Rainfall shocks, soil health, and child health outcomes

Siddharth Kishore

Abstract

This paper explores the impact of monsoon rainfall shocks on child health outcomes in rural India at varying levels of soil organic carbon. I combine high resolution spatial data on soil organic carbon content and weather with the Demographic and Health Survey (DHS) for India (2015-2016) to estimate the linkage between monsoon rainfall shock, soil health, and child health. Weather variables and soil determine crop productivity and thus affect human health through food access in low and middle income countries. I contribute to the literature by demonstrating direct and indirect impact of soil health on childhood outcome of wasting. Using a coarsened exact matching method, I estimate that having high soil health can result in 26 percent improvement in child wasting. I also demonstrate that having high soil health can moderate adverse impacts from weather shocks.

Keywords: Rainfall shock, soil organic carbon, wasting, matching methods

1 Introduction

India shows the poorest performance in South Asia for child health[1]. The increased risk of child malnourishment is evident in childhood wasting (i.e., children under the age of 5 with low weight for their height) and child stunting (i.e., children under the age of 5 years with low height for their age). It is well documented that environmental shocks experienced during childhood can have long-term effects on individual human capital accumulation and adult outcomes (Almond and Currie, 2011; Heckman, 2007; Rosales-Rueda, 2016; Maccini and Yang, 2009; Feeny et al, 2021). The research asks whether a higher level of soil organic carbon content (a measure of soil health) can moderate

[1] India poorly ranked on global hunger index score, which is calculated on four indicators - prevalence of undernourishment, child wasting, child stunting, and mortality of children aged below 5.

the negative impact of monsoon rainfall shocks through soil water retention in the agricultural systems on child health.

Crop yields are likely to decline as a consequence of increased weather volatility. The parts of the world that are going to be the most adversely affected are low and middle income countries in Asia and Africa. A higher soil quality as measured by soil organic carbon (soc), commonly used in literature, for example Bhargava et al (2018) and Kane et al (2021), provides resistance to extended dry periods and reduces cases of crop failure. Child nutrition depends on food quality which partially depends upon soil health. In this paper, I contribute to the literature by demonstrating the direct and indirect impacts of soil health. Due to having improved soil health, households would have access to increased food quality, which may improve child health outcomes. This reflects the direct impact of soil health. Soil health can help moderate the impact of adverse weather outcomes on food quantity. This is labeled as an indirect impact of soil health.

Berkhout et al (2019), based on their study in Sub-Saharan Africa, highlight the importance of soil nutrients such as zinc, copper and manganese in reducing the malnutrition in children. Although the exact relationship between soil quality and crop production in drought like condition is complex. Huang et al (2021) and Kane et al (2021) shows that a higher soil organic carbon content can moderate the impact of weather shocks through soil water retention in the agricultural systems.

This paper explores whether natural variation in organic soil carbon content moderates the impact of non-linear weather variables on child health. I use the anthropometric measurements such as height-for-age z score (HAZ), weight-for-age z score (WAZ), and Weight-for-height z score (WHZ) to understand the adverse effect of weather fluctuations on child health in rural India. For the purpose of my analysis, I focus on the weight-for-height z score because the current stage of child nutrition is reflected in it. In this study, the nutritional status and weather are dependent and is a likely mechanism of impact.

This paper is informed and contributes to two main strands of literature: first is the relationship between agronomy and climate; second is the relationship between child health and meteorological variables. I construct a non-linear function of meteorological variables, such as precipitation and temperature, that have been shown to impact agricultural yield and, hence, incomes[2]. Webb and Block (2012) have documented that regions with a higher agricultural growth have a lower incidence of child stunting. Others have also looked into child health outcomes by assessing anthropometric measurements in developing countries; in particular, Dimitrova and Muttarak (2020) and McMahon and Gray (2021) have studied the impact of monsoon rainfall on child health outcomes by linking the demographic and health survey data for India to the meteorological variables. My paper is different from earlier studies as I explore

[2]see for example (Taraz, 2018)

the linkage between rainfall shock, soil health, and child health through agricultural mechanism. I incorporated soil organic carbon content as an additional environment variable into the framework, discussed in the next section, to understand whether soil organic carbon content provides resilience (because of its water retention capacity) to rural households in response to monsoon rainfall shocks.

