Rainfall shocks, soil health, and child health outcomes

Siddharth Kishore*

Abstract

This paper estimates the moderating effect of soil organic carbon content (a measure of soil health) on child health in response to rainfall shocks in a low-income country setting. Focusing on rural India, I leverage the Demographic and Health Survey data set and high-resolution spatial data on soil organic carbon content and meteorological variables. Using a coarsened exact matching method, I show that a modest change in soil health can provide resistance to wasting in children during periods of low rainfall.

Keywords: Rainfall shock, soil organic carbon, child health, matching

1 Introduction

India consistently ranks low on the global hunger index, according to four indicators: malnutrition prevalence, child wasting, child stunting, and under-five mortality (Wiesmann, 2006). Many of India's villages in 2016 showed alarming levels of anthropometric measurements in children (Kim et al, 2021). According to the 2015-2016 India Demographic and Health Survey, 38% of children

*Ph.D. Candidate in the Department of Agricultural and Resource Economics at Colorado State University. I want to thank my advisor Dale Manning for his invaluable guidance and support. I am grateful to my committee members Jordan Suter, Alexandra Hill and Anita Pena for their detailed comments and support. Alma Calderón, Anna Dimitrova, Paul Wilson, Angela Mensah, Mubanga Chishimba and numerous seminar participants made helpful suggestions at various stages of this research. I thank Mostafa Shartaj and the participants at the Environmental Natural Resource Economics seminar at Colorado State University for their valuable suggestions. All errors are my own.

under the age of 5 are stunted (too short for their age) and 21% of children under the age of 5 are wasted (too thin for their height). Indian agricultural production is vulnerable to climate change and, without effective adaptation, can reduce food crop yields in the future by up to 9% (Guiteras, 2009). Moreover, in India's recent past, shortages of staple food crops, wheat and rice are associated with severe droughts and extreme rainfall (Zaveri and B Lobell, 2019; Auffhammer et al, 2012). Child nutrition and agricultural production in rural areas in the developing world are closely linked (Webb and Block, 2012). Bakhtsiyarava and Grace (2021) in Ethiopia demonstrated that more diversity in agricultural production during periods of low rainfall can reduce the risk of chronic food insecurity among children. Food shortages caused by crop failures due to extreme weather conditions and, therefore, nutritional deprivation can negatively impact children's health (Grace et al, 2012). Improved soil quality as measured by soil organic carbon (SOC), commonly used in the literature, increases agricultural production (Lal, 2006). Because of the water holding capacity, a high level of SOC offers long-term drought resistance and reduces the frequency of crop failures (Huang et al, 2021; Kane et al, 2021). SOC also provides agricultural profits for small landowners in developing countries (Bhargava et al, 2018). The research asks if SOC affects children's nutrition and health in a low-income country. Then, I explore to what extent SOC offers resilience during periods of low rainfall.

This article examines whether natural variation in soil organic carbon levels mitigates the impact of non-linear weather variables by crop growth on children's health. Focusing on rural India, I leverage the 2015 Demographic and Health Survey dataset and high-resolution spatial data on soil organic carbon content and meteorological variables. Following (Bakhtsiyarava and Grace, 2021), I evaluate the variation in anthropometric measurements, height-for-age (HAZ) and weight-for-height z-scores (WHZ) to measure child malnutrition in India. Inadequate nutrition can cause childhood stunting (if HAZ is below 2 standard deviation) and wasting (if WHZ is below 2 standard deviation). Unlike stunting (a measure of long-term inadequate nutrition), wasting (a measure of short-term inadequate nutrition) may be reversed by increasing nutritional intake (Victora, 1992). In this study, I focus on HAZ and WHZ to measure malnutrition linked to weather-induced food insecurity.

While the exact relationship between soil quality and crop production under dry conditions is complex and multidimensional. Huang et al (2021) and

Kane et al (2021) in the United States show that a higher soil organic carbon content can moderate the impact of weather shocks by retaining soil water in the agricultural systems. Children's nutrition also depends on food quality, which is partly dependent on soil micro-nutrients (Berkhout et al, 2019; Kim and Bevis, 2019). Berkhout et al (2019), based on their study in Sub-Saharan Africa, highlight the importance of soil micro-nutrients such as zinc, copper and manganese in reducing the malnutrition in children.

This article is informed and contributes to two main strands of the literature: the first is the relationship between soil agronomy and climate; the second is the relationship between children's health and SOC. While there are studies that examine the impact of climate on children's health in India (e.g., Dimitrova and Muttarak (2020) and McMahon and Gray (2021)), these studies have overlooked the importance of soil health. In this article, I contribute to the literature by demonstrating the direct and indirect effects of SOC. By enhancing the SOC, households would have access to greater food availability that could support children's nutrition and health. This is a direct result of SOC. The SOC may also help mitigate the impact of adverse weather conditions on food quantity. This is an indirect effect of SOC.

I find that higher soil organic carbon levels attenuate about 3% of the negative effect of rainfall shock on children's weight-for-height z-scores. I show that a small change in soil health can offer resistance to wasting in children during periods of low rainfall. I also explore heterogeneity in children's health outcomes by gender, household wealth index and land ownership, and climate zone. This suggests that efforts to improve soil quality should be adjusted to address these heterogeneous impacts. The results of the paper provide new evidence and inform policy-makers on the impact of high organic carbon in soils on children's health.

2 Conceptual Framework

Soil health has direct and indirect effects on childhood health. Weather-induced lower crop yields result in food shortages affecting food consumption and thus nutrition (Grace et al, 2012). The direct effect of high SOC levels is an increase in agricultural production (Lal, 2006), which contributes to food availability and supports nutrition. Moreover, SOC increases agricultural income (Bhargava et al, 2018) and can thus contribute to food security and nutrition for children. An indirect effect of high SOC levels is to ensure crop resilience following rainfall-induced agricultural shock and, therefore, to ensure food security. Because of the water holding capacity, a high level of SOC offers long-term drought resistance and reduces the frequency of crop failures (Huang et al, 2021; Kane et al, 2021).

The observed characteristics, such as that of the child, the mother and the household, affect the child's nutrition (Almond and Currie, 2011). These characteristics may affect children's nutrition differently depending on the level of education of the mother, the gender of the child and the wealth of the household. Moreover, SOC mitigation effects may vary depending on climate regions and the ability of households to cope with rain shocks. Later in the results section, I estimate the heterogeneity in children's health outcomes by region, climate zone, gender, household wealth and land ownership.

In addition, there may be unobserved covariates which may be correlated with children's nutrition and soil organic carbon levels and therefore may bias my results downwards. Figure 1 depicts a simple conceptual relationship between soil health and child health.

3 Data and Descriptive statistics

To demonstrate how soil organic carbon levels moderate the effect of monsoon activity on the health of Indian children, I leverage the Demographic and Health Survey dataset and high-resolution spatial data on soil organic carbon levels and weather variables.

3.1 Demographic and Health Data

I use the cross-sectional data from the fourth round of the Demographic and Health Survey (DHS) for India collected in 2015-2016. DHS uses a multi-stage stratified sampling design, with enumeration areas, hereinafter referred to as clusters (equivalent to census villages), being the smallest unit. In the clusters, households are randomly selected to be interviewed. DHS also collects the GPS locations of each cluster, enabling researchers to link DHS dataset to other geo-coded data, including soil organic carbon levels, precipitation, and temperature, at the cluster level. In order to preserve the anonymity of the villages, DHS randomly displaces the GPS coordinates of clusters up to 2 Km in urban areas and up to 5 Km in rural areas, and 1% of rural clusters are further displaced up to 10 Km. This displacement introduces measurement errors and may bias my results downwards.

131 of the 28,526 geo-referenced clusters did not have information and were dropped. I extracted environmental data using the DHS geo-referenced cluster for a 10-km buffer.[1]

DHS has a nationwide representative sample of children. In my analysis, the sample size for children aged 0 to 4 years was 259,627; 34,625 observations were excluded from the child data file that contained missing or invalid data. Invalid cases include children over plausible limits, age over plausible limits, and flagged cases. Additionally, observations with invalid woman's Body Mass Index (BMI) information (636 observations), missing data (6,447 observation) on caste, and not useful information (929 observations had "don't know" on caste) were excluded. Furthermore, I restrict the sample to focus exclusively on rural parts of the country as defined in the DHS dataset. To sum up, I analyzed a sample of 169,904 rural Indian children.

3.2 Rainfall Data

I draw monthly rainfall data from Climate Hazards Group Infrared Precipitation (CHIRPS) using DHS cluster geocordinates. CHIRPS is a quasi-global that extends over 50 S-50 N, with a gridded resolution of 0.05 degrees, from 1981 to near-real time precipitation time series (Funk et al, 2014).

There is not much guidance available in the literature about defining rain shock. For my purpose, I need to define a rainfall shock based on a threshold that lowers yields on India's major crops. Therefore, like Feeny et al (2021), I adopt an empirical strategy to determine the threshold. Using data from the International Crops Research Institute for the Semi-Arid Tropic (ICRISAT), I regress the natural log of the annual crop yield (Kg per hectare) from 2001 to 2015 on rainfall deciles controlling for year and district fixed effects.[2] The unit of analysis for the yield data is the district-year. As shown in the Figure 2, results indicate that rainfall below the 20th percentile reduces crop yield of grains and pulses in India.[3] I define rain shock as a monsoon rain that is below

[1] As a sensitivity test, I run every analysis for a 20-km buffer. Appendix Table A9 reports the main results.

