult members are willing to do unskilled manual work.

not postnatal, decisions have the largest effect on the sex ratio. 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 prior to 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.

We next 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, as well. In particular, we explore the relationship between rainfall during the year of birth and child anthropometrics.[4] We first confirm that rainfall during the year of birth is a significant predictor of height-for-age in India, 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 the implementation of 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.

Multiple mechanisms could plausibly explain these findings. First, NREGS increased rural wages (Imbert and Papp, 2015; Merfeld, 2019b). As such, increased income is one plausible mechanism. However, we find few effects of NREGS on our outcomes of interest in average rainfall years, suggesting this is unlikely to be the primary mechanism driving our results. Second, NREGS could lead to an increase in bargaining power for women,

[4]We have a single anthropometric measurement for children up to nine years of age.

given its focus on female employment. However, the evidence on this front is somewhat limited (Sukhtankar, 2016). Besides, the limited effects of NREGS during average rainfall years suggests that any bargaining power mechanism would have to manifest itself primarily through decisions related to rainfall, which seems unlikely. Moreover, we find that female workdays provided by NREGS does not correlate with the reduction in the effect of rainfall post implementation, nor does NREGS have a large effect on total female-specific employment, either in absolute terms or relative to men. Finally, Santangelo (2019) finds that NREGS attenuates the relationship between rural wages and rainfall in India. As such, the program appears to help families smooth consumption. Given the dynamics underlying the results in this paper, we believe this is indeed the primary mechanism for the effects of NREGS on the infant sex ratio and anthropometrics. However, this conclusion is somewhat speculative.

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. Families also tend to sex-select more in response to adverse shocks for their subsequent children if their first surviving child is a girl, though we find no concrete statistical evidence of differences based on the gender of the first surviving child. 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. We assume that the number of boys is unaffected by rainfall and that the number of girls never exceeds the number of boys; the empirical evidence presented in this paper generally supports both of these assumptions. Using a back-of-the-envelope calculation, we estimate that if NREGS had been implemented in the years 2001 to 2005, approximately 550 additional girls per district per year would have survived to one year of age, relative to boys. This calculation only quantifies gains based on the sex ratio and rainfall, so it is an admittedly crude measure of the true effects of the program.

This paper primarily contributes to three strands of existing literature. First, this work fits into the research on how sex selection and child health are affected by changing economic conditions (Rose, 1999; Bhalotra et al., 2016; Qian, 2008). Previous research has found that improving economic conditions may increase the number of girls born in countries with a strong preference for sons, like India and China (Qian, 2008). Also, improving economic conditions may lead to improved child health and survival more generally, including within developing countries (Tolonen, 2018). Changes in permanent economic conditions need not be the primary driver of these improvements, as both idiosyncratic and aggregate economic shocks are essential predictors of child health (Yang, 2008). Like Rose (1999), we confirm that transient shocks are a significant predictor of both the infant sex-ratio and surviving girls' health outcomes in India, at least before the implementation of NREGS. However, two recent papers find opposite results. Chari et al. (2019) finds that NREGS increased total infant mortality. Banerjee and Maharaj (2019) find that NREGS does not mitigate the effects of heat on infant mortality. These papers' findings are not necessarily in conflict with our findings, however, as they generally focus on post-natal outcomes (although many post-natal outcomes are of course driven by pre-natal decisions and investments). We look at the sex ratio and our results suggest sex-selective abortions – a completely prenatal decision – are a key driver of the sex ratio-rainfall relationship.

Second, the study contributes to the literature on the effectiveness of different government policies for the well-being of girls, such as greater political participation of women (Beaman et al., 2012; Bhalotra and Clots-Figueras, 2014; Kalsi, 2017) and financial incentives offered when families 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 affects decisions regarding human capital investments in children, especially girls. Given the importance of early-childhood investments for adult outcomes (Maluccio et al., 2009), social safety nets may be an effective way to improve child health in contexts in which household income is highly seasonal, and consumption smoothing is difficult.

Finally, these results add to the growing literature on the risk-mitigation effects of rural workfare programs and the subsequent impact on development outcomes (Fetzer, 2014; Foster and Gehrke, 2017). A growing body of evidence suggests that NREGS has had a substantial impact on several different development issues of interest, although concerns regarding corruption persist (Niehaus and Sukhtankar, 2013). In particular, the program appears to play an insurance role, protecting households against adverse shocks (Zimmermann, 2018). Along these lines, previous research has shown that the program allows families to reallocate resources away from low-risk, low-return employment and towards riskier, higher return opportunities (Gehrke, 2017; Merfeld, 2019a). This paper suggests that this reduction in risk and increased ability to smooth consumption have important implications for health outcomes for girls in India.

We organize the rest of this paper as follows. We describe our methods and data, including summary statistics, in section 2. We then turn to the results in section 3, where we discuss our main findings as well as the identifying assumptions before turning to the mechanisms and implications. We conclude in section 4.

2 Data and Empirical Strategy

We combine data from several sources for this paper. First, 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 district centroids to the closest grid in the CRU data to create yearly district-level rainfall. Our primary measure of rainfall shock is 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.

Second, the data on the gender of infants is from the National Sample Surveys (NSS), a nationally-representative household survey. The NSS is collected by the Government of India's Ministry of Statistics and Programme Implementation (MOSPI). We use the 2004-05, 2007-08, and 2011-12 "thick waves"[5] of the NSS. This data records the age 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. For example, the 2007-08 wave is used to construct the panel for all surviving children born between the 2004-05 and 2007-08 waves, while the 2011-12 wave is used to build the panel for surviving children born between the 2007-2008 and 2011-2012 waves. Our final sample includes 89,264 infants (less than one at the time of enumeration), 98,980 one-year-olds, and 107,764 two-year-olds. Due to the nature of the NSS data, we use the sample of surviving children in our analysis. However, this is similar to previous studies (Rose, 1999; Kalsi, 2017) that investigate female mortality. The sample of surviving children is advantageous as it minimizes bias in estimated effects due to misreporting of dead children, which is systematically higher for female children.