Rural economies in developing countries primarily revolve around the agricultural sector, and many of the agricultural production operations depend on monsoon rainfall[3]. Hence, any monsoon rainfall shock directly impacts the livelihood of rural households engaged in on-farm and in off-farm activities. This scenario is likely to be exacerbated by climate change, in particular, by a higher incidence of droughts[4]. Given the climatic conditions, farmers' adaptation is slow and is often drive through an external impetus, such as policies at the federal and or authorised local bodies. The resource-poor status of rural households leads to a lack of ability to adopt technology that could potentially mitigate the effect of short-term weather fluctuations and long-term climate change in developing countries. The weather fluctuation also impact the time-dependent nature of agricultural activities. For example, a misallocation of labor[5] is often observed, where the decision to hire labor at regular time is interrupted due to delay in the onset of monsoon rainfall. This may cascade to other operating processes such as sowing, weeding, and harvesting. The impact of frequent extreme weather events such as drought and deluge is severely felt by economically vulnerable regions, mainly due to lack of accountability and administrative neglect from local authorities. This study combines health data with environmental and climatic variables to estimate the impact of a better soil endowment in response to monsoon rainfall shock.

This paper contributes to the literature on climate change impacts on child health by assessing anthropometric measurements in developing countries. The empirical results indicate that a move from low to high soil quality reduces the effect of rainfall shock on child health outcomes. This paper highlights climate change issues, and the findings motivate climate change interventions at the policy level.

The remainder of this paper is organized as follows: section 2 presents the conceptual framework; section 3 goes over the data and descriptive statistics; sections 4 and 5 discuss the empirical framework and results; and section 6 presents concluding remarks.

---

[3] India receives most of its rainfall in the monsoon season from June through September.

[4] Much of the population residing in the global south have in recent decades witnessed an increased frequency of droughts. Source: EM-DAT Public, available at https://public.emdat.be/mapping

[5] The allocation of labor to the agricultural activities often disrupted because of uncertainty of monsoon season. Here, misallocation of labor means the shortage of hired labor.

2 Conceptual Framework

The health of agricultural sector in low- and middle- income countries is linked to the development of human health, education, and labor allocation (Shah and Steinberg, 2017; Mahajan, 2017).

Temperature increase, attributed to climate change, reduces global yields of major crops (Zhao et al, 2017). Decline in crop yields results in food shortages affecting the food intake and thus nutrition. It is the nutritional shock induced by extreme weather is the likely mechanism that negatively impacts the human health. A higher levels of soil organic carbon content can moderate the impact of weather shocks (Huang et al, 2021; Kane et al, 2021). Weather variables such as temperature and precipitation, and soil determine crop productivity and thus affect human health through food availability and access in low and middle income countries. Figure 1 shows a simple conceptual link between temperature, rainfall, soil health and child health.

Soil health have direct impact on food production. A decline in food production because of poor soil health directly impacts food availability, access and utilization, which in turn impacts the prevalence of undernourishment and malnutrition. Weather induced loss in crop yields results in increased food prices (Letta et al, 2022). An upward pressure on the food prices resulting from a decline in food availability further reduces access and utilization. This pathway connecting climatic variables to child health is well established in literature (Grace et al, 2012).

I incorporated soil quality as an additional environmental variable into the framework to understand whether soil organic carbon content provides resilience to households in response to monsoon rainfall shocks.

3 Data and Descriptive statistics

The climate and soil health (as measured by soil organic carbon content) data sources are combined with health data to investigate the impact of soil health on child health in response to rainfall shocks defined as the total rainfall in a growing season for a given year below 20th percentile of long-term rainfall within a geo-referenced DHS cluster. Table A7 presents the description of the variables included in the study.

3.1 Health Data

I use cross-sectional data from round four of the Demographic and Health Survey (DHS) for India collected in the year 2015-2016. The sample is representative at the national level.