[2] Crop yield data (unapportioned) are available at http://data.icrisat.org/dld/index.html

[3] In the appendix, Figure A1, I also show the negative effects of lower precipitation on selected staple and cash crops. Corn, soybeans and cotton appear to differ and not increase monotonously with precipitation, suggesting a non-linear response to weather conditions in some field crops.

the 20th percentile of the long-term historical mean within the DHS cluster (Shah and Steinberg, 2017).[4]

I used a measure of rainfall shock, which has already been used in the literature (Feeny et al, 2021; Dinkelman, 2017). Following (Dinkelman, 2017), I calculate the fraction of shocks:

$$\text{Fraction shocks}_{ij} = \frac{[\text{child's exposure to shocks in-utero through age 4}]_{ij}}{(\text{in-utero} + \text{child's age})_{ij}}.$$

where the subscripts i represent every child in the sample living in clusters j. By using the shock fraction, I capture the variation in the rain shock specific to the child living in the clusters.

A child under the age of 5 years may be exposed to one, many or no monsoon rainfall shock; the fraction of shocks captures that intensity of shock. For example, if a child of age 3 was exposed twice to rainfall shocks over his or her lifetime then the fraction of shocks for that child is given by 2/4. To measure the *in-utero* exposure to rainfall shock, I used the birthyear of the individuals observed in the DHS data.

To serve as a robustness check, I construct a population-weighted monthly rain measure based on gridded population data provided by the Center for International Earth Science Information Network (Center for International Earth Science Information Network - CIESIN - Columbia University, 2018).[5]

3.3 Growing Degree Days

Daily temperature was sourced from Indian Monsoon Data Assimilation and Analysis (IMDAA) reanalysis portal, managed by the National Centre for Medium Range Weather Forecasting (NCMRWF), India (Rani et al, 2021). Reanalysis Data Service (RDS) is a regional atmospheric reanalysis over the Indian subcontinent at a high resolution 0.12 x 0.12 from 1979-2018.[6] I have followed the formulation used in previous studies using meteorological measures which affect crop losses (Guiteras, 2009).[7] Using the maximum and

[4] India receives the majority of its rainfall during the monsoon from June to September.

[5] For my analysis, I use a resolution of 2.5 arc-minute for the year 2015. Data is available at https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-count-rev11/data-download

[6] Available at https://rds.ncmrwf.gov.in/datasets

[7] Following (Guiteras, 2009), I convert the daily mean temperature to GDD:

$$GDD(T)_j = \begin{cases} 0, & \text{if } T \leq 8C \\ T-8, & \text{if } 8C < T \leq 32C \\ 24, & \text{if } T \geq 32C \end{cases}$$

minimum daily temperature, the lower and upper threshold for calculating Growing Degree Days (GDD) during a growing season were set to $8C$ and $32C$, respectively.

3.4 Soil Data

Soil organic carbon data were obtained from OpenLandMap (Hengl and Wheeler, 2018).[8] Global soil maps were produced based on machine learning predictions from global soil profile compilations at a resolution of 250 m. Following (Huang et al, 2021), I extracted the mean soil organic carbon content around the DHS geo-coded clusters at four standard depths: 0, 10, 30, and 60 cm. I then calculated the depth-weighted soil organic carbon content at 0-60cm interval for the analysis.[9] The literature does not provide clear information about the threshold for classifying soil as high or low quality. Therefore, I have identified two categories of soil organic carbon content: low, below the 50th percentile, and high, above the 50th percentile.[10]

Figure 3 shows the soil organic carbon map for the rural DHS clusters. The missing area in the map indicates the null values for union territory Lakshadweep. Much of India is categorized as having low levels of soil organic carbon. The average soil organic carbon concentration is 0.945 %(g/Kg). Coastal regions in the west and east, most in the northeast and central plains are characterized by moderate to high soil carbon levels. Also, to explore what determines SOC variation, I do the Pearson correlation coefficient test between soil organic carbon and the historical enhanced vegetation index.[11] The Pearson coefficient of correlation between these two variables is 0.38 (p-val = 0.000).

3.5 Descriptive statistics

According to the World Health Organization, the z-scores compare the anthropometric measure of a child to the population mean value for his or her age and sex. The z-score value can be either negative or positive depending on whether

---

[8]Soil data are available at https://www.openlandmap.org

[9]Following (Huang et al, 2021), I used the trapezoidal rule to estimate the depth-weighted 0-60cm interval:

$$(S_{0-60cm})_j = \left(\frac{[(S_0 + S_{10}) * 10 * 0.5] + [(S_{10} + S_{30}) * 20 * 0.5] + [(S_{30} + S_{60}) * 30 * 0.5]}{60} \right)_j$$

[10]I also perform the sensitivity test for different threshold values such as 25th and 75th percentile of high soil organic carbon. Appendix table A7 and A8 report the results.

[11]I observe the enhanced vegetation index in the DHS dataset from 1985 to 2015 at 5-year intervals.

a child's anthropometric measurement is below or above the population average for the child's age and sex. Because the children in the sample have a negative value of z scores, suggesting infants with low birth weight, on average. The distribution of each anthropometric measure within the sample differs for boys and girls. Among boys, the height-for-age is -1.597, the weight-for-age z is -1.602 and the weight-for-height is -1.017. In girls, the height-for-age z score is -1.516, the weight-for-age is -1.572, and the weight-for-height is -0.963.

Figures 4a and 4b show the distribution of height-for-age (HAZ) and weight-for-height (WHZ) z scores of children under 5 years of age. The shaded portion in the figure shows the frequency indicating the absolute magnitude of child stunting and wasting. In my sample, approximately 41 per cent of children are stunted and approximately 21 per cent of children are wasted.

Table 1 reports the summary statistics for the data used in this study.[12] About 11 percent of children were exposed to at least one rainfall shock in their birth year and in-utero. Children aged 2 to 4 are more exposed to cumulative shocks ranging from 0.15 to 0.17. This means that children aged 2 to 4 may have been exposed to at least one rainfall shock in their lifetime. The average value of the fraction of shocks as an intensity measure is 0.13.

In my sample, the average age of children is 30 months. 51 per cent are boys and 49 per cent are girls. On average, mothers are 27 years of age and approximately half of the women have a high school or higher education. A little over half the households have agricultural land. Just under a third of households have potable water lines and a third have flush toilets. 23 per cent of families in my sample are poor. Table A3 in the appendix presents summary statistics for all control variables used in this study.

4 Empirical Framework

I estimate an OLS regression model to investigate the impact of high soil organic carbon levels on children's nutrition and health. mitigate the negative impact of shocks on children's health. The main specification is given by

$$h_{ij} = \beta_1 shock_{ij} + \beta_2 soc_j + \beta_3 (shock_{ij} * soc_j) + f(\theta)_{jt} + \xi \mathbf{X}_i + f(a)_i + \delta_d + \phi_{my} + \varepsilon_{ij} \tag{1}$$

[12] Appendix Table A2 describes the variables included in the research.

where h_{ij} denotes child health outcomes measured by the height-for-age, weight-for-age, and the weight-for-height z-scores for child i at the DHS cluster level, j; $shock_{ij}$ represents the fraction of rain shocks experience by child i residing at DHS cluster level, j; soc_j represents the mean soil organic carbon content at the DHS cluster level, j; $\mathbf{X}_i$ is the vector of explanatory variables including child, mother, and household characteristics. Child characteristics include age, gender and order of birth; mother characteristics include age, level of education and diet; and household characteristics include religion, social group, household income, and the wealth index (see Appendix Table A2 for a complete list of control variables); δ_d denotes district fixed effects and captures the time-invariant unobserved heterogeneity at the district level; ϕ_{my} denotes child birth year-month specific fixed effects and captures within cohort variations, and ε_{ij} denotes the disturbance terms. I cluster the standard errors at the level of DHS cluster (equivalent to Census village).

Additionally, I control precipitation and temperature derivatives (growth degree-days and harmful degree-days) during a growing season (June through September) throughout a child's life. $f(\theta)_{jt}$ is a function of monsoon rainfall and temperature at the DHS cluster level, j in time t. I followed (Dimitrova and Muttarak, 2020) to include a restricted cubic age spline, $f(a)_i$ with knots 6, 12, 18, 24, 36, and 48 months of age to control for non-linearity in children's growth trajectory. The key parameters are β_1, β_2, and β_3. β_1 represents the impact of cumulative periods of low precipitation on children's health; β_2 represents the direct impact of a high level of SOC on children's health; and β_3 represents the mitigation effects of a high level of SOC during cumulative periods of low rainfall.

In this study, I assume the soil endowments are exogenous. Because any change in agriculture, including climate change, takes a long time to get reflected in the soil system (Lal, 2004). This can mean that investment in soil or soil degradation by intensive cropping may take a long time to be reflected in the soil system. Also, because of India's low weather-induced internal migration rate (Viswanathan and Kumar, 2015). Because in my analysis, I look at short-term weather conditions on children's nutrition and health. That is a plausible assumption.

There may be a potential threat to identification. Some regions may experience larger declines in soil organic carbon content than others, resulting in measurement errors. For example, in wheat fields, stubble burning is often done after harvest, which can disrupt the natural cycle of soil organic carbon replenishment. However, because of the invariant time measure of the soil, I am unable to capture this variation.

Nevertheless, I take advantage of the coarsened exact matching method to estimate causal effects by reducing the covariate imbalance between treatment and control groups (Iacus et al, 2012). However, it may not circumvent the sample selection problem.