Third, the data on children's anthropometrics come from the 2011-12 wave of the India Human Development Survey (IHDS II). The IHDS II collects height and age of all children between 0 and 9 years of age of interviewed households.[6] Using this information, we construct gender-specific height-for-age measures using the Center for Disease Control's

[5]These waves are the largest, comprising hundreds of thousands of households. Smaller surveys are administered by MOSPI during some other years.

[6]The IHDS only collected this information for children who were present at the time of the survey.

(CDC) growth charts. Specifically, we use the *zanthro* command in Stata, developed by Vidmar et al. (2004). Our final sample is 18,141 children between 0 and 9 years of age in 2011-2012.

Fourth, 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. Fifth, as additional controls, we use several district-level variables from the 2001 Census. Lastly, we use data on the number of banks in a district from the Reserve Bank of India's directory on commercial banks.

2.1 Infant Sex-ratio

We empirically estimate impacts of the program on the infant sex ratio and surviving children's height-for-age. We cluster standard errors at the district level in all specifications. First, we use the following specification to test whether agricultural productivity shocks continue to predict the gender of an infant in India during the 21^{st} century:

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

where $Girl_{idt}$ is an indicator variable that is equal to one if a surviving infant i born in district d and year t is a girl and is equal to zero otherwise. $BirthYear_t$ is a vector of the year of birth dummies and captures year-specific shocks common to all districts. $District_d$ is a vector of the district of birth dummies and controls for time-invariant differences across districts.

$RainZscore_{dt}$ is our proxy for agricultural productivity shocks. We measure it as the deviation of district d's rainfall in year t from its 10-year mean and scaled by its 10-year standard deviation.[7] Therefore, in Equation 1, our primary coefficient of interest is β.

[7]Apart from a linear rainfall variable, we also estimate Equation 1 with other definitions of the rainfall shock to explore the possibility of a non-linear relationship between agricultural productivity shocks and infant sex-ratio. These extensions include rainfall shocks defined as follows: (a) six indicator variables that

Given findings in previous literature (Rose, 1999), we expect β to be positive, that is, higher rainfall will increase the probability that an infant born is a girl. This essentially argues that the gender of a surviving child observed in our sample is more likely to be a girl if they were born in a relatively high-rainfall year.

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 as-good-as random with respect to the infant sex ratio. In stricter specifications, we add a vector of interactions of 2001 Census variables at the district level, described above, with the year of birth dummies. 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 our measure of rainfall shock and with initial conditions. In further specifications, 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. Also, in an additional specification, we estimate Equation 1 conditional on rainfall shocks a year before and a year after birth to check that rainfall during the year of birth is the relevant productivity shock and also alleviate concerns about serial correlation in our measure of rainfall shock.

Second, we explore the effects of NREGS on the infant sex ratio. We use the following main specification to estimate NREGS's effect on the relationship between agricultural productivity shocks in a year and the gender of an infant born in that year:

$$\begin{aligned} Girl_{idt} =& \alpha + \beta_1 RainZscore_{dt} \times NREGS_{dt} + \beta_2 RainZscore_{dt} + \beta_3 NREGS_{dt} \\ &+ \delta_d District'_d + \tau_t BirthYear'_t + \epsilon_{idt}, \end{aligned} \tag{2}$$

where $NREGS_{dt}$ is an indicator that is equal to one if district d in year t received the

capture bins of deviation of annual rainfall from its long-run mean; (b) a single dummy that indicates good rain and is equal to one for rainfall above +1 SD; and (c) an ordinal rainfall variable with two cut points at -1 and +1 standard deviation, as in Jayachandran (2006).

guaranteed workfare program. β_1 is our primary coefficient of interest in Equation 2 and identifies the relationship between rainfall shocks and infant sex-ratio in districts that receive NREGS compared to the districts that do not. We hypothesize that NREGS helps households smooth consumption and that in turn reduces the use of sex-selection as an instrument of consumption-smoothing. Therefore, we expect β_1 to be negative; that is, NREGS attenuates the positive effect of rainfall on the probability that a surviving infant is a girl.

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)[8,9]. The phased in 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 β_1 is that the sensitivity of infant sex-ratio to our measure of rainfall shock does not trend differently in districts with and without exposure to NREGS. However, the phase in was not random and districts that receive the program at different points in time look quite different and time trends may differ by district and/or phase (Imbert and Papp, 2015). As such, we present additional evidence that the identifying assumptions hold in Section 2.3.

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

[9]We code $NREGS_{dt}$ 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.

2.2 Long-run Effects

Next, we investigate the effect of early-life agricultural productivity shocks on the surviving children's health, by gender, with the following specification:

$$\begin{aligned} HAZ_{idt} =& \alpha + \beta_1 Rainzscore_{dt} \times Girl_{idt} + \beta_2 Rainzscore_{dt} + \beta_3 Girl_{idt} \\ &+ \delta_d District_d + \tau_t Birthyear_t + \mu_{idt}, \end{aligned} \tag{3}$$

where HAZ_{idt} is height-for-age Z-score of child i born in district d and year t and $Girl_{idt}$ is an indicator for whether this child is a girl. β_1 in Equation 4 is our primary coefficient of interest, and we expect it to be positive if rainfall shock during the child's year of birth has long-lasting effects and girls received differentially more investment compared to boys in years with good rainfall.

