

UC Riverside

UC Riverside Previously Published Works

Title

Rainfall shocks, soil health, and child health outcomes

Permalink

<https://escholarship.org/uc/item/63p413xm>

Journal

Population and Environment, 45(3)

ISSN

0199-0039

Author

Kishore, Siddharth

Publication Date

2023-09-01

DOI

10.1007/s11111-023-00433-4

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Rainfall shocks, soil health, and child health outcomes

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. The results show that a high level of soil organic carbon significantly reduces the negative impact of rainfall shock on children's weight-for-height z scores, but not on height-for-age z scores.

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 (a measure of short-term inadequate nutrition), child stunting (a measure of long-term inadequate nutrition), 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 under the age of 5 are stunted (too short for their

21 age) and 21% of children under the age of 5 are wasted (too thin for their
22 height). Indian agricultural production is vulnerable to climate change and,
23 without effective adaptation, can reduce food crop yields in the future by up
24 to 9% (Guiteras, 2009). Moreover, in India's recent past, shortages of staple
25 food crops, wheat and rice are associated with severe droughts and extreme
26 rainfall (Zaveri and B Lobell, 2019; Auffhammer et al, 2012). Child nutrition
27 and agricultural production in rural areas in the developing world are closely
28 linked (Webb and Block, 2012). Bakhtsiyarava and Grace (2021) in Ethiopia
29 demonstrated that more diversity in agricultural production during periods of
30 low rainfall can reduce the risk of chronic food insecurity among children.

31 Food shortages caused by crop failures due to extreme weather condi-
32 tions, and the resulting nutritional deprivation can negatively impact children's
33 health (Grace et al, 2012). Improved soil quality as measured by soil organic
34 carbon (SOC), commonly used in the literature, increases agricultural produc-
35 tion (Lal, 2006). Because of the water holding capacity, a high level of SOC
36 offers long-term drought resistance and reduces the frequency of crop failures
37 (Huang et al, 2021; Kane et al, 2021). SOC also provides agricultural prof-
38 its for small landowners in developing countries (Bhargava et al, 2018). My
39 research asks if SOC affects children's nutrition and health in a low-income
40 country. Then, I explore to what extent SOC offers resilience during periods
41 of low rainfall.

42 This article examines whether natural variation in soil organic carbon lev-
43 els mitigates the impact of non-linear weather variables by crop growth on
44 children's health. Focusing on rural India, I leverage the 2015 Demographic
45 and Health Survey dataset and high-resolution spatial data on soil organic car-
46 bon content and meteorological variables. Following (Bakhtsiyarava and Grace,
47 2021), I evaluate the variation in anthropometric measurements, height-for-age

48 (HAZ) and weight-for-height z scores (WHZ) to measure child malnutrition
49 in India. Inadequate nutrition can cause childhood stunting (if HAZ is below
50 2 standard deviation) and wasting (if WHZ is below 2 standard deviation).
51 Unlike stunting, wasting may be reversed by increasing nutritional intake
52 (Victora, 1992). In this study, I focus on HAZ and WHZ scores to measure
53 malnutrition linked to weather-induced food insecurity.

54 While the exact relationship between soil quality and crop production
55 under dry conditions is complex and multidimensional. Huang et al (2021) and
56 Kane et al (2021) in the United States show that a higher soil organic carbon
57 content can moderate the impact of weather shocks by retaining soil water
58 in the agricultural systems. Children's nutrition also depends on food quality,
59 which is partly dependent on soil micro-nutrients (Berkhout et al, 2019; Kim
60 and Bevis, 2019). Berkhout et al (2019), based on their study in Sub-Saharan
61 Africa, highlight the importance of soil micro-nutrients such as zinc, copper
62 and manganese in reducing the malnutrition in children.

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

74 The results show that a high level of soil organic carbon significantly
75 reduces the negative impact of rainfall shock on children’s weight-for-height z
76 scores, but not on height-for-age z scores. I also explore heterogeneity in chil-
77 dren’s health outcomes by gender, household wealth index and land ownership,
78 and climate zone. This suggests that efforts to improve soil quality should be
79 adjusted to address these heterogeneous impacts. The results of the paper pro-
80 vide new evidence and inform policy-makers on the impact of high organic
81 carbon in soils on children’s health.

82 **2 Conceptual Framework**

83 Figure 1 depicts a simple conceptual connection between soil health and child-
84 hood nutrition. The figure can be used to examine the impact of a rainfall
85 shock with different levels of SOC. Because periods of low precipitation reduce
86 crop yields, food shortages affect food intake and thus nutrition (Grace et al,
87 2012). Higher SOC levels increase in agricultural production, particularly dur-
88 ing a drought (Lal, 2006), which contributes to food availability and supports
89 nutrition through consumption of output and income from crop sales that can
90 be used to purchase food. Because of the water holding capacity, a high level
91 of SOC offers long-term drought resistance and reduces the frequency of crop
92 failures (Huang et al, 2021; Kane et al, 2021). This reduction in crop fail-
93 ure increases agricultural income overall (Bhargava et al, 2018) and can thus
94 contribute to food security and nutrition for children by providing an extra
95 cushion against shocks.

96 Furthermore, the level of education of the mother, the gender of the child
97 and the wealth of the household can also influence the nutrition of the chil-
98 dren (Almond and Currie, 2011). Moreover, SOC mitigation effects may vary
99 depending on climate regions and the ability of households to cope with rain

100 shocks. Later in the results section, I estimate the heterogeneity in children’s
101 health outcomes by region, climate zone, gender, household wealth and land
102 ownership. Also, there may be unobserved covariates which may be correlated
103 with children’s nutrition and soil organic carbon levels and therefore may bias
104 my results downwards.

105 **3 Data and Descriptive statistics**

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

110 **3.1 Demographic and Health Data**

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

123 131 of the 28,526 geo-referenced clusters did not have information and were
124 dropped. I extracted environmental data using the DHS geo-referenced cluster
125 for a 10-km buffer.¹

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

136 **3.2 Rainfall Data**

137 I draw monthly rainfall data from Climate Hazards Group Infrared Precipitation
138 (CHIRPS) using DHS cluster geocoordinates. CHIRPS is a quasi-global
139 that extends over 50 S-50 N, with a gridded resolution of 0.05 degrees, from
140 1981 to near-real time precipitation time series (Funk et al, 2014).

141 There is not much guidance available in the literature about defining rain
142 shock. For my purpose, I need to define a rainfall shock based on a threshold
143 that lowers yields on India’s major crops. Therefore, like Feeny et al (2021), I
144 adopt an empirical strategy to determine the threshold. Using data from the
145 International Crops Research Institute for the Semi-Arid Tropic (ICRISAT),
146 I regress the natural log of the annual crop yield (Kg per hectare) from 2001
147 to 2015 on rainfall deciles controlling for year and district fixed effects.² The

¹As a sensitivity test, I run every analysis for a 20-km buffer. Appendix Table A10 reports the main results.

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

148 unit of analysis for the yield data is the district-year. As shown in the Figure
 149 2, results indicate that rainfall below the 20th percentile reduces crop yield of
 150 grains and pulses in India.³

151 Additionally, I also check the moderating effects of high SOC on crop yields.
 152 I interact with rainfall deciles and high SOC levels. The absolute impact of a
 153 high level of SOC is not statistically significant. But, the terms of interaction
 154 between precipitation deciles and high SOC are statistically significant for
 155 rainfall deciles 1 and 7. The results suggest that SOC moderated the impact
 156 of fluctuations in precipitation on yields in my analysis. Appendix Table A1
 157 report the results.

158 I define rain shock as a monsoon rain that is below the 20th percentile of
 159 the long-term historical mean within the DHS cluster (Shah and Steinberg,
 160 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:

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

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

164 A child under the age of 5 years may be exposed to one, many or no
 165 monsoon rainfall shock; the fraction of shocks captures that intensity of shock.
 166 For example, if a child of age 3 was exposed twice to rainfall shocks over his

³In 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.