4.1 Matching methods

The coarsened exact matching method estimates the average effect of treatment on the treated sample (Blackwell et al, 2009). I use data knowledge to search for a better match. The coarsened variables used were: a) child-specific (child's birth order, child's gender and age); b) mother-specific (mother's age and education level); and c) household-specific (religion, caste, source of drinking water, and toilet facility).[13] I apply the software package, *cem* created by (Blackwell et al, 2009) was used to calculate the weights and these weights were used in a simple weighted regression.[14] The treatment variable treat, is 1 for high soil organic carbon content (in treatment group) and 0 for low soil organic carbon content (control group). Here is the summary of the match: the number of balanced matched observations is 51,148 for treatment and control; and the unmatched observation is 33,802 out of 84,950 for control and 33,806 out of 84,954 for treatment.

5 Results

5.1 Rainfall shocks, soil health, and child health

Table 2 presents impact of rainfall shock and soil health on children's health. The OLS model takes into account the characteristics of the child, the mother, the household. Moreover, the model controls a child's lifetime exposure to rain and temperature during a growing season. The model includes district and month and year of birth fixed effects. Standard errors are clustered at the DHS cluster level. The shock fraction shows a significant negative association

[13] I also included the month of birth as part of the matching algorithm. I calculated if a child was born during the dry season (the first six months of the year) or the wet season (the last six months of the year). Then I included that as an additional variable in the matching algorithm. Appendix Table A10 presents the results. It reads findings similar to those of the main specification.

[14] The *cem* command with a *k2k* option in STATA produces a match result which has the same number of treated and control in each matched strata by dropping the observations randomly.

with child WHZ. A one standard deviation increase in rainfall shock exposure above the child average years of exposure implies that the child will have 0.029 (0.161*0.182 = 0.029)[15] lower weight-for-height z score. A high level of SOC has no effect on children's health at its main term, but substantially reduces the negative effect of the precipitation shock by 13.6 percentage points.

The interaction term between SOC and fraction of shocks, which captures the compensating effect of a high soil quality. The coefficient on the interaction term is 0.136 and significant at the 5% significance level suggesting that a higher soil organic carbon content moderates approximately 3% $\left(\frac{-0.161+0.136}{-0.991} = 0.025\right)$[16] of the negative effect of monsoon rainfall on child weight-for-height z score.

Columns 3 and 4 in Table 2 present the results after applying the coarsened exact matching weights to the OLS model. The shock fraction is negatively related to the child's WHZ. The interaction term between SOC and fraction of shocks shows a positive association. However, the key coefficients are not significant at the 5 per cent significance level in the matched sample. In addition, I find no significant association between SOC and the child's HAZ for the full and matched sample. Appendix Table A4, which uses the population-weighted monthly rain measures, reads similar effects on child health.

Figures 5a and 5b illustrates the average marginal effects of high soil organic carbon on anthropometric measures in children. The figure suggests that a high level of SOC reduces the negative impact of rain shocks on children's WHZ. The attenuation effect of high SOC levels during periods of low precipitation is greater for high shock intensity. Graph a in Figure 5 shows an interesting result: the child's height-for-age z score shows an upward slope suggesting that the cumulative period of dryness improved the height-for-age z scores. However, at high shock intensity, the average marginal effect is statistically insignificant (as shown in graph c). This may be due to a reduction in diseases that are common during monsoons such as diarrhoea and malaria. But it requires further research and the results have to be interpreted with caution.

There is a concern that soil organic carbon measurement may be confounded by other associated agronomic attributes. With SOC as the choice variable, it is difficult to remove concerns related to the omitted variable bias.

[15] The standard deviation for the shock fraction variable is 0.182.
[16] The mean dependent variable is -0.991.

Nevertheless, I approach this concern by including soil texture, slope and vegetative index as control variables in Equation 1.[17][18] In order to assess the influence of the different soil attributes used in this study on children's health, I ran a correlation between child WHZ and soil attributes. This demonstrates no concern for multicollinearity in the model. Table A5 in the appendix provides the correlation matrix for the soil attributes used in this study. Appendix Table A6 report the results. It reads similar effects on child health.

5.2 Heterogeneity

5.2.1 Heterogeneity by climate zone

The impact of soil organic carbon on children's health can vary according to climate zones in India. Following Dimitrova and Bora (2020), I constructed six major climate zones at the district level based on the basis of the climate classification Köppen Geiger.[19] They are tropical wet, tropical wet and dry, arid, semi-arid, humid sub-tropical, and mountainous. See Appendix A2 for a map of the main climatic zones in India.

Table 3 presents the heterogeneous effects of a high level of SOC on children's health in some climatic zones. Each column of Table 3 presents the regression results for the separate climatic zones. Cumulative rain shocks have a negative impact on the health of children living in semi-arid and humid subtropical climate zones. The impact is greater in semi-arid climate zones. Point estimate is -0.280 and significant at the 5% significance level. The interaction term of the shock fraction with SOC is positive and significant at the 5% significance level, suggesting mitigating effects of a high level of SOC. The shock fraction positively affects the WHZ of children in tropical wet and dry and a high level of SOC decreases WHZ during the cumulative periods of low rainfall.

The results suggest no impact of a high level of SOC on the child's HAZ in any major but semi-arid climatic zones. Furthermore, a high level of SOC lowers the child's HAZ during periods of low rainfall.

---

[17] I used OpenLandMap to extract clay, sand, and silt content in $\%(kg/kg)$ at a depth of 60cm in the DHS cluster (Hengl, 2018a,b,c).

[18] I used the enhanced vegetation index for 2015 available in the DHS dataset as a proxy for agricultural output.

[19] I am grateful to Anna Dimitrova for sharing the data and code with me.

5.2.2 Heterogeneity by gender

Table 4 presents the heterogeneous effects of rainfall shocks and soil health on children's health by gender. Each column in Table 4 shows the separate regression results for boys and girls. Cumulative rain shock has a negative impact on girls' and boys' WHZ scores. Girls are more affected by rain shocks, as suggested by the larger coefficient. The point estimation is -0.205 for girls and -0.112 for boys. A high level of SOC positively impacts WHZ scores for girls during cumulative periods of low rainfall. The p-value of the test of the difference in the coefficient across girls and boys for the interaction terms between high SOC and fraction of shocks is not statistically different from zero. In addition, the results show that a high level of SOC does not affect the HAZ scores for boys and girls.

5.2.3 Heterogeneity by household wealth index

I observe five different indices of wealth in the DHS data: the poorest, the poorer, the middle, the richer, and the richest. For my purpose, I code the poorest and the poorer as the poor and the middle, the richer, and the richest as the non-poor.

Table 4 presents the heterogeneous effects of rainfall shocks and soil health on children's health by household wealth index, as defined in the DHS data. Each column in Table 4 presents separate regression results for children from poor and non-poor households. The results suggest that low-income households are negatively affected by rain shocks. The point estimate is -0.197 and significant at the 5% significance level. A high level of SOC does not reduce the negative effect of the rainfall shock on poor households. In addition, the cumulative rainfall shock and a high level of SOC have no impact on children's HAZ.

5.2.4 Heterogeneity by land ownership

Agriculture is the main occupation in rural India. To see if the results are determined by farm households, I examine the heterogeneity by land ownership: has farmland and has no farmland.

Table 4 presents the results for households that own and do not own farmland. The results suggest that rain shocks negatively affect households that own land, suggesting they are rain-dependent. A high level of SOC does not reduce the negative impact of rainfall shock on households that own land. Moreover, the cumulative rainfall shock and a high level of SOC have no impact on children's HAZ. In addition, the p-value difference test suggests that those who have agricultural land do not differ statistically from those who do not.

5.3 Extended results

5.3.1 Impact of SOC on childhood stunting and wasting

I estimate the logistic regression model to predict whether a switch from low to high SOC reduces the probability of stunting or wasting in children in response to rain shocks. Table 5 presents the effect of rain shocks on the probability of stunting and wasting in children and estimates the moderating effects of SOC. The dependent variable is binary for stunted children whose height-for-age is less than -2 ($HAZ < -2$) is 1; 0 otherwise. Similarly, the binary for childhood wasting cases where the weight for height is less than -2 ($WHZ < -2$) is 1; 0 otherwise. Results suggest that children exposed to cumulative rain shocks (in-utero to 4 years of age) are more likely to be wasted and less likely to be stunted.

Columns 1 and 2 in Table 5 show the results for the logit regression. Odd ratios of coefficients are provided. The odds of child wasting increased by 31% (1.309-1 = 0.309) in periods of low rainfall. Whereas, the odds of stunting in children is reduced by 13% (0.870-1 = -0.13) during periods of low rainfall. The odds of wasting for children living in high-level SOC areas is 5% (1.053-1=0.05) higher than in low-level SOC areas. That is a surprising result. To check for sensitivity to SOC threshold. I run the logit regression for different SOC thresholds. Appendix Table A11 presents the results for the SOC threshold set at 25th, 50th, and 75th. We observe the mitigating effect of SOC during periods of low rainfall on childhood wasting at a threshold just above the 25th percentile and just above the 75th percentile. This means that a modest change in soil health may also improve children's health. The results suggest sensitivity to the SOC level. Therefore, my results must be interpreted cautiously.

The average marginal effects of the shock fraction at a low SOC level is 0.041 and significant at 1% significance level for children with wasting. Whereas, the average marginal effect of the shock fraction at a high SOC level is statistically not different from zero. Average marginal effects suggest that the probability of wasting in children living in low SOC areas increases by 0.04 percentage points during periods of low rainfall. The switch to high SOC levels attenuates this negative effect of cumulative rain shocks. The left and right graphs in Figure 6 show the average marginal effects on the probability of stunting and wasting in children.