To explore how the guaranteed workfare program changed the relationship between early-life agricultural productivity shocks and children's health by gender we use the following specification:

$$\begin{aligned} HAZ_{idt} =& \alpha + \beta_1 NREGS_{dt} \times Rainzscore_{dt} \times Girl_{idt} + \beta_2 Girl_{idt} \times Rainzscore_{dt} \\ &+ \beta_3 NREGS_{dt} \times Girl_{idt} + \beta_4 Rainzscore_{dt} \times NREGS_{dt} + \beta_5 NREGS_{dt} \\ &+ \beta_6 Girl_{idt} + \beta_7 Rainzscore_{dt} + \delta_d District_d + \tau_t Birthyear_t + \mu_{idt}. \end{aligned} \tag{4}$$

In Equation 4, β_1 is our primary coefficient of interest, and consistent with Equation 2 we hypothesize that it will be negative. That is, NREGS will attenuate the positive effect of good agricultural shocks during a year on the long-run health outcomes of girls born during that year relative to boys.

2.3 Identification Assumptions and Robustness Checks

We mention in Section 2.1 that our identification strategy is similar to a triple difference. Therefore, the identification of β_1 in Equation 2 relies on the assumption of parallel trends. In other words, for our estimate of β_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 identically in the absence of the program. This parallel trends assumption 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 low-caste individuals (Scheduled Castes and Scheduled Tribes).[10] To address this concern, we execute several tests.

First, we present additional results of ES2, where we include phase-specific rainfall trends (or district-specific rainfall trends for an even stricter specification). These flexible specifications allow the response of infant sex-ratio and height-for-age to rainfall shocks to trend differently for each phase of NREGS implementation (or district), which enables us to partly control for possible differential pre-treatment trends by treatment status.

Second, 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 actual implementation and dropping all observations after the NREGS treatment. So that the placebo test does not have any districts treated in reality[11]. By assigning NREGS to districts before 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,

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

[11] 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.

we estimate the following:

$$
\begin{aligned}
Girl_{idt} =& \alpha + \beta_1 RainZscore_{dt} \times placeboNREGS_{dt} + \beta_2 RainZscore_{dt} \\
&+ \beta_3 placeboNREGS_{dt} + \delta_d District'_d + \tau_t BirthYear'_t + \epsilon_{idt},
\end{aligned} \tag{5}
$$

where $placeboNREGS_{dt}$ 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 Equation 2 makes the coefficients directly comparable. This specification ensures that we use the same spatial variation as NREGS's implementation and also similar temporal spacing. If pre-program trends are responsible for our results, we expect to see similar results in the placebo test, that is, β_1 and β_2 in both Equation 2 and Equation 5 will be in the same direction. Importantly, this specification *tests for different trends in how our outcome variable responds to rainfall by treatment status*, which can confound our main estimates of interest and not just different trends in the outcome variable by treatment status. We perform a similar placebo test for the height-for-age variable.

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, the Reserve Bank of India implemented the Branch Authorization Policy in 2005 to expand the growth of bank branches in financially backward districts (Young, 2017). Using an older rural bank expansion policy, Rosenblum (2016) shows that greater access to bank branches affects the gender gap in child mortality outcomes. 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, construction of Anganwadi centers[12] is a permissible activity under NREGS and therefore may have accelerated with NREGS's implementation and can confound our estimates. To

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

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[13] in Table A2 and Table A3. 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 A3 as this information is only available for the IHDS sample.

2.4 Summary Statistics

We present summary statistics for the main variables used in the rainfall 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, on average.

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, as well as 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

[13]Highly banked districts are defined as those that have a population to bank ratio equal to more than the national median during a year.

districts than in phase-two, and phase-three districts, which may suggest that parents in poorer districts are less likely to sex-select.

3 Results

3.1 Main Results

We begin with a simple graphical representation of our primary motivation in Figure 1. The figure shows kernel-weighted polynomial regressions of infant gender (left panel) and child's height-for-age by gender (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 the survival and health of girls. 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.

We move to a more robust empirical examination of this relationship in Table 2. In all columns, the dependent variable is a dummy variable indicating whether a child is female. Since data are nationally representative, this is effectively equivalent to analyzing the gender of a randomly selected surviving child (Rose, 1999). In the first five columns, we restrict estimation to children under the age of one. In columns one through three, we examine the relationship between rainfall and child gender before the implementation of NREGS. Column one presents the results from the simplest specification of Equation 1. The coefficient on rain indicates 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) increases the probability that a randomly chosen infant is female by 1.4 percentage points. Adding more control variables for district characteristics and household characteristics in columns two and three increases the estimated effect size slightly; the coefficient in both columns 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 rainfall is approximately 2.08 standard deviations. Our results indicate that the probability a randomly-selected infant is a girl could increase by almost four percentage points when moving from the 25th percentile of rainfall to the 75th percentile. If accurate and if girls do not ever significantly outnumber boys, this suggests that adverse rainfall shocks could be a significant contributor to "missing" women in South Asia, a point to which we return later.

Column four in Table 2 removes the sample restriction of pre-NREGS years and estimates the relationship in Equation 1 for the years 2001-2011[14]. 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 something happened between 2006 and 2011 to attenuate this relationship. Next, we present empirical evidence that NREGS may be responsible.

In the fifth column of Table 2, we add the previous year's rainfall and the following year's rainfall to the regression. Both coefficients are small and statistically insignificant. The evidence is consistent with the argument that agricultural productivity shocks right around birth or during pregnancy are the most important determinant of (female) infant survival. We explore this possibility further in columns six and seven. Rose (1999) found that rainfall during up to school going age had a significant impact on the survival of female children relative to male children. While the results in column five suggest this is unlikely to be the case in our sample, we now test this explicitly. Column six investigates the effect of rainfall during the year that a child is one-year-old on the probability that the child is female. Column seven repeats this for the sample of two-year-olds. In neither column six nor seven is the coefficient on rainfall significant. The coefficient in column six

[14]Recall that in 2006 phase one districts, 2007 phase two districts, and 2008 phase three districts received NREGS.

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.

Table A1 presents several robustness checks of the relationship tested in Equation 1, 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, a dummy for good ($Z \geq 1$) rainfall, and an ordinal variable as in Jayachandran (2006). In all cases, conclusions are quite robust across these specification choices. In other words, the empirical evidence supports the contention 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.