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

167 or her lifetime then the fraction of shocks for that child is given by $2/4$. To
 168 measure the *in-utero* exposure to rainfall shock, I used the birthyear of the
 169 individuals observed in the DHS data. Appendix A2 shows how the shock
 170 fraction is distributed.

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

175 **3.3 Growing Degree Days**

176 Daily temperature was sourced from Indian Monsoon Data Assimilation and
 177 Analysis (IMDAA) reanalysis portal, managed by the National Centre for
 178 Medium Range Weather Forecasting (NCMRWF), India ([Rani et al, 2021](#)).
 179 Reanalysis Data Service (RDS) is a regional atmospheric reanalysis over the
 180 Indian subcontinent at a high resolution 0.12×0.12 from 1979-2018.⁶ I have
 181 followed the formulation used in previous studies using meteorological mea-
 182 sures which affect crop losses ([Guiteras, 2009](#)).⁷ Using the maximum and
 183 minimum daily temperature, the lower and upper threshold for calculating
 184 Growing Degree Days (GDD) during a growing season were set to $8C$ and
 185 $32C$, respectively.

⁵For 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>

⁶Available at <https://rds.ncmrwf.gov.in/datasets>

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

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

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).

⁸Soil data are available at <https://www.openlandmap.org>

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

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

¹⁰I also perform the sensitivity test for different threshold values such as 25th and 75th percentile of high soil organic carbon. Appendix table A8 and A9 report the results.

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

3.5 Descriptive statistics

Anthropometric data or body measurements for children, such as weight-for-age and weight-for-height, are taken and compared to a table in the World Health Organization (WHO) Child Growth Standards to calculate z-scores (WHO, 2006). The WHO Child Growth Standards are based on a sample of children from six countries: Brazil, Ghana, India, Norway, Oman and the United States of America. The z-score value can be either negative or positive depending on whether a child's anthropometric measurement is below or above the population average for the child's age and sex. 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 and weight-for-height 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.

¹²Appendix Table A3 describes the variables included in the research.

233 In my sample, the average age of children is 30 months. 51 per cent are
 234 boys and 49 per cent are girls. On average, mothers are 27 years of age and
 235 approximately half of the women have a high school or higher education. A
 236 little over half the households have agricultural land. Just under a third of
 237 households have potable water lines and a third have flush toilets. 23 per
 238 cent of families in my sample are poor. Appendix Table A4 presents summary
 239 statistics for all control variables used in this study.

240 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

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

241 where h_{ij} denotes child health outcomes measured by the height-for-age,
 242 weight-for-age, and the weight-for-height z-scores for child i at the DHS cluster
 243 level, j ; $shock_{ij}$ represents the fraction of rain shocks experience by child i
 244 residing at DHS cluster level, j ; soc_j represents the mean soil organic carbon
 245 content at the DHS cluster level, j ; \mathbf{X}_i is a set of explanatory variables including
 246 child, mother, and household characteristics. Child characteristics include
 247 age, gender and order of birth; mother characteristics include age, level of education
 248 and diet; and household characteristics include religion, social group,
 249 household income, and the wealth index (see Appendix Table A3 for a complete
 250 list of control variables); δ_d denotes district fixed effects and captures
 251 the time-invariant unobserved heterogeneity at the district level; ϕ_{my} denotes

252 child birth year-month specific fixed effects and captures within cohort varia-
253 tions, and ε_{ij} denotes the disturbance terms. I cluster the standard errors at
254 the level of DHS cluster (equivalent to Census village).

255 Additionally, I control precipitation and temperature derivatives (growth
256 degree-days and harmful degree-days) during a growing season (June through
257 September) throughout a child's life. $f(\theta)_{ij}$ is a non-linear function of precipi-
258 tation and temperature. I followed (Dimitrova and Muttarak, 2020) to include
259 a restricted cubic age spline, $f(a)_i$ with knots 6, 12, 18, 24, 36, and 48 months
260 of age to control for non-linearity in children's growth trajectory. The key
261 parameters are β_1 , β_2 , and β_3 . β_1 represents the impact of cumulative periods
262 of low precipitation on children's health; β_2 represents the direct impact of a
263 high level of SOC on children's health; and β_3 represents the mitigation effects
264 of a high level of SOC during cumulative periods of low rainfall.

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

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

279 coarsened exact matching method to estimate causal effects by reducing the
 280 covariate imbalance between treatment and control groups (Iacus et al, 2012).
 281 However, it may not circumvent the sample selection problem. I present the
 282 results for the matched sample in the online supplement section.

283 5 Results

284 5.1 Rainfall shocks, soil health, and child health

285 Table 2 presents impact of rainfall shock and soil health on children's health.
 286 The OLS model takes into account the characteristics of the child, the mother,
 287 the household. Moreover, the model controls a child's lifetime exposure to rain
 288 and temperature during a growing season. The model includes district and
 289 month and year of birth fixed effects. Standard errors are clustered at the
 290 DHS cluster level. Because the dependent variables (HAZ/WHZ scores) are
 291 measured as standard deviation, this is not a simple linear interpretation of a
 292 unit change in the exposure variable that results in a linear change in outcomes.
 293 The shock fraction shows a significant negative association with child WHZ.
 294 A one standard deviation increase in rainfall shock exposure above the child's
 295 mean years of exposure implies that the child will have negative WHZ score
 296 of -0.029 (-0.161×0.182) . The standard deviation for the fraction of shocks
 297 is 0.182. A high level of SOC has no effect on children's health at its main
 298 term, but substantially reduces the negative effect of the precipitation shock
 299 by 13.6 percentage points. The interaction term between SOC and fraction
 300 of shocks, which captures the compensating effect of a high soil quality. A
 301 one standard deviation increase in rainfall shock exposure above the child's
 302 average years of exposure in high SOC region leads to a positive WHZ score
 303 of 0.025 (0.136×0.182) . This implies that high levels of SOC have a moderator
 304 effect of 3% $((-0.161 + 0.136) / -0.991 = 0.025)$. The mean WHZ score is -0.991.

305 The p-value in the row shows the joint hypothesis test for High SOC and
306 High SOC*Fraction of shocks. The effect of a high and low level of SOC is
307 statistically same for the child's HAZ scores, but statistically different for the
308 WHZ scores.

309 Figures 5 illustrates the predictive margins and average marginal effects of
310 high SOC on HAZ and WHZ scores. Panels (a) and (b) show the predicted
311 margins of HAZ and WHZ scores stratified by high and low SOC levels at each
312 precipitation shock level. In panel (a), the predicted margins of the child's HAZ
313 scores are on an upward slope. This means that the predicted HAZ scores for
314 a child living in both low and high SOC areas become less negative at high
315 shock intensity. In contrast, in panel (b), the predicted margins of the child's
316 WHZ scores are declining. This means that the predicted WHZ scores for a
317 child living in both low and high SOC areas become more negative. However,
318 the magnitude of the predicted WHZ score is significantly smaller for a child
319 living in a high SOC area. Panels (c) and (d) tell us the difference in HAZ and
320 WHZ scores between high and low SOC groups at each precipitation shock
321 level. The difference between low and high SOC areas is approximately zero
322 for HAZ scores, while it is greater than zero for WHZ scores. This suggests
323 that a high level of SOC significantly affects the child's WHZ scores, but not
324 the HAZ scores.

325 Next, there is a concern that soil organic carbon measurement may be con-
326 founded by other associated agronomic attributes. With SOC as the choice
327 variable, it is difficult to remove concerns related to the omitted variable bias.
328 Nevertheless, I approach this concern by including soil texture, slope and veg-
329 etative index as control variables in Equation 1.¹³¹⁴ In order to assess the

¹³I 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).

¹⁴I used the enhanced vegetation index for 2015 available in the DHS dataset as a proxy for agricultural output.