Moreover, the results suggest that during periods of low rainfall, children are less likely to be stunted by 0.02 percentage points in regions with high or low SOC. This is reflected in the average negative marginal effects of stunting. It is noteworthy that cases of child stunting are chronic and difficult to explain simply by the agricultural process, including precipitation and soil quality. Furthermore, the results do not suggest any impact of a high level of SOC on stunting and wasting of children in the matched sample.

6 Conclusion

6.1 Summary

This article examines the relation between SOC and the impact of precipitation on children's health. The results demonstrate that a high level of SOC reduces the negative impact of rain shock on children's health in rural areas. Specifically, SOC affects the child's WHZ but has no effect on the HAZ. I find that SOC has an significant moderating effect on girls, but not on boys. Moreover, a high level of SOC ensures resilience in semi-arid and humid tropical climatic zones.

I find that the high level of SOC makes children resilient to wasting during periods of low rainfall. I also find that the shock fraction reduces the likelihood of child stunting, a long-term measure of health, and requires additional research. Results suggest a sensitivity to SOC threshold levels. Note that the regression results for the matched sample are sensitive to the variables used in the matching algorithm and the SOC threshold as treatment. Therefore, my results need to be interpreted cautiously.

6.2 Limitation

One limitation is that the soil organic carbon content variable used in this analysis is time invariant. Existing research shows that agricultural practices that cause pollution, such as stubble burning (Singh et al, 2019) and fertilizer use (Brainerd and Menon, 2014), can have negative impacts on children's health. Such agricultural practices may also have an impact on the concentrations of organic carbon in the soil. Therefore, estimates may be subject to upward bias because of the omitted variable. Due to a lack of data, I am unable to control for these practices. Nevertheless, it is important to explore these pathways in future research efforts.

6.3 Conclusions

The findings from this paper motivates the policies such as adoption of the best farming management practices, the use of soil health card to improve soil organic carbon concentration, and thus enhance food security. Additionally, given the significant moderating effect of high levels of soil organic carbon, decision makers should target areas with low levels of soil organic carbon. Child development programmes in India need to be strengthened to include soil quality and the impact of climate change on children's health. Food nutrients and soil conditions are interlinked through agriculture, and better soil quality helps reduce malnutrition.

6.4 Future work

Breastfeeding may provide nutrition to children as a response to food insecurity and other environmental shocks. By linking soil quality to breastfeeding practices, we could better understand how children's nutrition responds to shocks.

Weather fluctuations impact the time- and gender-dependent nature of agricultural activities (Mahajan, 2017; Afridi et al, 2022). For example, (Afridi et al, 2022) demonstrated that workdays in farm for Indian women were considerably shorter than those of men during a drought. Additionally, Mahajan (2017) show that rain-induced agricultural shocks affect women's wages differently than men's. Women's employment opportunities are closely related to children's nutrition (Debela et al, 2021). The SOC mitigation effect of the rain shock could increase women's resilience to employment opportunities in rural areas. That may be the subject of further research.

References

Afridi F, Mahajan K, Sangwan N (2022) The gendered effects of droughts: Production shocks and labor response in agriculture. Labour Economics 78:102,227. https://doi.org/https://doi.org/10.1016/j.labeco.

2022.102227, URL https://www.sciencedirect.com/science/article/pii/S0927537122001178

Almond D, Currie J (2011) Killing me softly: The fetal origins hypothesis. Journal of economic perspectives 25(3):153–72

Auffhammer M, Ramanathan V, Vincent JR (2012) Climate change, the monsoon, and rice yield in India. Climatic change 111(2):411–424

Bakhtsiyarava M, Grace K (2021) Agricultural production diversity and child nutrition in Ethiopia. Food Security 13(6):1407–1422

Berkhout ED, Malan M, Kram T (2019) Better soils for healthier lives? an econometric assessment of the link between soil nutrients and malnutrition in sub-saharan africa. PloS one 14(1):e0210,642

Bhargava AK, Vagen T, Gassner A (2018) Breaking ground: unearthing the potential of high-resolution, remote-sensing soil data in understanding agricultural profits and technology use in Sub-Saharan Africa. World Development 105:352–366

Blackwell M, Iacus S, King G, et al (2009) cem: Coarsened exact matching in stata. The Stata Journal 9(4):524–546

Brainerd E, Menon N (2014) Seasonal effects of water quality: The hidden costs of the green revolution to infant and child health in India. Journal of Development Economics 107:49–64

Center for International Earth Science Information Network - CIESIN - Columbia University (2018) Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11. URL https://doi.org/10.7927/H4JW8BX5, place: Palisades, New York

Debela BL, Gehrke E, Qaim M (2021) Links between maternal employment and child nutrition in rural Tanzania. American Journal of Agricultural Economics 103(3):812–830

Dimitrova A, Bora JK (2020) Monsoon weather and early childhood health in India. PLOS ONE 15(4):e0231,479. https://doi.org/10.1371/journal.pone.0231479, URL https://dx.plos.org/10.1371/journal.pone.0231479

Dimitrova A, Muttarak R (2020) After the floods: Differential impacts of rainfall anomalies on child stunting in India. Global Environmental Change 64:102,130. https://doi.org/10.1016/j.gloenvcha.2020.102130, URL https://linkinghub.elsevier.com/retrieve/pii/S0959378020307135

Dinkelman T (2017) Long-run health repercussions of drought shocks: evidence from south african homelands. The Economic Journal 127(604):1906–1939

Feeny S, Mishra A, Trinh TA, et al (2021) Early-Life exposure to rainfall shocks and gender gaps in employment: Findings from Vietnam. Journal of Economic Behavior & Organization 183:533–554. https://doi.org/10.1016/j.jebo.2021.01.016, URL https://linkinghub.elsevier.com/retrieve/pii/S0167268121000275

Funk CC, Peterson PJ, Landsfeld MF, et al (2014) A quasi-global precipitation time series for drought monitoring. US Geological Survey data series 832(4):1–12

Grace K, Davenport F, Funk C, et al (2012) Child malnutrition and climate in Sub-Saharan Africa: An analysis of recent trends in Kenya. Applied Geography 35(1-2):405–413. https://doi.org/10.1016/j.apgeog.2012.06.017, URL https://linkinghub.elsevier.com/retrieve/pii/S0143622812000768

Guiteras R (2009) The impact of climate change on Indian agriculture. Manuscript, Department of Economics, University of Maryland, College Park, Maryland

Hengl T (2018a) Clay content in % (kg / kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. https://doi.org/10.5281/zenodo.2525663, URL https://doi.org/10.5281/zenodo.2525663

Hengl T (2018b) Sand content in % (kg / kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. https://doi.org/10.5281/zenodo.2525662, URL https://doi.org/10.5281/zenodo.2525662

Hengl T (2018c) Silt content in % (kg / kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. https://doi.org/10.5281/zenodo.2525676, URL https://doi.org/10.5281/zenodo.2525676

Hengl T, Wheeler I (2018) Soil organic carbon content in x 5 g / kg at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. https://doi.org/10.5281/zenodo.2525553, URL https://doi.org/10.5281/zenodo.2525553

Huang J, Hartemink AE, Kucharik CJ (2021) Soil-dependent responses of US crop yields to climate variability and depth to groundwater. Agricultural Systems 190:103,085. https://doi.org/10.1016/j.agsy.2021.103085, URL https://linkinghub.elsevier.com/retrieve/pii/S0308521X2100038X

Iacus SM, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. Political analysis 20(1):1–24

Kane DA, Bradford MA, Fuller E, et al (2021) Soil organic matter protects US maize yields and lowers crop insurance payouts under drought. Environmental Research Letters 16(4):044,018

Kim K, Bevis L (2019) Soil fertility and poverty in developing countries. Choices 34(2):1–8

Kim R, Bijral AS, Xu Y, et al (2021) Precision mapping child undernutrition for nearly 600,000 inhabited census villages in India. Proceedings of the National Academy of Sciences 118(18):e2025865,118. https://doi.org/10.1073/pnas.2025865118, URL http://www.pnas.org/lookup/doi/10.1073/pnas.2025865118

Lal R (2004) Soil carbon sequestration impacts on global climate change and food security. science 304(5677):1623–1627

Lal R (2006) Enhancing crop yields in the developing countries through restoration of the soil organic carbon pool in agricultural lands. Land degradation & development 17(2):197–209

Mahajan K (2017) Rainfall shocks and the gender wage gap: Evidence from Indian agriculture. World Development 91:156–172

McMahon K, Gray C (2021) Climate change, social vulnerability and child nutrition in South Asia. Global Environmental Change 71:102,414

Rani SI, Arulalan T, George JP, et al (2021) Imdaa: High-resolution satellite-era reanalysis for the indian monsoon region. Journal of Climate 34(12):5109–5133

Shah M, Steinberg BM (2017) Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. Journal of Political Economy 125(2):527–561

Singh P, Dey S, Chowdhury S, et al (2019) Early life exposure to outdoor air pollution: effect on child health in India. Brookings India URL http://hdl.handle.net/11540/9783.