Two tables in the appendix present several tests for standard error-related concerns. We first test for district-level autocorrelation for rainfall in Table A7. 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 also be spatially correlated within a year, column one of Table A8 presents results from column three of Table 2, adding two-way clustered standard errors at both the district level and the state-year level. Qualitative conclusions are unchanged.

We next move to an analysis of the effects of NREGS on the relationship between rainfall and newborn gender in Table 3. We hypothesize that NREGS could attenuate the relationship between rainfall and child gender if the program helps households smooth consumption in the face of negative rainfall shocks. In columns one through four, the coefficient on "Year of birth rainfall (Z)" is always positive, suggesting the effect of rainfall on the probability of being female is positive prior to NREGS's implementation, consistent with the estimates in Table 2. The coefficient on NREGS is never close to significant, indicating NREGS did not have a direct impact on the sex ratio during "normal" years, when rainfall is approximately equal to the district's ten-year mean. In fact, the coefficient is actually negative in the first three columns, suggesting NREGS, if anything, decreased the probability a randomly-selected child was female. However, there is no ex ante reason to expect NREGS to decrease the number of girls, and the insignificant results in columns one through three and the positive coefficient in column four are consistent with this.

The coefficient of interest is the interaction term between rainfall and NREGS, in the first row, which represents the change in the effect of rainfall on the sex ratio following implementation of the program.[15] The interaction term is negative, suggesting this relationship decreases markedly following 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 (results not shown), suggesting NREGS almost completely reverses the relationship between rainfall and the sex ratio. Additionally, the coefficients are very stable across specifications. Column two adds year-of-birth fixed effects, column three adds household variables, and column four adds phase linear trends to insulate the estimates from possible differences in trends by phase prior to program implementation. Reassuringly, the results remain surprisingly consistent across these columns.

Columns one through four utilize the entire panel we have constructed, from 2001-2011. While we include district and year-of-birth fixed effects, there may still remain concerns that we are isolating variation in years far removed from NREGS implementation. To test this possibility, in column five we restrict estimation only to the years 2005-2009, one year prior to NREGS to one year following the final phase of NREGS. Though the results are slightly more imprecise, conclusions are unchanged. In fact, the interaction term

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

is now slightly larger than the rainfall coefficient, though the linear combination is not significantly different from zero, similar to columns one through four. 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 NREGS is actually implemented; this is exactly what we observe in column five.

Table A8 presents additional specifications from Table 3, adding two-way clustered standard errors to account for possible spatial correlation in rainfall. Standard errors are clustered at both the district and state-year level. Qualitative conclusions are unchanged.

The previous results focused on the sex ratio. However, it may be that the effect of rainfall at the time of birth also extends to long-run indicators of human capital investments for *surviving* girls, like anthropometrics. Figure 1 has already presented suggestive evidence that this is indeed the case. Table 4 presents a number of different specifications exploring this possibility more formally. The dependent variable in all columns is height-for-age, defined using CDC growth charts.[16] Column one estimates the effect of year-of-birth rainfall on height-for-age in 2012. The coefficient on rainfall is positive but small and insignificant. The coefficient on the female dummy is negative and significant, confirming the existence of differential investment in boys and girls in India, independent of rainfall.

The coefficient on female in conjunction with our previous results of the effect of rainfall on newborn gender raises the possibility that rainfall may differentially affect height-for-age for boys and girls. To explore this possibility, the specification in column two adds an interaction between female and rainfall. The coefficient on rainfall – which now represents the effect of rainfall on boys' height-for-age – decreases to almost exactly zero. Moreover, the coefficient on the interaction term between rainfall and female is positive and marginally significant, and the linear combination of this coefficient with the coefficient on rainfall is also significant (results not shown; p=0.049), suggesting rainfall is significantly

[16]We translate height and age into the height-for-age z-score using the user-written *zanthro* command in Stata (Vidmar et al., 2004).

correlated with girls' height-for-age. Putting these two results together, rainfall during year of birth apparently 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. If so, then the true results are actually much stronger than the results in Table 4 would indicate (Barcellos et al., 2014).

Column three again removes the pre-NREGS restriction and estimates the relationship over the years 1998-2012. Similar to Table 2, the result is no longer significant and, in fact, actually reverses, though the coefficients are small in magnitude. This again supports the contention that something changed between 2006 and 2012. Column four 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. Consistent with the pooled results in column one, it does not appear that NREGS affects the relationship between rainfall and height-for-age. Nonetheless, the coefficients are in the expected direction and the coefficient on rainfall – which now represents the effect of rainfall on height-for-age prior to NREGS implementation – is marginally significant.

Column five allows the effects of NREGS to vary by gender, a possibility suggested by the previous results in this paper. We find further evidence that NREGS impacts human capital investments differently for girls and boys. In particular, the triple interaction of *NREGS* × *Female* × *Rainfall* is negative and significant, indicating NREGS attenuated the relationship between rainfall and height-for-age more for girls than for boys. In addition, the effect of NREGS on boys' height-for-age during normal rainfall years (the coefficient on NREGS) is insignificant, while the effect of NREGS on girls' height-for-age relative to boys' (the coefficient on Female times NREGS) is positive. In other words, there is some evidence that NREGS increased girls' height-for-age relative to boys' during average rainfall years, though the coefficient is only marginally significant at the ten-percent level, so some caution is warranted.

Since these results use the IHDS, we are also able to control for village fixed effects, which we do in column five. Many health outcomes are determined at levels below the district – due to differences in medical care, nutrient availability, etc. – so the inclusion of village fixed effects might be expected to alter the estimated impact of the program. However, it appears that the inclusion of village fixed effects has no effect on our substantive conclusions and increases precision – consistent with the argument that many healthcare-related outcomes are determined at a more local level – providing further evidence that NREGS improves girls' human capital outcomes in poor agricultural years, relative to boys'.