The sample size for children aged 0-4 years was 259,627; 34,625 observations were excluded from the child data file that had missing or invalid information. 131 out of 28,526 geo-referenced clusters had no information and were excluded. After the merge between environmental and health data, there were 223,977 observations. I restricted the sample size to rural areas to focus on

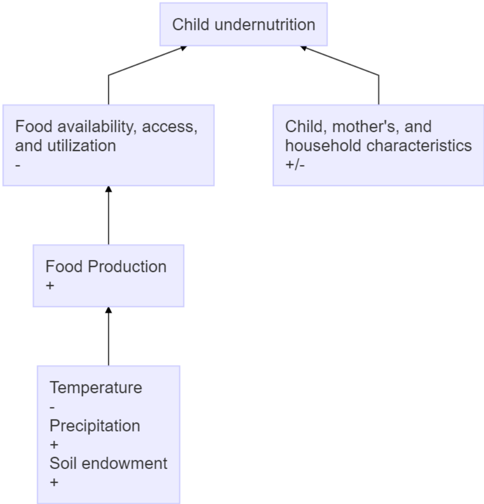


Fig. 1: A simple conceptual link between temperature, precipitation, soil health and child undernourishment.

the agricultural productivity (53,437 observations were excluded). 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. I performed analysis on a sample size of 169,904.

For the purpose of my analysis, I use DHS cluster geo-cordinates for a buffer area of 10[6] km.

3.2 Rainfall Data

Monthly rainfall data was extracted from the Climate Hazards Group Infrared Precipitation (CHIRPS) using DHS cluster geo-coordinates for a buffer area 10 km. CHIRPS is a quasi-global that spans 50S-50N, gridded 0.05 degree resolution, 1981 to near-real time precipitation time series (Funk et al, 2014). I constructed rainfall deciles using long-term rainfall in the growing season, from June through September, over a period of 36 years from 1981 through 2016.

For my purpose, I need to define a rainfall shock that lowers the crop yields of major crops in India. I use unapportioned data from 2001 to 2015 obtained from International Crops Research Institute for the Semi-Arid Tropic (ICRISAT)[7]. I then regress the natural log of annual crop yield from 2001 to 2015 on rainfall deciles. Figure 2 shows that monsoon rainfall below 20th[8]

[6] As a sensitivity test, I perform all the analysis for a buffer area of 20 km. Appendix table A6 report the results.
[7] Crop yield data is available at http://data.icrisat.org/dld/index.html
[8] Shah and Steinberg (2017) have also used rainfall below 20th percentile as a threshold to define drought.

percentile lowers the crop yields (kg per hectare) of cereals and pulses[9] in India. I define rainfall shock as monsoon rainfall that is below 20th percentile of long-run historical average within the DHS cluster.

I used a measure of rainfall shock that has been used in the literature before (Feeny et al, 2021; Dinkelman, 2017). I constructed a variable fraction of shocks to study the effect of monsoon rainfall shock on child health. 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 the rainfall shock over his lifetime then the fraction of shocks for that child is given by 2/4[10].

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)[11].

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[12]. I followed the formulation used in earlier studies that use weather measures that affect crop losses (Guiteras, 2009)[13]. Using daily maximum and minimum temperature, the lower and upper threshold to calculate the Growing Degree Days (GDD) in a growing season were set at $8C$ and $32C$, respectively. The daily temperature were extracted for a buffer area of 10 km around DHS cluster and is converted to degree Celsius from Kelvin using R package *weathermetrics*.

[9] In appendix, figure A1, I also show the negative effects of less rainfall on selected stable and cash crops. Maize, soybean, and cotton seem to be different and not monotonically increasing with rainfall suggesting that weather-yield response of some major crops are non-linear.

[10] I calculate fraction of shocks as:

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

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

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

[13] I used the following formula to convert daily mean temperature into GDD:

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

3.4 Soil Data

Soil organic carbon content data were collected from the OpenLandMap[14] (Hengl and Wheeler, 2018). The maps were generated based on machine learning predictions from global compilations of soil profiles at 250 − m resolution. I extracted the average soil organic carbon content at the buffer of 10 km around the geo-coded DHS clusters at four standard depths: 0, 10, 30, and 60 cm. I then calculated the depth-weighted soil organic carbon content at 0-60 cm[15] intervals for the analysis. There is not much guidance in the literature on the threshold used to categorize soil quality. I constructed two categories of soil organic carbon content: low, below the 25th percentile, and high, above the 25th percentile.[16]

Figure A2 shows the map of soil organic carbon of rural DHS clusters. The missing in the map indicates the null values for union territory Lakshadweep. A large part of the country that is habituated is characterized by low soil organic carbon content areas. The West and East coastal region, much of the north eastern region, and the central plains are characterized by moderate to high soil carbon content areas.

3.5 Summary statistics

Figures A3 and A4 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.