Victora CG (1992) The association between wasting and stunting: an international perspective. The Journal of nutrition 122(5):1105–1110

Viswanathan B, Kumar KK (2015) Weather, agriculture and rural migration: evidence from state and district level migration in India. Environment and Development Economics 20(4):469–492

Webb P, Block S (2012) Support for agriculture during economic transformation: Impacts on poverty and undernutrition. Proceedings of the National Academy of Sciences 109(31):12,309–12,314. https://doi.org/10.1073/pnas.0913334108, URL http://www.pnas.org/cgi/doi/10.1073/pnas.0913334108

Wiesmann D (2006) A global hunger index measurement concept, ranking of countries, and trends. FCND discussion paper ; 212, International food policy research institute (IFPRI). Food consumption and nutrition division (FCND), Washington, DC

Zaveri E, B Lobell D (2019) The role of irrigation in changing wheat yields and heat sensitivity in India. Nature communications 10(1):1–7

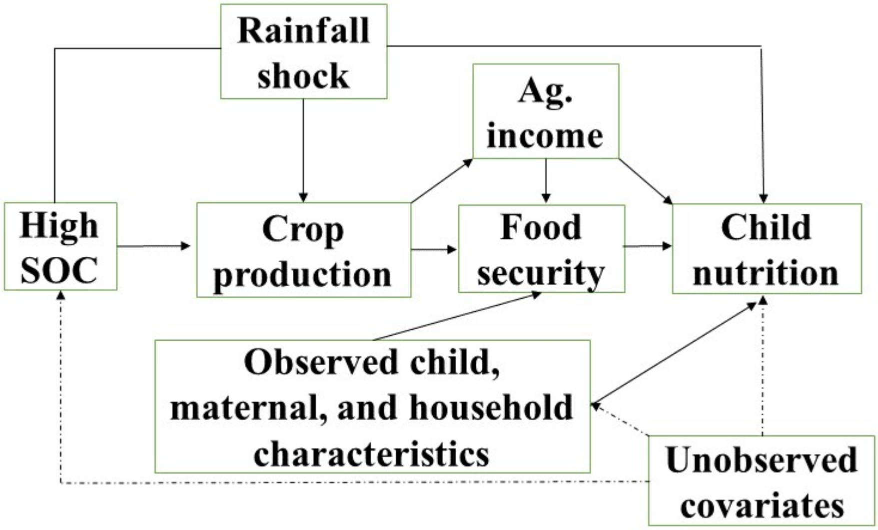


Fig. 1: A simple conceptual relationship between soil and children's health.

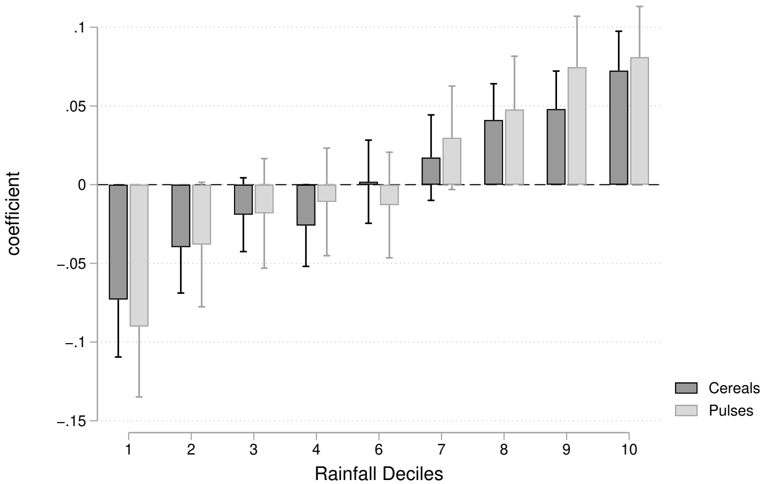


Fig. 2: Coefficient for rainfall deciles and 95% CI in India. The dependent variable is the natural logarithm of annual crop yield (kg per hectare) from 2001 to 2015. The specification includes district and year fixed effects. The 5th decile is selected as reference.

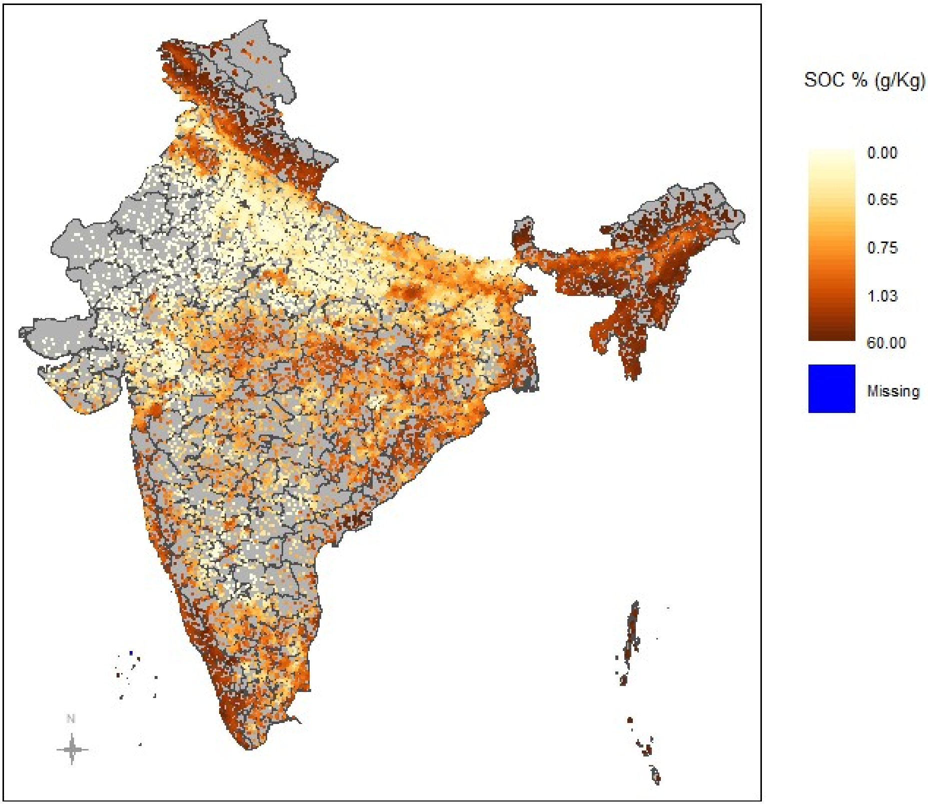
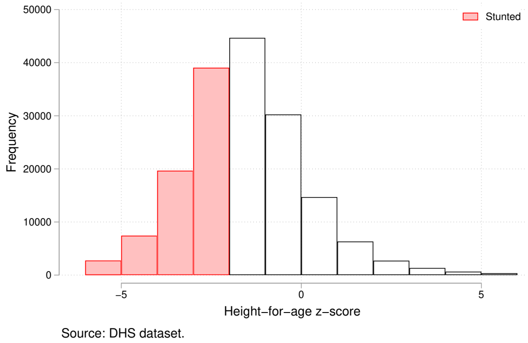
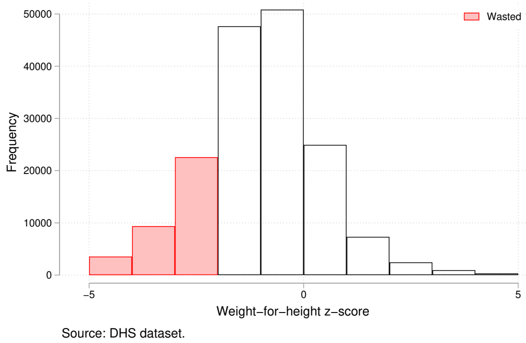


Fig. 3: The dots represent the average soil organic carbon content of the DHS rural clusters in India. The missing in the map indicates the null values for union territory Lakshadweep. The dark lines in the background are the district borders.

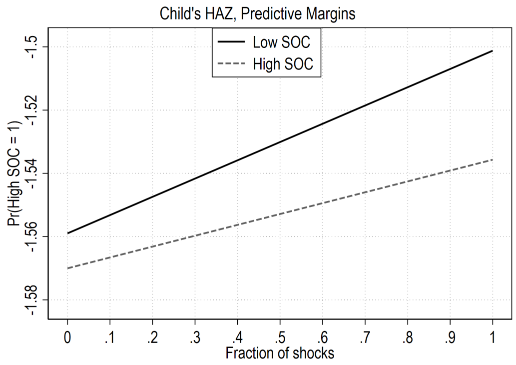


(a) Stunted (HAZ< −2)

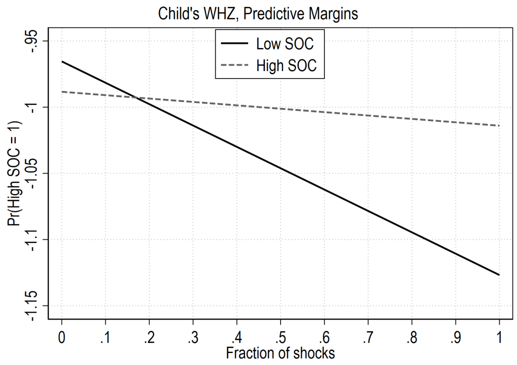


(b) Wasted (WHZ< −2)

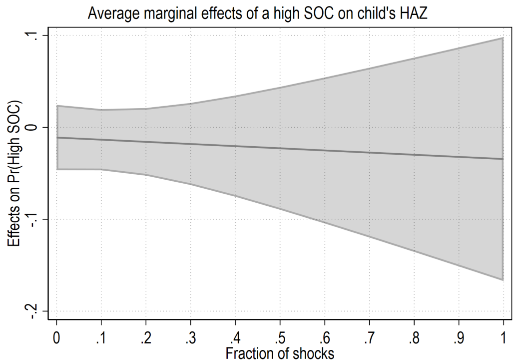
Fig. 4: Distribution of childhood health outcomes. Source: Own calculations based on DHS dataset (2015-2016).