3.2 Possible Explanations

In this section, we discuss several possible explanations for our findings: failure of the identifying assumptions, changes in women's bargaining power, and consumption smoothing.

3.2.1 Failure of the Identifying Assumptions

The first possible explanation for our findings is that the difference-in-differences assumption of parallel trends does not hold. Since we have many years of data, one way to control for possible violations in this assumption is to include rainfall trends in our main results. We do this in Table A2 and Table A3 of the appendix. The inclusion of these trends allows the effects of rainfall on the dependent variables of interest to be trending differentially in each district. Importantly, these are not simple year trends, but are instead trends in the effect of rainfall, a more direct test of the identification assumptions. The models are otherwise identical to the results in Table 3 and Table 4. 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.

Table A2 presents results for the sex ratio. Column one 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 actually slightly larger than our previous results. Table A3 presents the results for height-for-age. We modify the specifications slightly by allowing the rainfall time trends to differ by gender. Column one is again the phase-specific trends specification while column two is the more flexible district-specific trends specification. The coefficient of interest, the triple interaction between NREGS, rainfall, and female, remains negative in both columns but is slightly attenuated. However, the standard errors are twice as large as standard errors in other columns, making inference difficult.

We move to a more common test of pre-trends in Table 5. As described in subsection 2.3, we test the plausibility of our identification assumptions by constructing a "placebo" NREGS variable, with implementation moved up by three years and dropping districts when they receive treatment (to avoid any of the years of implementation). If pre-trends are responsible for our results, we would expect to see similar results in Table 5 as we did in the previous results section. Columns one and two of Table 5 present the corresponding results for the gender sample, with district-specific rainfall trends added in column two. The first coefficient in each column is the coefficients of interest. In both columns, the coefficients are actually positive and are significantly so in the first column. In other words, the effect of rainfall on the sex ratio was actually increasing in treated districts relative to untreated districts in the three years just prior to NREGS implementation. This is suggestive evidence that pre-trends are not responsible for our results.

Columns three and four present the corresponding estimates for the IHDS sample and height-for-age. Column four 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 three, the coefficient is small in magnitude while in column four, the coefficient is positive, suggesting our results may actually be underestimating the true effect of NREGS. However, it is worth noting that the coefficient in column four is quite imprecisely estimated, making inference difficult. In conjunction with the results in Table A2 and Table A3, there appears to be little evidence that differential trends are responsible for our results, but we interpret the height-for-age results with a bit of caution due to the slight attenuation observed in Table A3.

Finally, another possible explanation is that concurrent events and/or policies were implemented along with NREGS. For example, access to both banking and health centers was expanded significantly around the dates of implementation. In addition, both banks and health centers could plausibly affect our outcomes of interest. To explicitly test for this, we add controls for bank and health center expansion in the last columns of Table A2 and Table A3. Our substantive conclusions are completely unchanged by these inclusions.

3.2.2 Women's Bargaining Power

Women were originally intended 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). In addition, according to one government advisor, the mandated minimum program wage often double the prevailing wage rate for women at the time.[17] 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. In addition, if the future (expected) earnings of girls increase, parents may respond by increasing investments in girls' human capital – e.g. education and nutrition (Balakrishnan, 2017; Heath and Mobarak, 2015) – 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.

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

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 (i.e. when rainfall is equal to the district's ten-year mean). This coefficient is insignificant in all of the sex-ratio specifications. This is *prima facie* evidence that NREGS did not increase the bargaining power of women, at least in average rainfall years, since this means any bargaining power would have to affect our outcomes only through effects of rainfall. We present additional suggestive evidence in Table A5 of the appendix. For each district, we calculate the total number of female workdays provided by NREGS during the first year of implementation, under the assumption that this number is correlated with women's employment and bargaining power. The results in columns one and two show that this figure is negatively correlated with the effect of rainfall on the sex ratio prior to NREGS. In other words, in districts with higher amounts of female NREGS employment during implementation, there is *less* sex selection prior to the program, consistent with our assumption. However, using this employment number in the differences-in-differences specification, we find no evidence of an attenuation of the effect of rainfall for higher levels of employment; in fact, the coefficient is positive.

Finally, using the NSS data we also estimate simple differences-in-differences for employment for both men and women with the 2004/05 (pre-NREGS) and 2007/08 (during implementation) waves. NREGS is significantly correlated with more employment for men relative to women and is only positively associated with more public works for women in absolute terms. As such, while we cannot rule it out, there is little evidence suggesting that an increase in women's bargaining power is the main mechanism driving the results.

3.2.3 Consumption Smoothing

The final possible channel through which NREGS affects these gendered outcomes is through its ability to help households smooth consumption. Santangelo (2019) shows that NREGS decouples wages and consumption from rainfall. We replicate these results in Table A4 using the Additional Rural Incomes Survey Rural Economic Demographic Survey.[18] Importantly, the triple interaction between rainfall, NREGS, and post is negative and strongly significant, suggesting the effect of rainfall on consumption is more negative following implementation of NREGS. In conjunction with the results in Santangelo (2019), this suggests an increased ability to smooth consumption during bad rainfall years following implementation of NREGS is a possible mechanism explaining our results.

3.3 Birth Order Heterogeneity

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 is increasing in birth order, with relatively little sex selection of first-born children (Bhalotra and Cochrane, 2010). If this is the case, we would expect to see insignificant effects of rainfall on the gender of first-born children and, correspondingly, no 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 one explores the effects of rainfall on the gender of all children of the household head. We do this to confirm that our results hold when examining the children of the head only. The magnitude on rainfall is positive and strongly significant and of a similar magnitude to our rainfall coefficient in column three Table 2, suggesting that our results hold when restricting attention to just children of the head.