330 influence of the different soil attributes used in this study on children’s health,
331 I ran a correlation between child WHZ and soil attributes. This demonstrates
332 no concern for multicollinearity in the model. Table A6 in the appendix pro-
333 vides the correlation matrix for the soil attributes used in this study. Appendix
334 Table A7 report the results. It reads similar effects of high SOC on child health
335 outcomes.

336 As a robustness check, I perform regressions on the matched sample after
337 applying the matching algorithm (discussed in the online supplement) and on
338 the population-weighted monthly rainfall measurements. Appendix Table B1
339 present the results after applying the coarsened exact matching weights to the
340 OLS model. The sign of the estimated coefficients is identical to that of the
341 main results. However, the key coefficients are not significant at the 5 per cent
342 significance level in the matched sample. Next, Appendix Table A5, which uses
343 the population-weighted monthly rain measures, reads similar effects on child
344 health.

345 **5.2 Heterogeneity**

346 As mentioned earlier in the conceptual section, I use the subset of the sample
347 to explore these heterogeneous effects of different climate zones, gender and
348 the household wealth index, common in the development and climate litera-
349 ture. I also explore the heterogeneous effect that soil health has on the health
350 outcomes of children in households with and without agricultural land. To
351 check whether the differences between heterogeneous groups are statistically
352 different, I perform a simple statistical test (two-sample t-test). The results
353 suggest that the differences in mean HAZ/WHZ scores between groups differ
354 statistically.

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.¹⁵ They are tropical wet, tropical wet and dry, arid, semi-arid, humid sub-tropical, and mountainous. The Köppen classification map is based on local vegetation which, in turn, is based on local precipitation and temperature. The tropical rainforest and the tropical monsoon are reclassified as tropical humid while the tropical savannah is reclassified as tropical humid and dry. The wet season in summer and the dry season in winter are the characteristics of the humid and dry tropical region. A one-way ANOVA test for average differences in HAZ/WHZ scores indicates a significant difference between different climate zones.¹⁶ Appendix Figure [A3](#) shows the major climatic zones of India.

Heterogeneous effects across climate zones suggest that high SOC has a significant effect on children’s WHZ scores in semi-arid and humid sub-tropical areas. The impact of high SOC is larger in semi-arid areas. In contrast, cumulative precipitation shocks are positively associated with child WHZ scores in the wet and dry tropical climate area. This can be due to a reduction of diseases that are common during monsoon weather such as diarrhea and malaria. But this requires further research, and the results should be interpreted cautiously. The results do not suggest any impact of a high level of SOC on children’s HAZ scores in major but semi-arid climate zones. Table [3](#) summarizes the heterogeneous effects across climate zones.

¹⁵I am grateful to Anna Dimitrova for sharing the data and code with me.

¹⁶see Appendix Table [A12](#)

379 **5.2.2 Heterogeneity by gender**

380 Next, I disaggregate the sample into boys and girls. There is evidence of gender
381 discrimination in the literature in response to different types of household
382 shocks, including environmental shocks. SOC may have no direct relationship
383 to gender. But the moderating effect of a high level of SOC against crop failure
384 contributes to children's food and nutritional security.

385 Table 4 presents the heterogeneous effects of rainfall shocks and soil health
386 on children's HAZ and WHZ scores by gender. Cumulative rain shock has a
387 negative impact on girls' and boys' WHZ scores. Girls are more affected by rain
388 shocks, as suggested by the larger coefficient. The point estimation is -0.205
389 for girls and -0.112 for boys. A high level of SOC mitigates the negative impact
390 of precipitation shock on girls' WHZ scores, but not for boys. This implies
391 that the nutrition of girls is addressed where resilience to climate-induced food
392 insecurity exists through a high level of SOC. On the other hand, the results
393 show no effect of a high level of SOC on children's HAZ scores.

394 **5.2.3 Heterogeneity by household wealth index**

395 Household characteristics, such as household wealth, directly impact the
396 resilience of households to absorb shocks, including environmental shocks. Poor
397 households have less resilience than non-poor households. In the DHS dataset,
398 I observe five different indices: the poorest, the poorer, the middle, the richer,
399 and the richest. For my purpose, I code the poorest and the poorer as the poor
400 and the middle, the richer, and the richest as the non-poor.

401 Table 4 presents the heterogeneous effects of rainfall shocks and soil health
402 on children's HAZ and WHZ scores by household wealth index, as defined in
403 the DHS data. Unsurprisingly, the results indicate that poor households are
404 negatively affected by precipitation shocks. The point estimate is -0.197 and

405 significant at the 5% significance level. Children WHZ scores in poor rural
406 households are negatively impacted by rainfall shocks. A high level of SOC
407 does not significantly reduce the adverse effect of the rainfall shock on poor
408 households. The ability of rural households to influence SOC and directly ben-
409 efit from high SOC may depend on their association with farms. I explore that
410 in more detail in the next section by disaggregating the sample by landowner.

411 **5.2.4 Heterogeneity by land ownership**

412 To show the relationship between SOC and landowners, I look at the hetero-
413 geneity by land ownership: has agricultural land and does not have agricultural
414 land. Table 4 presents the results for households that own and do not own
415 farmland. The results suggest that rain shocks negatively affect households
416 that own land, suggesting they are rain-dependent. However, a high level of
417 SOC does not reduce the negative impact of rainfall shock on households that
418 own land.

419 **6 Conclusion**

420 **6.1 Summary**

421 This paper examines the relationship between SOC and the impact of rainfall
422 shocks on children's health. Based on the conceptual framework developed in
423 this paper, I examine the impact of rainfall shock with different levels of SOC.
424 The results show that a high level of SOC significantly reduces the negative
425 impact of rainfall shock on children's WHZ scores in rural India. My findings
426 are consistent with two separate literatures: (1) studies (e.g., [Dimitrova and](#)
427 [Muttarak \(2020\)](#)) which show the negative impact of precipitation shock on
428 children's health and (2) studies (e.g., [Berkhout et al \(2019\)](#)) that show the

429 importance of soil quality in the reduction of malnutrition in low- and middle-
430 income countries. In this paper, I show a significant moderating effect of a
431 high level of SOC offering resilience from the rainfall shock on short-term
432 inadequate nutrition in rural areas of India. I find significant reduction in
433 children's negative WHZ scores and thus a resistance to child wasting during
434 periods of low precipitation in a rain-fed farming country. However, I find no
435 effect of a high level of SOC on children's HAZ scores suggesting that a high
436 SOC does not reduce chronic malnutrition among children.

437 Analyses of heterogeneous impacts suggest that high SOC significantly
438 moderates the precipitation shocks in semi-arid and humid sub-tropical cli-
439 mate zones. The results suggest that the nutrition of girl child is addressed in
440 high level of SOC areas in response to rainfall shock. The results also suggest
441 that a high level of SOC does not significantly mitigate the negative impact of
442 rainfall shocks for children from poor households and farm landowners. One
443 plausible reason of these results is that poor households and farm landown-
444 ers in rural areas of India are characterized by a small landholding size (the
445 average landholding in India is 1 hectare) and therefore we may not expect a
446 stronger effect of a high SOC.

447 **6.2 Limitation**

448 A major limitation of this paper is that the soil organic carbon content variable
449 used is time invariant. Existing research shows that agricultural practices that
450 cause pollution, such as stubble burning ([Singh et al, 2019](#)) and fertilizer use
451 ([Brainerd and Menon, 2014](#)), can have negative impacts on children's health.
452 These agricultural practices may also affect soil organic carbon concentrations,
453 leading to endogeneity issues in the estimates. Due to a lack of data, I am

454 unable to control for these agricultural practices and address the endogeneity
455 problem.

456 **6.3 Conclusions**

457 Since it takes longer to reflect changes in soil organic carbon concentrations,
458 policies may include both long-term and short-term measures. One long-term
459 policy to enhance SOC would be to incentivize the adoption of agricultural
460 best management practices. This can increase resilience to shocks over time,
461 particularly as climate changes. Indian child development programs could be
462 improved by considering the impact of climate change on the incidence of
463 droughts, and consequently, on children's health.