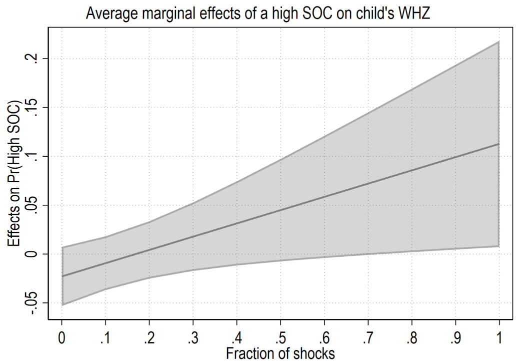
(a) Height-for-age z-scores



(b) Weight-for-height z-scores

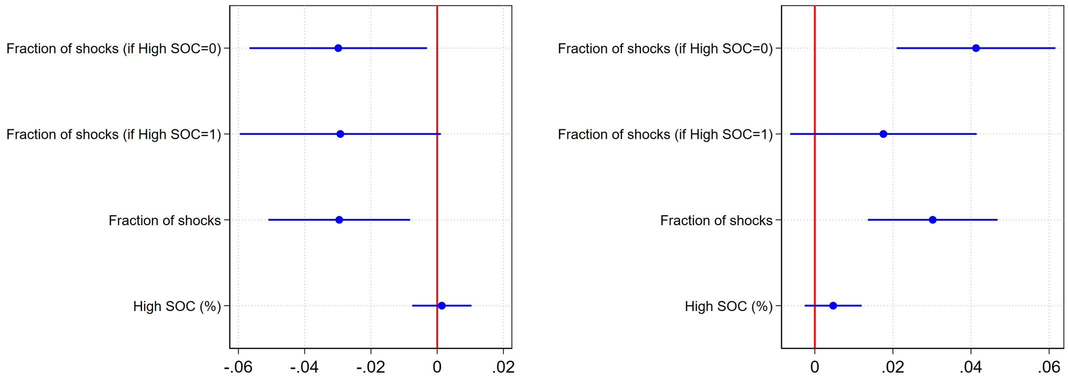


(c) Height-for-age z-scores



(d) Weight-for-height z-scores

Fig. 5: Average marginal effects of high SOC levels on anthropometric measurements in childhood. A full sample is used to create the plots.



(a) Stunting (b) Wasting

Fig. 6: Average marginal effects on the probability of childhood stunting and wasting. The x-axis is the probability level. The high SOC level is fixed above the 50th percentile.

Table 1: Summary statistics.

	Observation	Mean	Std. Dev.
Child health measures			
Height-for-age z score	169,904	-1.558	1.681
Weight-for-age z score	169,904	-1.588	1.208
Weight-for-height z score	169,904	-0.991	1.381
Child health outcomes, yes=1			
Stunted (HAZ< -2)	169,904	0.405	0.491
Wasted (WHZ< -2)	169,904	0.209	0.406
Rainfall below 20th percentile, yes=1			
Rainfall shock - in-utero	169,904	0.110	0.313
Rainfall shock - birth year	169,904	0.110	0.312
Rainfall shock - 1st year	137,807	0.125	0.331
Rainfall shock - 2nd year	103,642	0.148	0.355
Rainfall shock - 3rd year	69,621	0.168	0.374
Rainfall shock - 4th year	33,951	0.167	0.373
Fraction of shocks	169,904	0.134	0.182
Soil health measure			
Soil organic carbon (SOC) %(g/Kg)	169,897	0.945	0.675
25th percentile level of SOC	169,904	0.633	
50th percentile level of SOC	169,904	0.733	
75th percentile level of SOC	169,904	0.965	

Note: The rain shock for the 1st to the 4th year have different observations to adjust the age of the child. The sample is composed of 33,951 4-year-olds, 69,621 3-year-olds, 103,642 2-year-olds, 137,807 1-year-olds and 169,904 in-utero. Source: DHS and CHIRPS data.

Appendix A Additional Figures and Tables

Table 2: Impact of high levels of SOC on the health of children.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.058	-0.161***	0.011	-0.063
	(0.050)	(0.042)	(0.063)	(0.053)
High SOC (%)	-0.011	-0.023	-0.008	-0.023
	(0.018)	(0.015)	(0.021)	(0.018)
High SOC $\times$ Fraction of shocks	-0.023	0.136**	-0.071	0.057
	(0.072)	(0.059)	(0.089)	(0.071)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.573	-1.059
SD dependent var.	1.681	1.381	1.667	1.366
Observations	169,904	169,904	102,296	102,296
R-square	0.148	0.090	0.144	0.079

Levels of significance: $p < 0.01^{***}$, $p < 0.05^{**}$, $p < 0.1^{*}$. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 50th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time.

Table 3: Heterogeneity by selected climate zones

	HAZ			
	Tropical wet	Tropical wet and dry	Semi arid	Humid sub-tropical
Fraction of shocks	-1.204	-0.167	0.130	0.062
	(1.661)	(0.167)	(0.130)	(0.061)
High SOC (%)	0.451	-0.009	0.023	-0.031
	(0.284)	(0.029)	(0.051)	(0.027)
High SOC × Fraction of shocks	1.084	0.175	-0.767**	0.052
	(1.661)	(0.183)	(0.322)	(0.094)
Mean dependent var.	-1.258	-1.538	-1.516	-1.647
Observations	7036	40,607	25,517	86,254
R-square	0.146	0.130	0.144	0.160

	WHZ			
	Tropical wet	Tropical wet and dry	Semi arid	Humid sub-tropical
Fraction of shocks	-0.325	0.292**	-0.280**	-0.133***
	(0.813)	(0.139)	(0.111)	(0.051)
High SOC (%)	0.314*	0.016	-0.054	-0.021
	(0.175)	(0.026)	(0.041)	(0.023)
High SOC × Fraction of shocks	0.200	-0.416***	0.547**	0.180**
	(0.819)	(0.153)	(0.275)	(0.075)
Mean dependent var.	-0.861	-1.197	-1.025	-0.934
Observations	7036	40,607	25,517	86,254
R-square	0.079	0.075	0.072	0.093

Levels of significance: $p < 0.01$***, $p < 0.05$**, $p < 0.1$*. Robust standard errors in parentheses are clustered at the DHS cluster level. The high SOC level is fixed above the 50th percentile. Each regression includes district and month-birth year specific fixed effects. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. Arid and Mountain are limited by very small sample to provide meaningful estimates and hence excluded.

Table 4: Heterogeneities on the full sample.

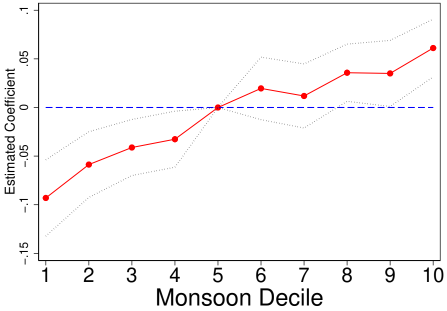
	Boys		Girls	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.011	-0.112**	0.108	-0.205***
	(0.065)	(0.057)	(0.067)	(0.055)
High SOC (%)	-0.023	-0.005	0.003	-0.041**
	(0.022)	(0.020)	(0.024)	(0.020)
High SOC × Fraction of shocks	0.022	0.110	-0.068	0.152**
	(0.093)	(0.079)	(0.094)	(0.077)
Mean dependent. var.	-1.597	-1.017	-1.516	-0.963
Observations	87,643	87,643	82,259	82,259
R-square	0.142	0.096	0.165	0.093
	Poor		**Non-poor**	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.060	-0.197***	0.056	-0.110*
	(0.069)	(0.056)	(0.069)	(0.060)
High SOC (%)	-0.000	-0.012	0.019	-0.029
	(0.026)	(0.022)	(0.024)	(0.020)
High SOC × Fraction of shocks	0.002	0.114	-0.038	0.133*
	(0.104)	(0.081)	(0.092)	(0.080)
Mean dependent. var.	-1.847	-1.135	-1.321	-0.873
Observations	76,633	76,633	93,259	93,259
R-square	0.128	0.088	0.137	0.090
	Has ag. land		**Has no ag. land**	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.083	-0.190***	0.033	-0.110*
	(0.063)	(0.054)	(0.075)	(0.062)
High SOC (%)	-0.015	-0.025	-0.006	-0.012
	(0.023)	(0.020)	(0.026)	(0.021)
High SOC × Fraction of shocks	-0.112	0.119	0.089	0.132
	(0.090)	(0.076)	(0.104)	(0.083)
Mean dependent. var.	-1.511	-0.976	-1.617	-1.009
Observations	94,065	94,065	75,838	75,838
R-square	0.152	0.100	0.153	0.089

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 50th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time. See Appendix Table A10 for heterogeneities on the matched sample.

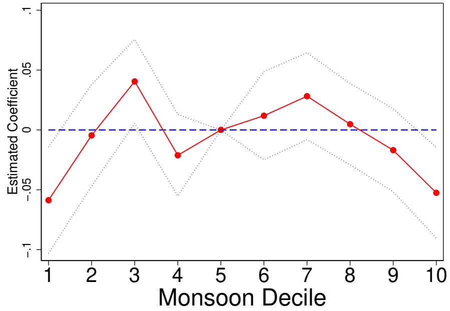
Table 5: Impact of high SOC on the likelihood of childhood stunting and wasting: Logit estimates.