We present the rainfall results using only the first-born child of the household head in column two of Table 6. Consistent with previous literature, the coefficient is negative,

[18] http://adfdell.pstc.brown.edu/arisreds_data/readme.txt

small and insignificant[19]. This result suggests that rainfall is uncorrelated with the gender of first-born children in India. In column three, 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 children. In other words, the magnitude must be larger for non-first-born children to compensate for the null effect among first-born 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, we believe the magnitude and direction of the coefficients can help 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; both the coefficients are marginally significant ($p<0.15$). Consistent with the null effect of rainfall shock on first-born children in column two, column five finds no impact of NREGS on the relationship between rainfall shock and the gender of the first-born child. Column six 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 five and six arguably strengthen the identification assumptions, as the results suggest any spurious trends would have to be specific to non-first-born children only.

Child gender preferences may lead to differential effects of rainfall on future children based on the existing gender composition of siblings. Table 7 explores heterogeneity along this dimension. The first two columns split the sample by the gender of the first surviving

[19]Note that the "first-born" child in our sample is constructed using the oldest living child still in the household. This construction ignores children who may have left the household before the survey. However, as the most aged children in our constructed data are five years old, their migration out of the household is very unlikely.

child in the household. Taking the coefficients at face value, it appears that sex selection on subsequent children is more likely during bad rainfall years if the first-born is a girl than if the first-born child is a boy. Column three pools the samples and includes an interaction term between rainfall and whether the first-born is a boy. Interestingly, the sign of the coefficient flips and the coefficient is far from significant. As such, we are unable to reject equal levels of sex selection based on rainfall by gender of the first-born child. Column four also includes a triple interaction with NREGS. It does not appear that NREGS differentially affects the impact of rainfall on the sex ratio based on the gender of the first-born child. However, our estimates are somewhat noisy so it is difficult to put too much confidence into this conclusion.

3.4 Pre- vs. Post-Natal Decisions

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 suggests 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 obvious possibility is sex-selective abortions. Though abortions for reasons of sex selection are generally illegal, the accessibility of abortions more generally has increased in the last few decades. To see whether sex-selective abortions, rather than postnatal outcomes between birth and one year of age, may be driving the relationship, we collect data on state-level abortion rates in 2000. If sex-selective abortions explain much of the finding, we expect to see the effect of rainfall on the sex ratio to be higher in states with higher abortion rates.

We test this hypothesis in Table 8. The first column confirms this is indeed the case: higher abortion rates are correlated with stronger effects of rainfall on the sex ratio. However, it may be the case that abortion access is also correlated with infant mortality, another possible driver of the relationship. In column two, we add an interaction between a state's infant mortality rate in 2000 and rainfall. The coefficient on the rainfall/abortion rate variable is almost completely unchanged, while the interaction between rainfall and the infant mortality rate, while imprecisely estimated, is quite small in magnitude. This is consistent with sex-selective abortions being the main driving force behind the relationship, as well as with recent literature that points to the role ultrasound technology has played in sex selection (Bhalotra and Cochrane, 2010; Chen et al., 2013). The facts that infant mortality is not correlated with the relationship and that we explore gender-specific outcomes rather than aggregate infant outcomes may explain why our results appear to diverge from those in Chari et al. (2019) and Banerjee and Maharaj (2019).

Column three presents 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; results are consistent with columns one and two. Columns four and five test whether NREGS had larger effects on the rainfall/sex ratio relationship when abortion rates are higher. The results are, unfortunately, very imprecisely estimated, but the signs and magnitudes of the coefficients are consistent with the general story.

3.5 NREGS, Rainfall, and "Missing" Girls

All the results presented in this paper suggest that adverse rainfall shocks had profound impacts on girls in the early 2000s. Importantly, 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 the probability of having a girl.

We estimate these numbers by assuming that the birth of boys is unaffected by rainfall. We analyze the plausibility of this assumption in Table A6. 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 survey 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 to 2001 to 2005) and include district fixed effects, year fixed effects, and the census variables interacted with year dummies.

The results in Table A6 show that an increase in rainfall increases the number of surviving girls born but does not significantly affect the number of surviving boys born during a year. A one-standard-deviation increase in the rainfall shock at the district centroid is associated with an increase in the number of girls by approximately 4.5 percent. However, column 2 shows no significant correlation between rainfall and the number of boys born each 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 more generally in response to rainfall shocks. If this were the case, we would expect to see a significantly positive coefficient on rainfall for boys, as well. This result also provides support to the hypothesis that decreasing early-life investment in female children is a primary consumption smoothing strategy but not reducing early-life investment in male children. These results are consistent with the results presented in autoreftable2.

We next assume that the number of boys and the number of girls in a district is equal when rainfall is approximately two standard deviations above its ten-year mean. Though the data actually suggest the number of girls is slightly higher at +2 standard deviations, we believe a more plausible assumption is that girls never outnumber boys, or that the "natural" sex ratio is approximately even. We choose +2 standard deviations since this is the level of rainfall at which the number of girls relative to boys is the highest (Figure 1). We next take the effect of rainfall on the probability of being a girl from the coefficient in column three 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 $Z = 2$. 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 1.4 million girls. This number translates to around 288,000 girls per year in all of India or around 550 girls per district per year. In other words, if NREGS had been available in the years 2001 to 2005, we estimate that about 1.4 million more girls relative to boys would have been alive in 2006. Though large, this calculation does not even take into account any improvement in the life of the surviving girls who receive more investments around birth and are thus healthier. However, this also does not take into account the possibility that NREGS has a detrimental impact on the infant mortality rate in absolute terms. As such, we are estimating a relative estimate (of girls relative to boys), not an absolute estimate.

4 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) reduces consumption volatility. Third, as a consequence, the introduction of NREGS mitigates the effect of income shocks on the sex selection of infants. Fourth, we find that before the advent of NREGS, a positive agricultural shock also more positively related to the health of surviving female children compared to male children. Lastly, NREGS mitigates this relationship between income shocks and health of girls.

This paper establishes that policies that are successful in providing tools for consumption-smoothing 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 insurance 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 the success of such policies is very sensitive to the design of these policies (Anukriti, 2018; Balakrishnan, 2017). Therefore, risk-mitigation and similar policies that help households smooth consumption may be an attractive development intervention, with favorable consequences for th e sex ratio and female health investments, as well.