464 In the short term, soil health in a region could be used to inform the
465 likely impacts of precipitation shocks, which could better target relief efforts.
466 Nutrition and soil conditions are linked to agriculture, and high soil quality
467 contributes to reducing malnutrition, particularly during precipitation shocks.
468 Therefore, there may be a greater need for food relief aid in low SOC areas.

469 **References**

- 470 Almond D, Currie J (2011) Killing me softly: The fetal origins hypothesis.
471 *Journal of economic perspectives* 25(3):153–72
- 472 Auffhammer M, Ramanathan V, Vincent JR (2012) Climate change, the
473 monsoon, and rice yield in India. *Climatic change* 111(2):411–424
- 474 Bakhtsiyarava M, Grace K (2021) Agricultural production diversity and child
475 nutrition in Ethiopia. *Food Security* 13(6):1407–1422
- 476 Berkhout ED, Malan M, Kram T (2019) Better soils for healthier lives? an
477 econometric assessment of the link between soil nutrients and malnutrition

- 478 in sub-saharan africa. PloS one 14(1):e0210,642
- 479 Bhargava AK, Vagen T, Gassner A (2018) Breaking ground: unearthing
480 the potential of high-resolution, remote-sensing soil data in understand-
481 ing agricultural profits and technology use in Sub-Saharan Africa. World
482 Development 105:352–366
- 483 Blackwell M, Iacus S, King G, et al (2009) cem: Coarsened exact matching in
484 stata. The Stata Journal 9(4):524–546
- 485 Brainerd E, Menon N (2014) Seasonal effects of water quality: The hidden
486 costs of the green revolution to infant and child health in India. Journal of
487 Development Economics 107:49–64
- 488 Center for International Earth Science Information Network - CIESIN -
489 Columbia University (2018) Gridded Population of the World, Version 4
490 (GPWv4): Population Count, Revision 11. URL [https://doi.org/10.7927/
491 H4JW8BX5](https://doi.org/10.7927/H4JW8BX5), place: Palisades, New York
- 492 Dimitrova A, Bora JK (2020) Monsoon weather and early childhood health in
493 India. PLOS ONE 15(4):e0231,479. [https://doi.org/10.1371/journal.pone.
494 0231479](https://doi.org/10.1371/journal.pone.0231479), URL <https://dx.plos.org/10.1371/journal.pone.0231479>
- 495 Dimitrova A, Muttarak R (2020) After the floods: Differential impacts of rain-
496 fall anomalies on child stunting in India. Global Environmental Change
497 64:102,130. <https://doi.org/10.1016/j.gloenvcha.2020.102130>, URL [https://
498 linkinghub.elsevier.com/retrieve/pii/S0959378020307135](https://linkinghub.elsevier.com/retrieve/pii/S0959378020307135)
- 499 Dinkelman T (2017) Long-run health repercussions of drought shocks: evidence
500 from south african homelands. The Economic Journal 127(604):1906–1939

- 501 Feeny S, Mishra A, Trinh TA, et al (2021) Early-Life exposure to rainfall
502 shocks and gender gaps in employment: Findings from Vietnam. *Journal of Economic Behavior & Organization* 183:533–554. [https://doi.org/10.](https://doi.org/10.1016/j.jebo.2021.01.016)
503 [1016/j.jebo.2021.01.016](https://doi.org/10.1016/j.jebo.2021.01.016), URL [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S0167268121000275)
504 [S0167268121000275](https://linkinghub.elsevier.com/retrieve/pii/S0167268121000275)
- 506 Funk CC, Peterson PJ, Landsfeld MF, et al (2014) A quasi-global precipitation
507 time series for drought monitoring. *US Geological Survey data series*
508 832(4):1–12
- 509 Grace K, Davenport F, Funk C, et al (2012) Child malnutrition and climate in
510 Sub-Saharan Africa: An analysis of recent trends in Kenya. *Applied Geog-*
511 *raphy* 35(1-2):405–413. <https://doi.org/10.1016/j.apgeog.2012.06.017>, URL
512 <https://linkinghub.elsevier.com/retrieve/pii/S0143622812000768>
- 513 Guiteras R (2009) The impact of climate change on Indian agriculture.
514 Manuscript, Department of Economics, University of Maryland, College
515 Park, Maryland
- 516 Hengl T (2018a) Clay content in % (kg / kg) at 6 standard depths (0, 10, 30,
517 60, 100 and 200 cm) at 250 m resolution. [https://doi.org/10.5281/zenodo.](https://doi.org/10.5281/zenodo.2525663)
518 [2525663](https://doi.org/10.5281/zenodo.2525663), URL <https://doi.org/10.5281/zenodo.2525663>
- 519 Hengl T (2018b) Sand content in % (kg / kg) at 6 standard depths (0, 10, 30,
520 60, 100 and 200 cm) at 250 m resolution. [https://doi.org/10.5281/zenodo.](https://doi.org/10.5281/zenodo.2525662)
521 [2525662](https://doi.org/10.5281/zenodo.2525662), URL <https://doi.org/10.5281/zenodo.2525662>
- 522 Hengl T (2018c) Silt content in % (kg / kg) at 6 standard depths (0, 10, 30,
523 60, 100 and 200 cm) at 250 m resolution. [https://doi.org/10.5281/zenodo.](https://doi.org/10.5281/zenodo.2525676)
524 [2525676](https://doi.org/10.5281/zenodo.2525676), URL <https://doi.org/10.5281/zenodo.2525676>

- 525 Hengl T, Wheeler I (2018) Soil organic carbon content in x 5 g / kg
526 at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m res-
527 olution. <https://doi.org/10.5281/zenodo.2525553>, URL [https://doi.org/10.](https://doi.org/10.5281/zenodo.2525553)
528 [5281/zenodo.2525553](https://doi.org/10.5281/zenodo.2525553)
- 529 Huang J, Hartemink AE, Kucharik CJ (2021) Soil-dependent responses of
530 US crop yields to climate variability and depth to groundwater. *Agricul-*
531 *tural Systems* 190:103,085. <https://doi.org/10.1016/j.agsy.2021.103085>,
532 URL <https://linkinghub.elsevier.com/retrieve/pii/S0308521X2100038X>
- 533 Iacus SM, King G, Porro G (2012) Causal inference without balance checking:
534 Coarsened exact matching. *Political analysis* 20(1):1–24
- 535 Kane DA, Bradford MA, Fuller E, et al (2021) Soil organic matter pro-
536 tects US maize yields and lowers crop insurance payouts under drought.
537 *Environmental Research Letters* 16(4):044,018
- 538 Kim K, Bevis L (2019) Soil fertility and poverty in developing countries.
539 *Choices* 34(2):1–8
- 540 Kim R, Bijral AS, Xu Y, et al (2021) Precision mapping child undernutri-
541 tion for nearly 600,000 inhabited census villages in India. *Proceedings of*
542 *the National Academy of Sciences* 118(18):e2025865,118. [https://doi.org/](https://doi.org/10.1073/pnas.2025865118)
543 [10.1073/pnas.2025865118](https://doi.org/10.1073/pnas.2025865118), URL [http://www.pnas.org/lookup/doi/10.1073/](http://www.pnas.org/lookup/doi/10.1073/pnas.2025865118)
544 [pnas.2025865118](https://doi.org/10.1073/pnas.2025865118)
- 545 Lal R (2004) Soil carbon sequestration impacts on global climate change and
546 food security. *science* 304(5677):1623–1627
- 547 Lal R (2006) Enhancing crop yields in the developing countries through
548 restoration of the soil organic carbon pool in agricultural lands. *Land*