Table 1 reports the summary statistics for the data used in this study. The table shows that about 41 and 21 percent of children in rural India suffer from stunting and wasting, respectively. About 11 percent of children were exposed to rainfall shock in their birth year and in-utero[17]. The average value for the fraction of shocks as a measure of intensity is 0.13.

Appendix table A1 reports the summary statistics of all the control variables used in this study.

3.6 Matching methods

The coarsened exact matching method estimates the sample average treatment effect on the treated (Blackwell et al, 2009). I employ knowledge of the data in order to find better match. The coarsened variables used were: a) child-specific (child birth order, gender and age of child); b) mother-specific (mother's age

---

[14] Soil data is available at https://www.openlandmap.org

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

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

[16] I also perform the sensitivity test for different threshold values such as 50th and 75th percentile of high soil organic carbon. Appendix table A4 and A5 report the results.

[17] To measure the *in-utero* exposure to rainfall shock, I used the birth year of individuals observed in the DHS data.

and education); and c) household-specific (religion, caste, source of drinking water, and toilet facility). I apply the software package, *cem*[18] created by (Blackwell et al, 2009) was used to calculate weights and those weights were used in a simple weighted regression. 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). I take the advantage of coarsened exact matching method to improve the estimation of causal effects by reducing the imbalance in covariates between treated and control groups (Iacus et al, 2012).

4 Empirical Framework

I estimated an OLS regression model to find the impact of monsoon rainfall shocks on child health outcomes. The main specification is given by

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

where h_{ij} denotes child health outcomes measured by the height-for-age, weight-for-age, and the weight-for-height z-scores for children i at the DHS cluster level, j; $shock_{jt}$ represents the fraction of monsoon rainfall shocks at DHS cluster level, j in time t; soc_j represents the soil organic carbon content at the DHS cluster level, j; $f(\theta)_{jt}$ is a function of monsoon rainfall and temperature at the DHS cluster level, j in time t; $\mathbf{X}_i$ is the vector of explanatory variables including child, mother, and household characteristics; δ_d denotes district fixed effects and captures the 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.

In this study, I assume the soil endowments to be exogenous. The treatment variable is high soil organic carbon content.

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.

5 Results and Discussion

5.1 Baseline model

Table 2 presents the effect of monsoon rainfall shock and soil health on child health outcomes measured by weight-for-height z-scores. Children exposed to monsoon rainfall shocks *in-utero* through age 4 are more likely to weigh less for their height compared to the mean weight-for-height z score, as shown in all the specifications[19]. The coefficient on the fraction of shocks in column 3 is -0.234

[18]The *cem* command with an option *k2k* in STATA produces a matching result that has the same number of treated and control in each matched strata by randomly dropping observations. The following is the matching summary: the number of balanced matched observation is 40,129 for treated and control, and the unmatched observation is 2354 out of 42,483 for control (0) and 87,292 out of 127,421 for treated (1).

[19]Appendix table A2, which uses the population-weighted monthly rain measures, reads similar effects on child health.

and statistically different from zero at the 1 percent significance level. A one standard deviation increase in rainfall shock exposure above the child average years of exposure implies that the child will have 0.043 (0.234*0.182[20]) lower likelihood of weight-for-height z score. This suggests that a greater exposure to rainfall shock *in-utero* through age 4 negatively affects the child health. The key coefficient is on the interaction term between soil organic carbon content and fraction of shocks, which captures the compensating effect of a better soil quality. The coefficient on the interaction term is 0.186 and significant at the 1 percent significance level suggesting that a higher soil organic carbon content moderates 5 percent ($\frac{-0.234+0.186}{-0.991} = 0.048$) of the negative effect of monsoon rainfall on child weight-for-height z score. The average marginal effect of fraction of shocks is -0.094, suggesting that a high soil organic carbon content lowers the effect of monsoon rainfall shocks on child health.

Column 4 in table 2 report the results after applying the coarsened exact matching weights to the OLS model to estimate the causal effect of soil quality on child health in response to monsoon rainfall shocks. A one standard deviation increase in rainfall shock exposure above the child average years of exposure implies that child will have 0.039 (0.215*0.182) lower likelihood of weight-for-height z score. The average marginal effect of fraction of shocks is -0.134 and significant at the 1 percent significance level. The effect is larger in using the matched specification. The average years of exposure on children is also higher in the matched specification.

Figure 3 show the average marginal effects of high soil organic carbon on child weight-for-height z score. The average marginal effects of high soil organic carbon is larger for a greater cumulative exposure to rainfall shock.