Our results suggest that NREGS would have saved approximately 550 girls per district per year from the years 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? Unfortunately, we are not able to answer this question, which remains a big question for future research.

Several questions remain from our analyses. First, while we show that a girl is more likely to survive a poor rainfall year following implementation of NREGS, it is not clear whether this finding also suggests that household fertility will increase. In other words, is the surviving girl added to the counterfactual number of children, or does the surviving girl "replace" one of them? Given that NREGS was rolled out over just three years, we are unable to explore long-term fertility changes using our data and empirical methodology.

Second, this paper explores sex selection in response to a negative 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 take place through other channels. Given that one form of sex selection decreases following the implementation of the program, it seems reasonable to assume that sex selection, on average, must have also reduced. However, there are many other mechanisms apart from income shocks that may lead to sex selection, and our paper does not address these different possibilities.

References

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

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

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*.

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*.

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.

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.

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.

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).

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

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.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *The Economic Journal*, 119(537):734–763.

Merfeld, J. D. (2019a). Moving up or Just Surviving? Non-Farm Self-Employment in India. *American Journal of Agricultural Economics, forthcoming*.

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

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.

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. (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*.

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.

Tolonen, A. (2018). Local industrial shocks and infant mortality. *The Economic Journal*.

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.

Yang, D. (2008). International migration, remittances and household investment: Evidence from philippine migrantsâĂŹ exchange rate shocks. *The Economic Journal*, 118(528):591–630.

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*.

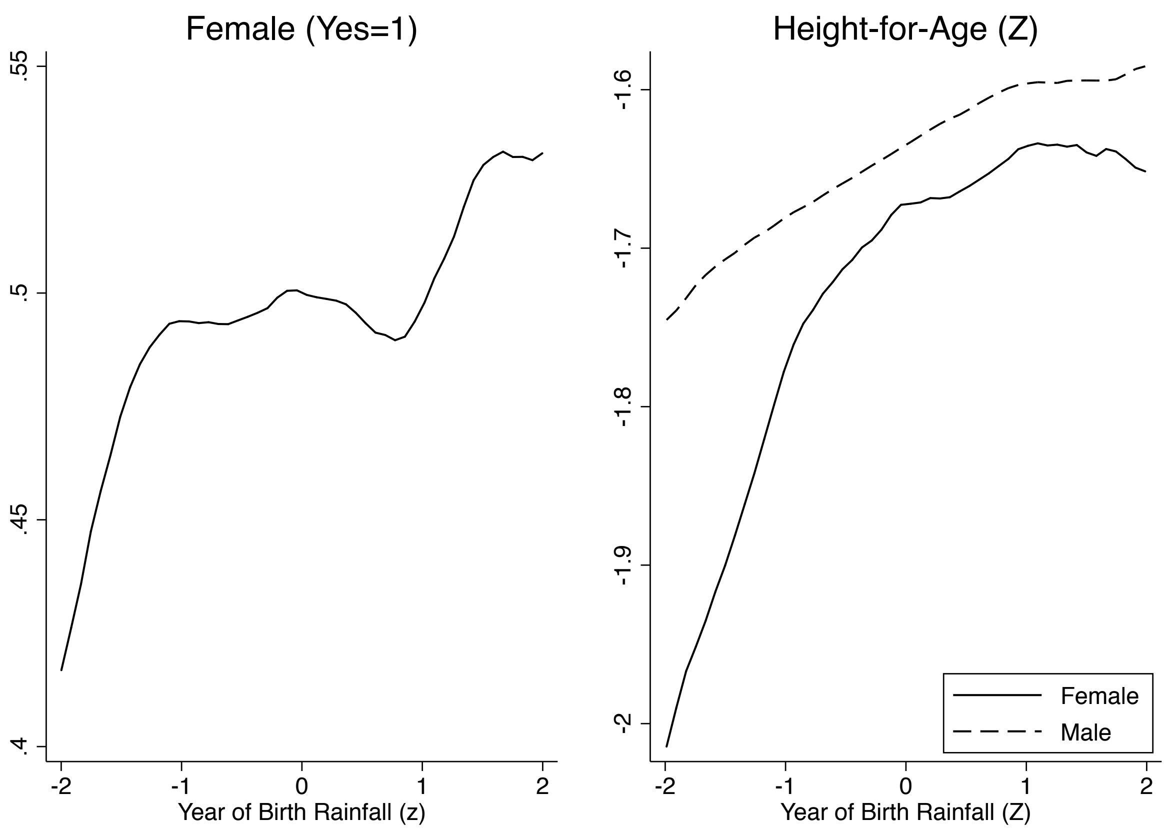
Figures & Tables

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: Gender of Older Siblings and Sex Selection

	(1) First-born boy	(2) First-born girl	(3) All	(4) All
Year of birth rainfall (Z)	0.018	0.032***	0.024	0.045
	(0.012)	(0.010)	(0.025)	(0.041)
First born was boy times Rainfall			0.018	0.020
			(0.050)	(0.047)
Rainfall (z) times NREGS				−0.045
				(0.043)
First born boy times NREGS				−0.076
				(0.047)
First born boy times NREGS times rain				−0.013
				(0.052)
First born boy			0.103**	0.116***
			(0.051)	(0.045)
NREGS				0.036
				(0.057)
District FE	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes
Observations	13900	13498	27398	36202

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 8: 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