- 549 degradation & development 17(2):197–209
- 550 McMahon K, Gray C (2021) Climate change, social vulnerability and child
551 nutrition in South Asia. *Global Environmental Change* 71:102,414
- 552 Rani SI, Arulalan T, George JP, et al (2021) Imdaa: High-resolution
553 satellite-era reanalysis for the indian monsoon region. *Journal of Climate*
554 34(12):5109–5133
- 555 Shah M, Steinberg BM (2017) Drought of opportunities: Contemporaneous and
556 long-term impacts of rainfall shocks on human capital. *Journal of Political*
557 *Economy* 125(2):527–561
- 558 Singh P, Dey S, Chowdhury S, et al (2019) Early life exposure to outdoor
559 air pollution: effect on child health in India. *Brookings India* URL [http:](http://hdl.handle.net/11540/9783)
560 [//hdl.handle.net/11540/9783](http://hdl.handle.net/11540/9783).
- 561 Victora CG (1992) The association between wasting and stunting: an interna-
562 tional perspective. *The Journal of nutrition* 122(5):1105–1110
- 563 Viswanathan B, Kumar KK (2015) Weather, agriculture and rural migration:
564 evidence from state and district level migration in India. *Environment and*
565 *Development Economics* 20(4):469–492
- 566 Webb P, Block S (2012) Support for agriculture during economic transforma-
567 tion: Impacts on poverty and undernutrition. *Proceedings of the National*
568 *Academy of Sciences* 109(31):12,309–12,314. [https://doi.org/10.1073/pnas.](https://doi.org/10.1073/pnas.0913334108)
569 [0913334108](https://doi.org/10.1073/pnas.0913334108), URL <http://www.pnas.org/cgi/doi/10.1073/pnas.0913334108>
- 570 WHO (2006) Reliability of anthropometric measurements in the WHO Multi-
571 centre Growth Reference Study. *Acta Paediatrica* 95:38–46

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

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

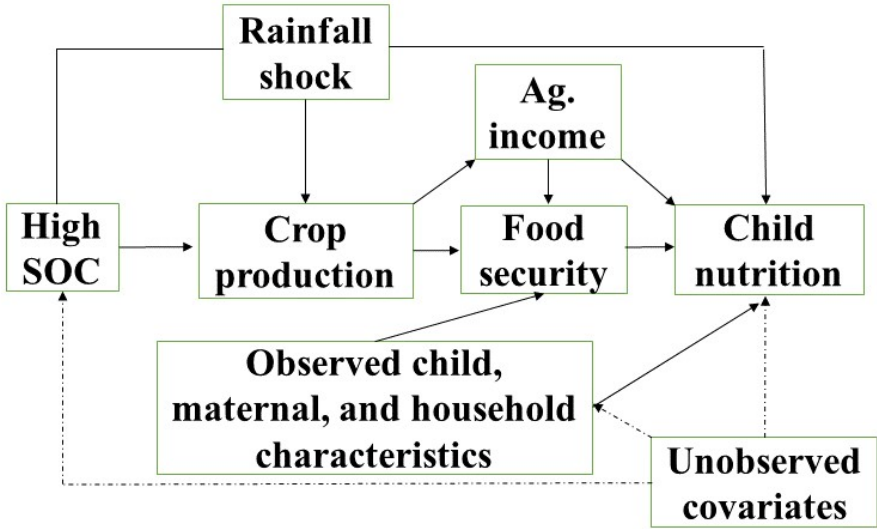


Fig. 1: A simple conceptual relationship between soil and children’s health.

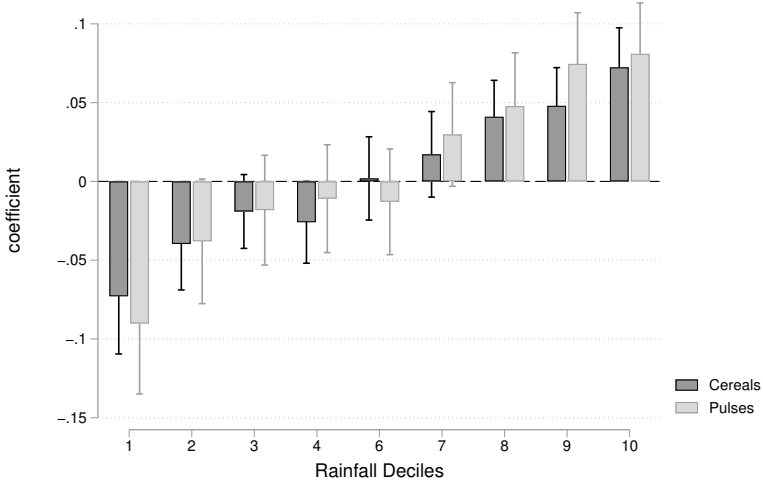


Fig. 2: Coefficient for rainfall deciles and 95% CI in India. The dependent variable is the natural logarithm of annual crop yield (kg per hectare) from 2001 to 2015. The specification includes district and year fixed effects. The 5th decile is selected as reference.

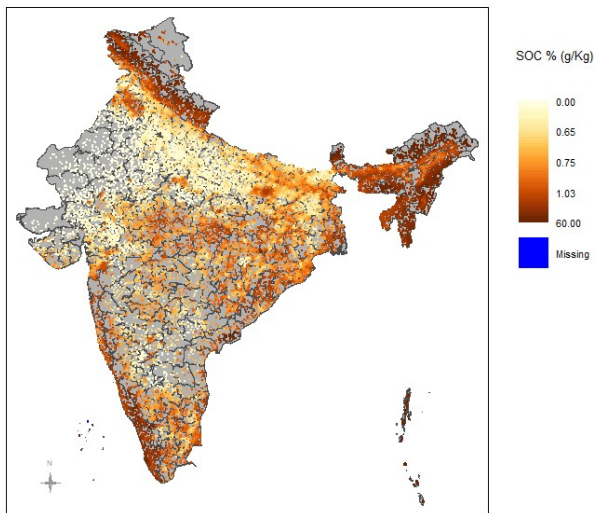


Fig. 3: The dots represent the average soil organic carbon content of the DHS rural clusters in India. The missing in the map indicates the null values for union territory Lakshadweep. The dark lines in the background are the district borders.

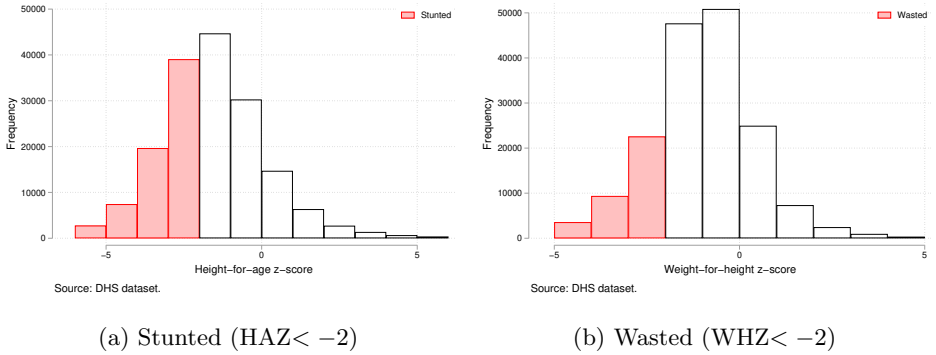
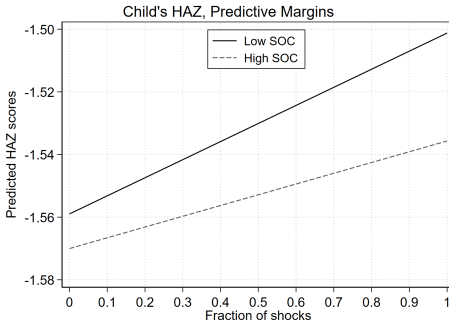


Fig. 4: Distribution of childhood health outcomes. Source: Own calculations based on DHS dataset (2015-2016).