	Full		Matched	
	Stunted	Wasted	Stunted	Wasted
Fraction of shocks	0.870**	1.309***	0.875*	1.091
	(0.056)	(0.088)	(0.070)	(0.094)
High SOC (%)	1.006	1.053**	1.002	1.038
	(0.023)	(0.027)	(0.027)	(0.031)
High SOC × Fraction of shocks	1.003	0.855	1.101	1.061
	(0.091)	(0.082)	(0.121)	(0.123)
AME of the shock fraction at a high SOC=1	-0.029*	0.017	-0.008	0.024
	(0.015)	(0.012)	(0.019)	(0.015)
AME of the shock fraction at a high SOC=0	-0.030**	0.041***	-0.029*	0.014
	(0.014)	(0.010)	(0.017)	(0.014)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent var.	0.405	0.209	0.407	0.221
SD dependent var.	0.491	0.406	0.491	0.415
Observations	169,898	169,879	102,289	102,147

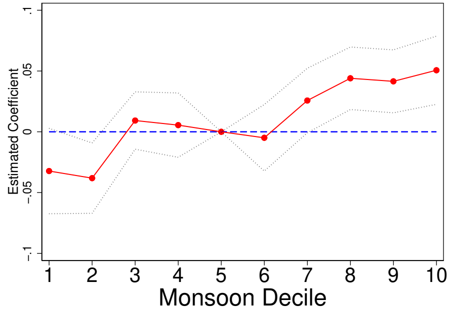
Levels of significance: $p < 0.01$***, $p < 0.05$**, $p < 0.1$*. Odd ratios are reported. Robust standard errors in parentheses are clustered at the DHS cluster level. The high SOC level is fixed above the 50th percentile. AME refers to average marginal effects. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls.



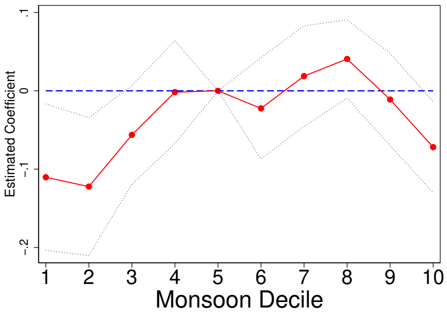
(a) Rice



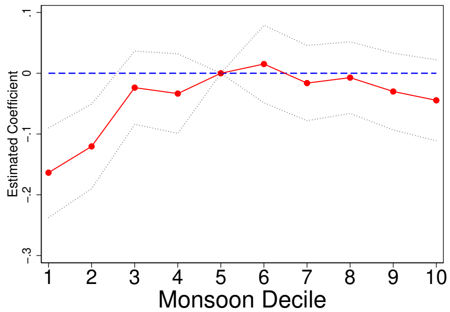
(b) Maize



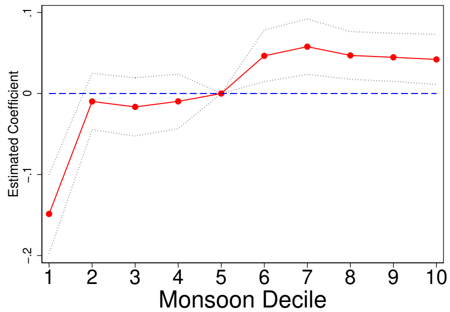
(c) Wheat



(d) Soybean



(e) Cotton



(f) Groundnut

Fig. A1: Effects of monsoon rainfall on crop yields

Notes: The dependent variable is the natural logarithm of annual crop yield (kg per hectare) from 2001 to 2015. The specification includes district and year fixed effects. The figure plots the point estimate are plotted with 95% confidence intervals. The 5th decile is selected as reference. The monsoon rainfall deciles were constructed using monthly Climate Hazards Group InfraRed Precipitation (CHIRPS) data in a growing season (June through September) from year 1982 to 2015 (Funk et al, 2014).

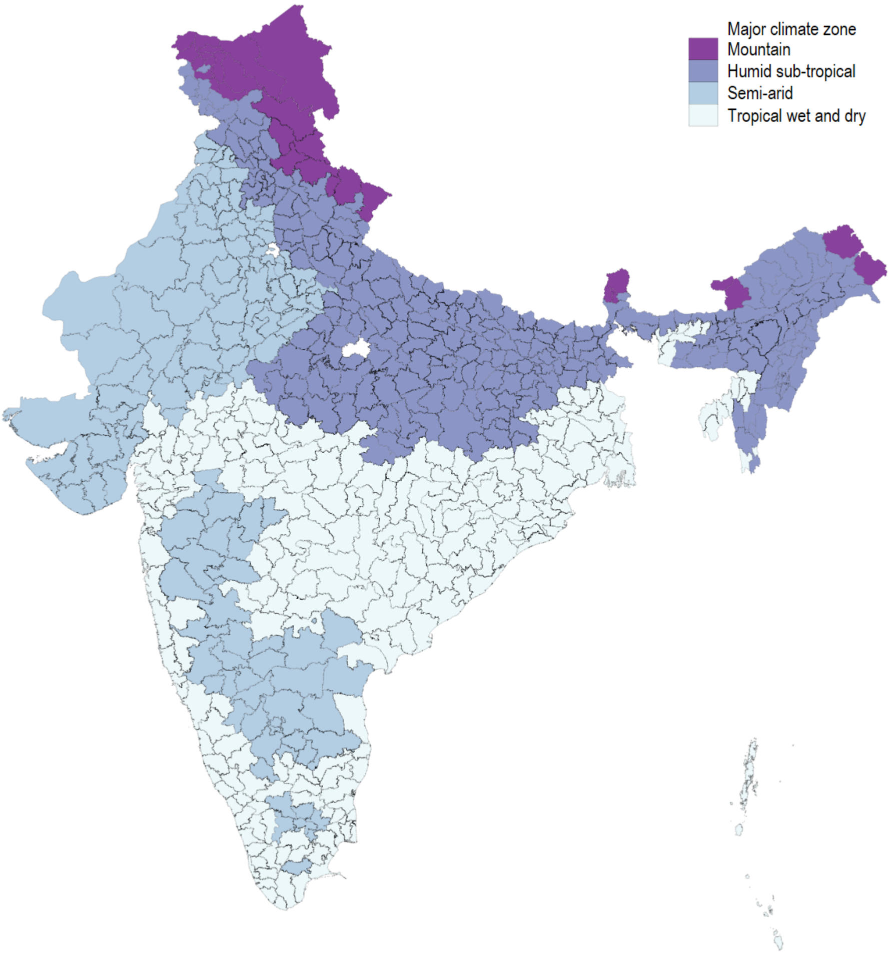


Fig. A2: Major climate zones in India based on Köppen Geiger climate classification. The arid zone is in the western part of India depicted in the semi-arid zone.Tropical wet is in the western coastal portion of India represented in the tropical wet and dry zone. Source: Own calculation.

Table A1: Description for variables included in the study

Variable	Description
Child-specific	
bord	Order of birth
malechild	Dummy for male child
childsizelarge	Dummy for child was large at birth
childsizeavg	Dummy for child was average at birth
numfemalesib	Number of female siblings
nummalesib	Number of male siblings
$age510$	Dummy for child with sibling between the age 5 and 10 years
$age1115$	Dummy for child with sibling between the age 11 and 15 years
$age16$	Dummy for child with sibling above 16 years
$hw1$	Child's age in months
Woman-specific	
$v012$	Woman's age in years
womanpriedu	Dummy for woman has primary education
womansecedu	Dummy for woman has secondary or higher level education
womanbmi	Woman's body mass index
womaneatfruits	Dummy for woman consumes fruits daily or weekly
womaneatveges	Dummy for woman consumes vegetables daily or weekly
womaneateggs	Dummy for woman consumes eggs daily or weekly
womaneatmeat	Dummy for woman consumes chicken/meat/fish daily or weekly
womansmoke	Dummy for woman smokes
womandrink	Dummy for woman drinks alcohol
womanprenataldoc	Dummy for had prenatal care with doctor
Household-specific	
$v104$	Years lived in place of residence
$hv220$	Age of household head in years
hhheadmale	Dummy for male household head
hhhindu	Dummy for household religion is Hinduism
hhmuslim	Dummy for household religion is Islam
hhscst	Dummy for household belongs to SC/ST
hhradio	Dummy for household owns a radio/transistor
hhtv	Dummy for household owns a television
hhrefri	Dummy for household owns a refrigerator
hhmotorcycle	Dummy for household owns a motorcycle
hhcar	Dummy for household owns a car
hhelec	Dummy for household has electricity
$hv244$	Dummy for household owns agricultural land
hhirragland	Dummy for household irrigate agricultural land
$sh52a$	Dummy for household owns cows/bulls/buffaloes
$sh52b$	Dummy for household owns camels
$sh52c$	Dummy for household owns horses/donkeys/mules
$sh52d$	Dummy for household owns goats
$sh52e$	Dummy for household owns sheep
$sh52f$	Dummy for household owns chickens/ducks
hhpipewater	Dummy for source of drinking water: piped water
hhgroundwater	Dummy for source of drinking water: ground water
hhsurfacewater	Dummy for source of drinking water: surface water
hhrainwater	Dummy for source of drinking water: rain water, tanker water, etc
hhflushtoilet	Dummy for toilet facility: flush toilet
hhpit	Dummy for toilet facility: pit toilet/latrine
hhnofacility	Dummy for toilet facility: no facility/bush/field
hhpoorest	Dummy for household wealth index: poorest
hhpoorer	Dummy for household wealth index: poorer
hhmiddle	Dummy for household wealth index: middle
hhricher	Dummy for household wealth index: richer

Notes: For the analysis, $hw1$ was transformed with restricted cubic spline and knots are placed at the interval of 6, 12, 18, 24, 36, and 48.

Table A2: Description for variables included in the study

Variable	Description
Weather-specific	
childrain	June-September daily accumulation of rainfall over child's life time
childgdd	Growing degree days over child's life time
childhdd	Harmful degree days over child's life time

Notes: For the analysis, *childrain* and *childgdd* was transformed by squaring the variable; *childhdd* was transformed by taking a square root of the variable.