Appendix: Results

Figure A1: Sex Ratio in India

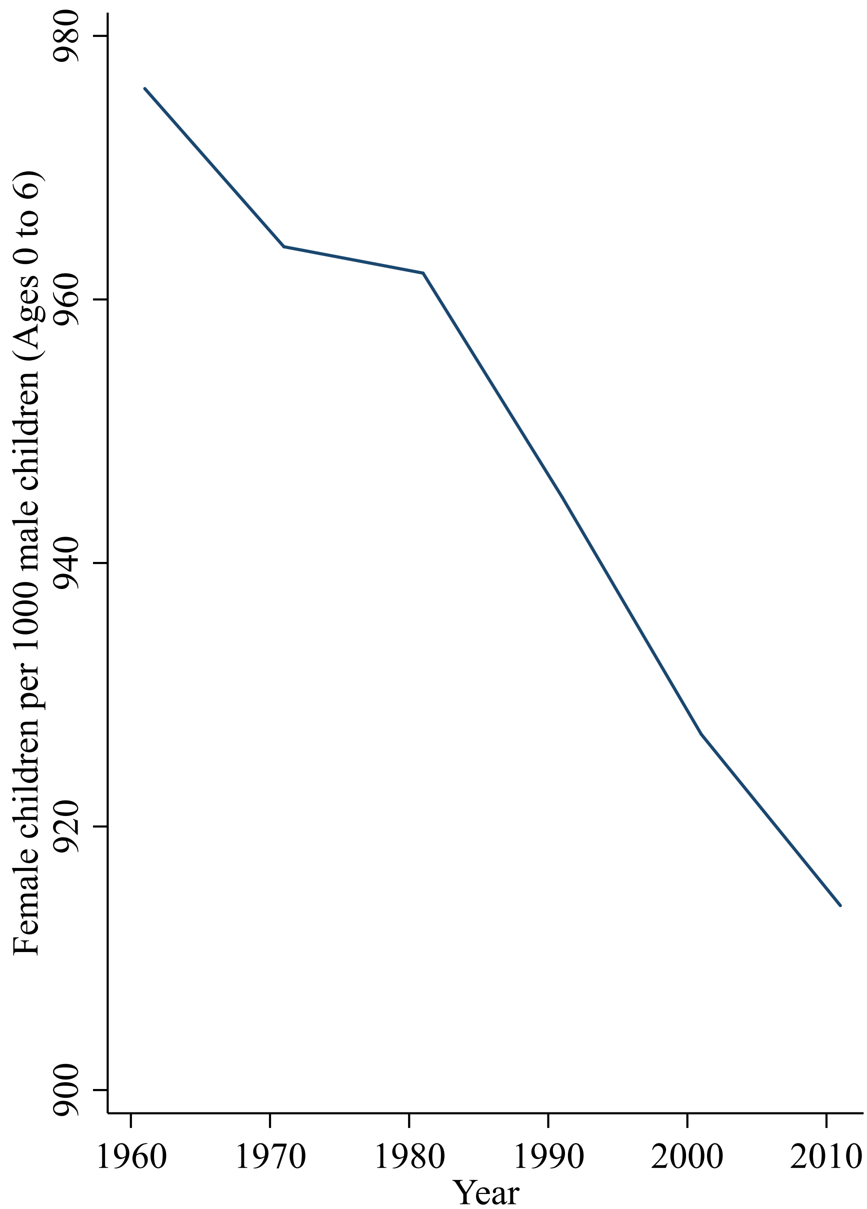


Table A1: Rainfall Robustness

	Rainfall Robustness				NREGS	
	Model 1	Model 2	Model 3	Model 4	All Years	2005-2009
Year of birth rainfall (Z)	0.026***					
	(0.007)					
Rain <-2		−0.085***				
		(0.029)				
Rain between -1 and -2		−0.050**				
		(0.024)				
Rain between 0 and -1		−0.045**				
		(0.019)				
Rain between 0 and 1		−0.024				
		(0.019)				
Rain between 1 and 2		−0.010				
		(0.020)				
Good year (Z>1)			0.027*			
			(0.015)			
Ordinal rainfall (cuts -1 and 1)				0.021**	0.060*	0.073
				(0.010)	(0.035)	(0.046)
NREGS times Ordinal rainfall					−0.037	−0.051
					(0.037)	(0.045)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes	Yes
State/Year of Birth FE	Yes	No	No	No	No	No
Observations	65,791	65,791	65,791	65,791	88,547	38,451

Standard errors are in parentheses and are clustered at the district level. The first column repeats results from Table 2 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 A2: 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 A3: 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 A4: Rainfall, Household Consumption, and NREGS

	Total Consumption	Food Consumption
Rainfall times NREGS times Post	−0.243***	−0.194**
	(0.081)	(0.080)
District FE	Yes	Yes
District Vars	Yes	Yes
Household Vars	Yes	Yes
Observations	7,996	7,992

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 A5: Female Bargaining Power and NREGS

	Outcome=Female		Employment (All Adults)			Employment (Women Only)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Babies	Babies	Self	Wage	Public Works	Self	Wage	Public Works
Year of birth rainfall (Z)	0.029***	0.066**						
	(0.007)	(0.028)						
NREGS		−0.022	0.049**	0.036**	0.013***	0.035	0.002	0.009***
		(0.043)	(0.021)	(0.017)	(0.003)	(0.023)	(0.012)	(0.003)
Female			−0.504***	−0.250***	−0.002***			
			(0.014)	(0.009)	(0.000)			
NREGS times Rain		−0.057**						
		(0.028)						
NREGS times Female days		−0.001						
		(0.010)						
Rain times Female days	−0.006**	−0.010						
	(0.003)	(0.012)						
NREGS times Rain times Female days		0.010						
		(0.012)						
NREGS times Female			−0.080***	−0.063***	−0.006***			
			(0.018)	(0.016)	(0.002)			
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Vars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	65791	88547	659450	659450	659450	328259	328259	328259

Standard errors are in parentheses and are clustered at the district level. The dependent variable in columns one and two is whether a baby is female. In column one, only pre-NREGS years are included, while all years are included in column two. Columns three through five estimate the effects of NREGS on employment. Observations are individual adults. Columns six through eight are similar, but estimation is restricted to women only.

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

Table A6: 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 A7: 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 A8: 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