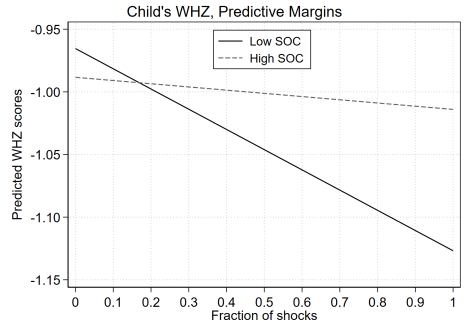
Table 1: Summary statistics.

	Observation	Mean	Std. Dev.
<i>Child health measures</i>			
Height-for-age z score	169,904	-1.558	1.681
Weight-for-height z score	169,904	-0.991	1.381
<i>Rainfall below 20th percentile, yes=1</i>			
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
<i>Soil health measure</i>			
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	

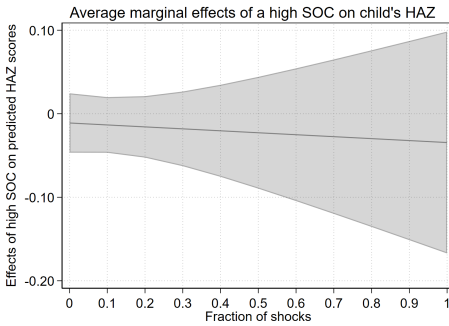
Note: The rain shock for the 1st to the 4th year have different observations to adjust the age of the child. The sample is composed of 33,951 4-year-olds, 69,621 3-year-olds, 103,642 2-year-olds, 137,807 1-year-olds and 169,904 in-utero. Source: DHS and CHIRPS data.



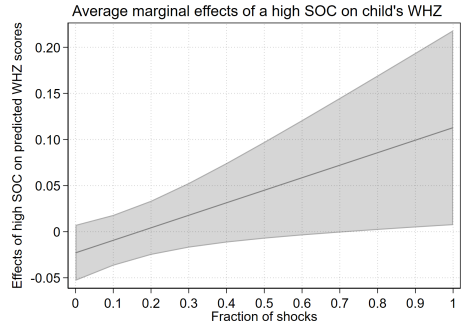
(a) Height-for-age z-scores



(b) Weight-for-height z-scores



(c) Height-for-age z-scores



(d) Weight-for-height z-scores

Fig. 5: Predictive margins stratified by high and low SOC levels and average marginal effects of high SOC levels on HAZ and WHZ scores.

Notes: Panel (a) / (b) was derived by predicting HAZ/WHZ scores to specified precipitation shock values per high/low SOC level. Panel (c) / (d) was obtained from a partial derivative of HAZ/WHZ scores with respect to High SOC. Panels (c) and (d) are shaded with a 95% confidence Interval.

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

	HAZ	WHZ
Fraction of shocks	0.058 (0.050)	-0.161*** (0.042)
High SOC	-0.011 (0.018)	-0.023 (0.015)
High SOC \times Fraction of shocks	-0.023 (0.072)	0.136** (0.059)
DHS controls	Yes	Yes
Weather controls	Yes	Yes
P-val: High SOC + High SOC \times Fraction of shocks = 0	0.611	0.036
Mean dependent. var.	-1.558	-0.991
SD dependent var.	1.681	1.381
Observations	169,904	169,904
R-square	0.148	0.090

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 level of SOC is a dummy variable, 1 for the value above the 50th percentile of SOC and 0 otherwise. 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 (1.661)	-0.167 (0.167)	0.130 (0.130)	0.062 (0.061)
High SOC	0.451 (0.284)	-0.009 (0.029)	0.023 (0.051)	-0.031 (0.027)
High SOC × Fraction of shocks	1.084 (1.661)	0.175 (0.183)	-0.767** (0.322)	0.052 (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 (0.813)	0.292** (0.139)	-0.280** (0.111)	-0.133*** (0.051)
High SOC	0.314* (0.175)	0.016 (0.026)	-0.054 (0.041)	-0.021 (0.023)
High SOC × Fraction of shocks	0.200 (0.819)	-0.416*** (0.153)	0.547** (0.275)	0.180** (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

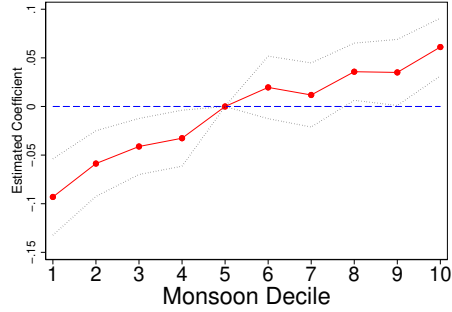
Levels of significance: $p < 0.01^{***}$, $p < 0.05^{**}$, $p < 0.1^*$. Robust standard errors in parentheses are clustered at the DHS cluster level. The high level of SOC is a dummy variable, 1 for the value above the 50th percentile of SOC and 0 otherwise. Each regression includes district and month-birth 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. A one-way ANOVA test for average differences in HAZ/WHZ scores indicates a significant difference between different climate zones (see Appendix Table A12).

Table 4: Heterogeneity by individual and household characteristics

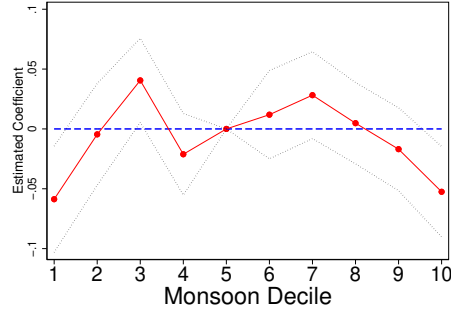
	Boys		Girls	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.011 (0.065)	-0.112** (0.057)	0.108 (0.067)	-0.205*** (0.055)
High SOC	-0.023 (0.022)	-0.005 (0.020)	0.003 (0.024)	-0.041** (0.020)
High SOC × Fraction of shocks	0.022 (0.093)	0.110 (0.079)	-0.068 (0.094)	0.152** (0.077)
Mean dependent. var.	-1.597	-1.017	-1.516	-0.963
Observations	87,643	87,643	82,259	82,259
R-square	0.142	0.096	0.165	0.093
Difference of average HAZ scores by gender				-0.081***
Difference of average HAZ scores by gender				-0.054***
	Poor		Non-poor	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.060 (0.069)	-0.197*** (0.056)	0.056 (0.069)	-0.110* (0.060)
High SOC	-0.000 (0.026)	-0.012 (0.022)	0.019 (0.024)	-0.029 (0.020)
High SOC × Fraction of shocks	0.002 (0.104)	0.114 (0.081)	-0.038 (0.092)	0.133* (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
Difference of average HAZ scores by wealth				0.525***
Difference of average WHZ scores by wealth				0.262***
	Has ag. land		Has no ag. land	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.083 (0.063)	-0.190*** (0.054)	0.033 (0.075)	-0.110* (0.062)
High SOC	-0.015 (0.023)	-0.025 (0.020)	-0.006 (0.026)	-0.012 (0.021)
High SOC × Fraction of shocks	-0.112 (0.090)	0.119 (0.076)	0.089 (0.104)	0.132 (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
Difference of average HAZ scores by landowner				-0.106***
Difference of average WHZ scores by landowner				-0.033***

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 level of SOC is a dummy variable, 1 for the value above the 50th percentile of SOC and 0 otherwise. 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 difference in mean HAZ/WHZ scores by heterogeneous groups was obtained from a two-sample t-test with equal variances.