Table A3: Summary statistics (N = 169,904)

	Mean	Std. Dev.
Child birth order number	2.343	1.521
Male child	0.516	0.500
Child with greater than average size at birth	0.165	0.371
Child with average size at birth	0.691	0.462
Number of female siblings	0.828	1.050
Number of male siblings	0.662	0.852
Number of child with sibling between the age 5 and 10 years	0.691	0.878
Number of child with sibling between the age 11 and 15 years	0.176	0.506
Number of child with sibling above 16 years	0.062	0.358
Child's age in months	29.895	17.034
Woman's age in years	27.079	5.178
Woman has primary edu	0.156	0.363
Woman has secondary or higher edu	0.494	0.500
Woman's body mass index	20.775	3.465
Woman consumes fruits daily or weekly	0.333	0.471
Woman consumes vegetables daily or weekly	0.945	0.227
Woman consumes eggs daily or weekly	0.340	0.474
Woman consumes chicken/meat/fish daily or weekly	0.356	0.479
Woman smokes	0.007	0.084
Woman drinks alcohol	0.024	0.153
Access to prenatal care with doctor	0.361	0.480
Years lived in place of residence	15.460	25.387
Age of household head	44.360	15.216
Male household head	0.879	0.326
Household religion is Hinduism	0.744	0.436
Household religion is Islam	0.137	0.344
Household belongs to SC/ST	0.420	0.494
Household owns a radio/transistor	0.086	0.280
Household owns a television	0.495	0.500
Household owns a refrigerator	0.165	0.371
Household owns a motorcycle	0.311	0.463
Household owns a car	0.042	0.200
Household has electricity	0.814	0.389
Household owns ag. land	0.554	0.497
Irrigated ag land only	0.278	0.448
Household owns cows/bulls/buffaloes	0.523	0.499
Household owns camels	0.004	0.064
Household owns horses/donkeys/mules	0.007	0.086
Household owns goats	0.225	0.417
Household owns sheep	0.022	0.148
Household owns chickens/ducks	0.220	0.414
Source of drinking water: piped water	0.295	0.456
Source of drinking water: ground water	0.626	0.484
Source of drinking water: surface water	0.054	0.226
Toilet facility: flush toilet	0.337	0.473
Toilet facility: pit toilet/latrine	0.105	0.306
Toilet facility: no facility/bush/field	0.541	0.498
Wealth index: poorest	0.232	0.422
Wealth index: poorer	0.219	0.414
Wealth index: middle	0.200	0.400
Wealth index: richer	0.180	0.384

Source: DHS and CHIRPS data.

Table A4: Alternative main regression results using population-weighted rain measures

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.027	-0.143***	-0.016	-0.046
	(0.050)	(0.042)	(0.062)	(0.052)
High SOC (%)	-0.011	-0.020	-0.008	-0.021
	(0.018)	(0.015)	(0.021)	(0.018)
High SOC × Fraction of shocks	-0.017	0.102*	-0.073	0.038
	(0.072)	(0.058)	(0.089)	(0.071)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.573	-1.059
SD dependent var.	1.681	1.381	1.667	1.366
Observations	169,904	169,904	102,296	102,296
R-square	0.148	0.090	0.144	0.079

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 50th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time.

Table A5: Means, standard deviation and Pearson correlation matrix for soil attributes (N = 169,897)

	Means	SD	WHZ	SOC	Clay	Sand	Silt	EVI	Slope
WHZ	-0.99	1.38	1.00						
SOC	0.94	0.67	0.12^a	1.00					
Clay	32.44	5.33	-0.09^a	-0.08^a	1.00				
Sand	38.18	5.58	0.02^a	0.02^a	-0.57^a	1.00			
Silt	29.39	5.08	0.07^a	0.06^a	-0.43^a	-0.50^a	1.00		
EVI	2927.33	702.22	0.10^a	0.38^a	0.02^a	-0.15^a	0.14^a	1.00	
Slope	0.29	111.22	0.00	-0.25^a	0.00	0.00	0.00	0.21^a	1.00

Note: $^a p < .01$. EVI: Enhanced Vegetation Index for 2015.

Table A6: Robustness check: confounding variables included as controls

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.056	-0.166***	0.070	-0.152***
	(0.050)	(0.042)	(0.063)	(0.053)
High SOC (%)	-0.011	-0.023	-0.005	-0.016
	(0.018)	(0.016)	(0.021)	(0.018)
High SOC × Fraction of shocks	-0.022	0.135**	-0.089	0.098
	(0.072)	(0.059)	(0.090)	(0.072)
DHS controls	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.572	-1.061
SD dependent var.	1.681	1.381	1.667	1.369
Observations	169,897	169,897	102,296	102,296
R-square	0.148	0.090	0.142	0.080

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 50th percentile. DHS controls include child, mother, and household level characteristics. Other controls include confounding variables such as soil texture, slope, and vegetation.

Table A7: Sensitivity test for various thresholds: High soil organic carbon content above 25 percentile.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.153**	-0.233***	0.155**	-0.214***
	(0.063)	(0.055)	(0.067)	(0.059)
High SOC (%)	0.012	-0.026	0.016	-0.017
	(0.022)	(0.018)	(0.027)	(0.022)
High SOC × Fraction of shocks	-0.147**	0.186***	-0.233***	0.151**
	(0.072)	(0.061)	(0.086)	(0.076)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.573	-1.059
SD dependent var.	1.681	1.381	1.667	1.366
Observations	169,904	169,904	80,253	80,253
R-square	0.148	0.090	0.145	0.094

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 25th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time. The match summary consists of: the number of balanced matched observations is 40129 for treatment and control; and the unmatched observation is 2354 out of 42483 for control and 87292 out of 127421 for treatment.

Table A8: Sensitivity test for various thresholds: High soil organic carbon content above 75 percentile.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.048	-0.114***	0.122	-0.091
	(0.042)	(0.037)	(0.093)	(0.080)
High SOC (%)	-0.015	-0.022	-0.022	-0.000
	(0.028)	(0.023)	(0.034)	(0.029)
High SOC × Fraction of shocks	-0.003	0.066	-0.122	-0.020
	(0.085)	(0.072)	(0.124)	(0.106)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.573	-1.059
SD dependent var.	1.681	1.381	1.667	1.366
Observations	169,904	169,904	45,498	45,498
R-square	0.148	0.090	0.145	0.094

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 75th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time. The match summary consists of: the number of balanced matched observations is 22749 for treatment and control; and the unmatched observation is 104676 out of 127425 for control and 19730 out of 42479 for treatment.

Table A9: Sensitivity test for different DHS cluster level: 20 km

	(1) Full	(2) Full	(3) Full	(4) Matched
Fraction of shocks	-0.241***	-0.242***	-0.260***	-0.229***
	(0.054)	(0.054)	(0.055)	(0.060)
High SOC (%)	-0.017	-0.023	-0.023	-0.008
	(0.018)	(0.018)	(0.018)	(0.022)
High SOC × Fraction of shocks	0.129**	0.154**	0.163***	0.109
	(0.061)	(0.061)	(0.061)	(0.076)
Marginal effects	-0.144***	-0.127***	-0.137***	-0.174***
	(0.034)	(0.033)	(0.034)	(0.047)
Mean dependent variable		-0.991		-1.075
Average years of exposure		0.133		0.150
DHS controls	No	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes
Observations	169,904	169,904	169,904	80,254
Adjusted R^2	0.067	0.086	0.086	0.068

Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. The high SOC level is fixed above the 25th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time. All regressions include district and month-birth year specific fixed effects. The matching summary includes: 40,129 matched out of 42,483 observations for control and 40,129 matched out of 127,421 for treated.

Table A10: Including dry and rainy seasons as an additional variable in the matching algorithm.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.058	-0.161***	0.030	-0.102*
	(0.050)	(0.042)	(0.063)	(0.053)
High SOC (%)	-0.011	-0.023	-0.011	-0.016
	(0.018)	(0.015)	(0.021)	(0.018)
High SOC × Fraction of shocks	-0.023	0.136**	-0.072	0.036
	(0.072)	(0.059)	(0.091)	(0.072)
DHS controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Mean dependent. var.	-1.558	-0.991	-1.580	-1.065
SD dependent var.	1.681	1.381	1.665	1.366
Observations	169,904	169,904	97,441	97,441
R-square	0.148	0.090	0.147	0.080

Levels of significance: $p < 0.01^{***}$, $p < 0.05^{**}$, $p < 0.1^{*}$. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. The high SOC level is fixed above the 50th percentile. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time. The match summary consists of: the number of balanced matched observations is 48721 for treatment and control; and the unmatched observation is 36229 out of 84950 for control and 36233 out of 84954 for treatment.

Table A11: Sensitivity check for different SOC thresholds

	Child wasting at various SOC thresholds.		
	25th	50th	75th
Fraction of shocks	1.482***	1.309***	1.290***
	(0.131)	(0.088)	(0.076)
High SOC (%)	1.008	1.053**	1.065
	(0.032)	(0.027)	(0.041)
High SOC × Fraction of shocks	0.764***	0.855	0.750**
	(0.075)	(0.082)	(0.093)
DHS controls	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Mean dependent. var.	0.209	0.209	0.209
SD dependent var.	0.406	0.406	0.406
Observations	169,879	169,879	169,879

Levels of significance: $p < 0.01^{***}$, $p < 0.05^{**}$, $p < 0.1^{*}$. Odds ratios are reported. Robust standard errors in parentheses are clustered at the DHS cluster level. Each regression includes district and month-birth year specific fixed effects. DHS controls include child, mother, and household level characteristics. Weather controls include non-linear transformation of precipitation and temperature over child's life time.