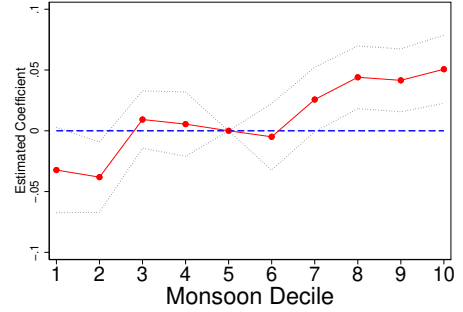
578 **A Additional Figures and Tables**



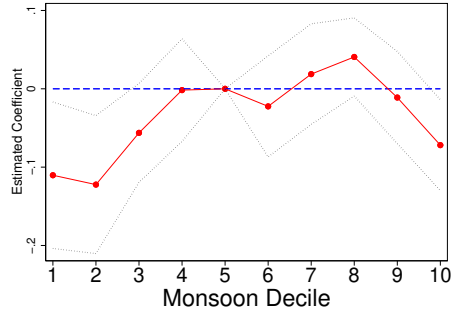
(a) Rice



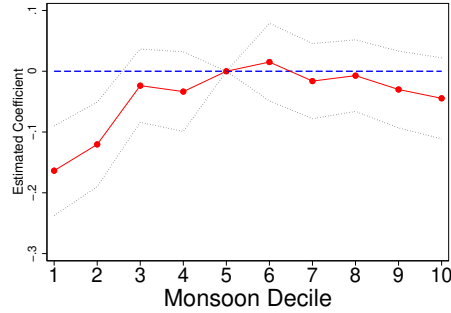
(b) Maize



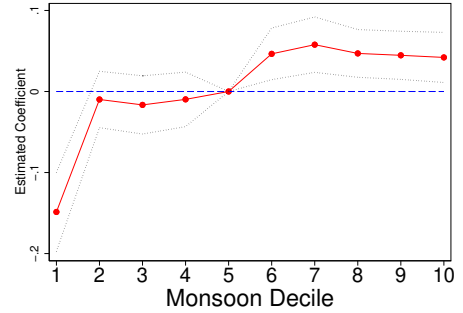
(c) Wheat



(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).

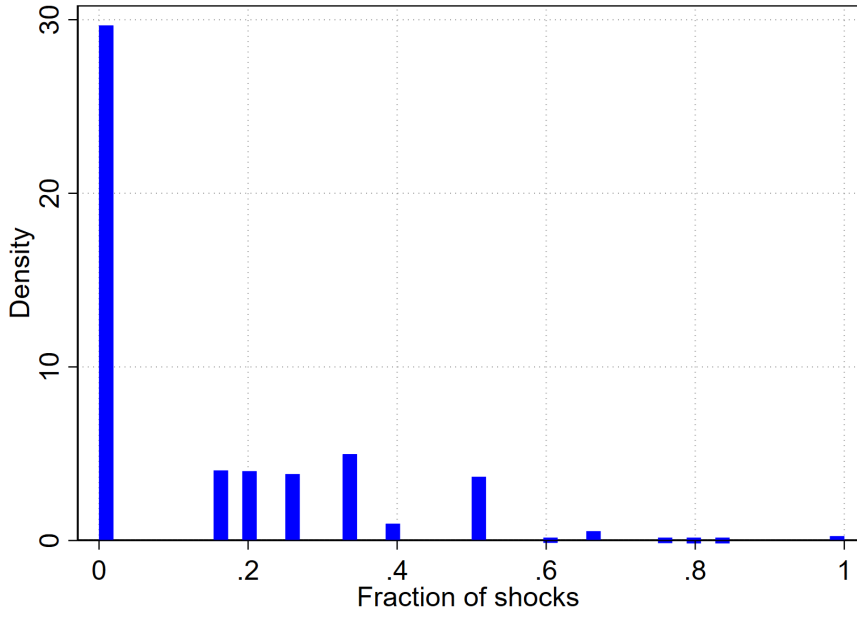


Fig. A2: The distribution of fractional shocks.

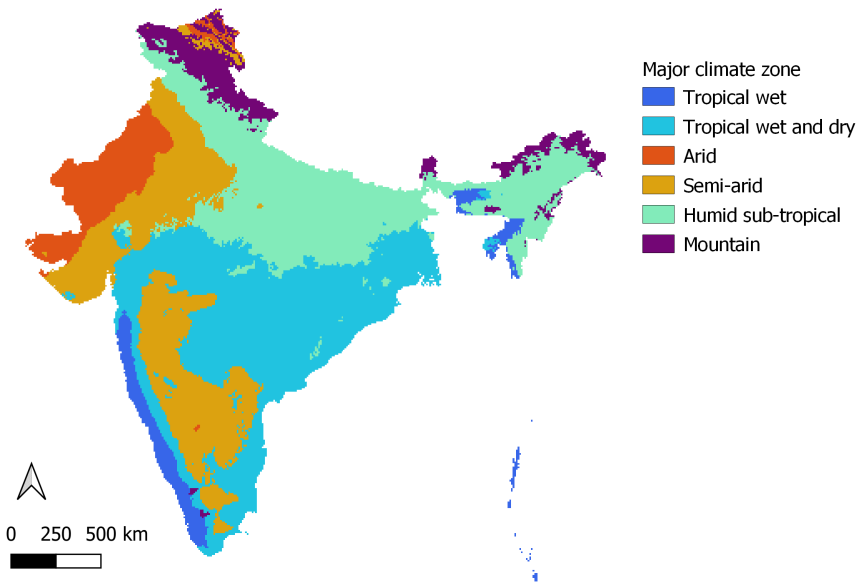


Fig. A3: Major climate zones in India based on Köppen Geiger climate classification.

Table A1: Moderating impacts of high SOC on crop yields

	Cereal
SOC (%)	0.018 (0.020)
Rainfall decile 1	-0.129*** (0.029)
Rainfall decile 1 x SOC	0.038** (0.018)
Rainfall decile 2	-0.051* (0.027)
Rainfall decile 2 x SOC	0.010 (0.020)
Rainfall decile 3	-0.016 (0.024)
Rainfall decile 3 x SOC	0.004 (0.019)
Rainfall decile 4	-0.045 (0.030)
Rainfall decile 4 x SOC	0.034 (0.022)
Rainfall decile 6	-0.016 (0.025)
Rainfall decile 6 x SOC	0.015 (0.021)
Rainfall decile 7	-0.049 (0.031)
Rainfall decile 7 x SOC	0.060*** (0.023)
Rainfall decile 8	0.049** (0.021)
Rainfall decile 8 x SOC	-0.001 (0.020)
Rainfall decile 9	0.071** (0.028)
Rainfall decile 9 x SOC	-0.015 (0.026)
Rainfall decile 10	0.084*** (0.025)
Rainfall decile 10 x SOC	-0.001 (0.019)
Observations	7091
Adjusted R^2	0.460

Levels of significance: $p < 0.01$ ***, $p < 0.05$ ** , $p < 0.1$ * . Robust standard errors in parentheses are clustered at the district level. The dependent variable is the natural logarithm of annual crop yield (kg per hectare) from 2001 to 2015. The specification includes state and year fixed effects. SOC is continuous. The 5th decile is selected as reference.

Table A2: Description for variables included in the study

Variable	Description
Child-specific	
<i>bord</i>	Order of birth
<i>malechild</i>	Dummy for male child
<i>childsizelarge</i>	Dummy for child was large at birth
<i>childsizavg</i>	Dummy for child was average at birth
<i>numfemalesib</i>	Number of female siblings
<i>nummalesib</i>	Number of male siblings
<i>age510</i>	Dummy for child with sibling between the age 5 and 10 years
<i>age1115</i>	Dummy for child with sibling between the age 11 and 15 years
<i>age16</i>	Dummy for child with sibling above 16 years
<i>hw1</i>	Child's age in months
Woman-specific	
<i>v012</i>	Woman's age in years
<i>womanpriedu</i>	Dummy for woman has primary education
<i>womansecedu</i>	Dummy for woman has secondary or higher level education
<i>womanbmi</i>	Woman's body mass index
<i>womaneatfruits</i>	Dummy for woman consumes fruits daily or weekly
<i>womaneatveges</i>	Dummy for woman consumes vegetables daily or weekly
<i>womaneat eggs</i>	Dummy for woman consumes eggs daily or weekly
<i>womaneatmeat</i>	Dummy for woman consumes chicken/meat/fish daily or weekly
<i>womansmoke</i>	Dummy for woman smokes
<i>womandrunk</i>	Dummy for woman drinks alcohol
<i>womanprenataldoc</i>	Dummy for had prenatal care with doctor
Household-specific	
<i>v104</i>	Years lived in place of residence
<i>hv220</i>	Age of household head in years
<i>hhheadmale</i>	Dummy for male household head
<i>hhhindu</i>	Dummy for household religion is Hinduism
<i>hhmuslim</i>	Dummy for household religion is Islam
<i>hhscst</i>	Dummy for household belongs to SC/ST
<i>hhradio</i>	Dummy for household owns a radio/transistor
<i>hhtv</i>	Dummy for household owns a television
<i>hhrefri</i>	Dummy for household owns a refrigerator
<i>hhmotorcycle</i>	Dummy for household owns a motorcycle
<i>hhcar</i>	Dummy for household owns a car
<i>hhelec</i>	Dummy for household has electricity
<i>hv244</i>	Dummy for household owns agricultural land
<i>hhirragland</i>	Dummy for household irrigate agricultural land
<i>sh52a</i>	Dummy for household owns cows/bulls/buffaloes
<i>sh52b</i>	Dummy for household owns camels
<i>sh52c</i>	Dummy for household owns horses/donkeys/mules
<i>sh52d</i>	Dummy for household owns goats
<i>sh52e</i>	Dummy for household owns sheep
<i>sh52f</i>	Dummy for household owns chickens/ducks
<i>hhpipewater</i>	Dummy for source of drinking water: piped water
<i>hhgroundwater</i>	Dummy for source of drinking water: ground water
<i>hhsurfacewater</i>	Dummy for source of drinking water: surface water
<i>hhrainwater</i>	Dummy for source of drinking water: rain water, tanker water, etc
<i>hhflush toilet</i>	Dummy for toilet facility: flush toilet
<i>hhpit</i>	Dummy for toilet facility: pit toilet/latrine
<i>hhnofacility</i>	Dummy for toilet facility: no facility/bush/field
<i>hhpoorest</i>	Dummy for household wealth index: poorest
<i>hhpoorer</i>	Dummy for household wealth index: poorer
<i>hhmiddle</i>	Dummy for household wealth index: middle
<i>hhricher</i>	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 A3: Description for variables included in the study

Variable	Description
Weather-specific	
<i>childrain</i>	June-September daily accumulation of rainfall over child's life time
<i>childgdd</i>	Growing degree days over child's life time
<i>childhdd</i>	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 A4: 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 A5: Alternative main regression results using population-weighted rain measures

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.027 (0.050)	-0.143*** (0.042)	-0.016 (0.062)	-0.046 (0.052)
High SOC	-0.011 (0.018)	-0.020 (0.015)	-0.008 (0.021)	-0.021 (0.018)
High SOC × Fraction of shocks	-0.017 (0.072)	0.102* (0.058)	-0.073 (0.089)	0.038 (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 A6: 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 A7: Robustness check: confounding variables included as controls

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.056 (0.050)	-0.166*** (0.042)	0.070 (0.063)	-0.152*** (0.053)
High SOC	-0.011 (0.018)	-0.023 (0.016)	-0.005 (0.021)	-0.016 (0.018)
High SOC × Fraction of shocks	-0.022 (0.072)	0.135** (0.059)	-0.089 (0.090)	0.098 (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 A8: Sensitivity test for various thresholds: High soil organic carbon content above 25 percentile.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.153** (0.063)	-0.233*** (0.055)	0.155** (0.067)	-0.214*** (0.059)
High SOC	0.012 (0.022)	-0.026 (0.018)	0.016 (0.027)	-0.017 (0.022)
High SOC × Fraction of shocks	-0.147** (0.072)	0.186*** (0.061)	-0.233*** (0.086)	0.151** (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 A9: Sensitivity test for various thresholds: High soil organic carbon content above 75 percentile.

	Full		Matched	
	HAZ	WHZ	HAZ	WHZ
Fraction of shocks	0.048 (0.042)	-0.114*** (0.037)	0.122 (0.093)	-0.091 (0.080)
High SOC	-0.015 (0.028)	-0.022 (0.023)	-0.022 (0.034)	-0.000 (0.029)
High SOC × Fraction of shocks	-0.003 (0.085)	0.066 (0.072)	-0.122 (0.124)	-0.020 (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 A10: Sensitivity test for different DHS cluster level: 20 km

	(1)	(2)	(3)	(4)
	Full	Full	Full	Matched
Fraction of shocks	-0.241*** (0.054)	-0.242*** (0.054)	-0.260*** (0.055)	-0.229*** (0.060)
High SOC	-0.017 (0.018)	-0.023 (0.018)	-0.023 (0.018)	-0.008 (0.022)
High SOC × Fraction of shocks	0.129** (0.061)	0.154** (0.061)	0.163*** (0.061)	0.109 (0.076)
Marginal effects	-0.144*** (0.034)	-0.127*** (0.033)	-0.137*** (0.034)	-0.174*** (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 A11: Sensitivity check for different SOC thresholds

	Child wasting at various SOC thresholds.		
	25th	50th	75th
Fraction of shocks	1.482*** (0.131)	1.309*** (0.088)	1.290*** (0.076)
High SOC	1.008 (0.032)	1.053** (0.027)	1.065 (0.041)
High SOC × Fraction of shocks	0.764*** (0.075)	0.855 (0.082)	0.750** (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.

Table A12: ANOVA test

	Sum of square	Degree of freedom	Mean square	F	Prob > F
<i>HAZ scores</i>					
Between groups	3108.95	5	621.79	221.46	0.000
Within groups	476230.63	169619	2.81		
<i>WHZ scores</i>					
Between groups	4316.99	5	863.40	458.79	0.000
Within groups	319203.80	169619	1.88		

Note: A one-way ANOVA test was performed to test if there is a difference in the mean HAZ/WHZ scores between different climate zones.

B Online Supplement

The main idea of this paper was to use soil organic carbon (as a measure of soil health) as a moderator in response to rainfall shock. In the long run, the soil organic carbon may not be exogenous and may be correlated to an omitted variable, resulting in biased estimates. I take advantage of the coarsened exact matching method to estimate causal effects by reducing the covariate imbalance between treatment (high SOC region) and control (low SOC region) groups.

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).¹⁷ 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. 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. 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.

¹⁷I 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 B2 presents the results. It reads findings similar to those of the main specification.

The estimating equation is similar to the equation (1):

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

603 where the terms are defined same as the equation (1). I applied the package
 604 in STATA, *cem*, to compute the weights and those weights were used in a
 605 simple weighted regression.

Table B1: Impact of high levels of SOC on the health of children.

	HAZ	WHZ
Fraction of shocks	0.011 (0.063)	-0.063 (0.053)
High SOC	-0.008 (0.021)	-0.023 (0.018)
High SOC × Fraction of shocks	-0.071 (0.089)	0.057 (0.071)
DHS controls	Yes	Yes
Weather controls	Yes	Yes
Mean dependent. var.	-1.573	-1.059
SD dependent var.	1.667	1.366
Observations	102,296	102,296
R-square	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 matched weights, *cem*, are applied on all regressions. 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 B2: Including dry and rainy seasons as an additional variable in the matching algorithm.

	HAZ	WHZ
Fraction of shocks	0.030 (0.063)	-0.102* (0.053)
High SOC	-0.011 (0.021)	-0.016 (0.018)
High SOC \times Fraction of shocks	-0.072 (0.091)	0.036 (0.072)
DHS controls	Yes	Yes
Weather controls	Yes	Yes
Mean dependent var.	-1.580	-1.065
SD dependent var.	1.665	1.366
Observations	97,441	97,441
R-square	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 matched weights, *cem*, are applied on all regressions. 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.