There may be a concern that the soil organic carbon measure is likely to be confounded by other related agronomic attributes. I address this concern by including the soil texture[21], slope, and vegetation index[22] as control variables in the equation 1. Appendix table A9 report the results. It reads similar effects on child health.

Table 3 report the results of average marginal effects of rainfall shock on timing of child outcomes. The results suggest that the timing of exposure to rainfall shock matters for child health outcomes; in particular, the weight-for-age z-score for children less than 3 years old is negatively associated with shocks and is in line with the literature. The moderating effect of high soil organic carbon is realized the most for children below 2 years old. The average marginal effect of rainfall shock in first year for height-for-age z score outcome is 0.177 and significant at the 5 percent level.

[20]The standard deviation for the fraction of shocks variable is 0.182.

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

[22]I used enhanced vegetation index for 2015 available at the DHS dataset as a proxy for the agricultural production.

5.2 Heterogeneity by region

Table 4 shows the heterogeneous effects of monsoon rainfall shock and soil health on child WHZ score by region[23]. The variables of the model are the same as the base specification. Each column presents regression results after processing the dataset using *cem* weights for a region defined earlier. I restricted the sample size to each region based off of the higher and lower incidence of child WHZ score. For example, the western region has the lowest weight-for-height for children aged 5 years and below (mean WHZ score is -1.277) followed by central plains (mean WHZ score is -1.198) and the eastern region (mean WHZ score is -1.046). The interaction term for eastern region is 1.069 and significant at the 5 percent significance level, and the average marginal effect of fraction of shocks at the mean soil organic carbon value is -0.085 and is nonsignificant at the 10 percent level. The interaction term for the northern region that has the lowest incidence of negative average WHZ score is -1.045 and significant at the 5 percent significance level. The high soil quality in the eastern region provides resilience in response to rainfall shocks. The average marginal effect of fraction of shocks for the eastern region is -0.380 and nonsignificant at the 10 percent level. The Indian green revolution started in the northern region which comprised of Indian states of Punjab and Haryana (Brainerd and Menon, 2014). The states in northern region has irrigation facility though high usages of crop chemicals and monocropping, this region grows irrigated wheat, has greater declines in the soil health than others.

The results suggest a regional disparity across India, the southern part has a relatively lower incidence of negative child health outcomes compared to other parts in India. Much of the southern part represents a low to moderate level of soil organic carbon content and a lower monsoon rainfall. The average marginal effect of fraction of shocks is -0.043 and nonsignificant at the 10 percent level.

5.3 Heterogeneity by household wealth index

Table 5 presents the heterogeneous effects of monsoon rainfall shock and soil health on child WHZ score by household wealth index as defined in the DHS data. The results suggest that both poor and non-poor households are negatively affected by rainfall shocks. However, the magnitude of the effect is larger for poor households. The point estimate is -0.198 for the poor households and is -0.188 for the non-poor households. Both estimates are significant at the 5 percent significance level. The average marginal effect of fraction of shocks for the poor household is -0.132 and significant at the 1 percent significance level. Better soil quality does reduce the negative effect of rainfall shock. Typical

[23] The north region is comprised of Indian states: Punjab and Haryana; the north central (NC) region is comprised of Uttar Pradesh, Bihar, and Jharkhand; the central plains (CP) region is comprised of Chhattisgarh, madhya Pradesh, and Rajasthan; the eastern region comprises of Assam, Odisha, and West Bengal; the western region is comprised of Gujarat, Maharashtra, Goa, and union territories such as Dadra and Nagar Haveli, and Daman and Diu; the southern region is comprised of Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Telangana, and Puducherry

poor households are characterized by low landholding size (if any land ownership), and thus we would not expect a larger effect of soil organic carbon content on child health through the agricultural income effect.

5.4 Heterogeneity by gender

Table 6 presents the heterogeneous effects of monsoon rainfall shocks and soil health on child WHZ score by gender. The fraction of shocks negatively impacts the female and male child WHZ scores. Female children bear a larger effect of rainfall shocks, as suggested by the larger magnitude on the coefficient. The point estimate is -0.219 for female child and is -0.188 for male child. The average marginal effect of shocks is -0.140 for female child and significant at the 1 percent significance level. The p-value of the test of difference in the coefficient across female and male groups for the interaction terms between high SOC and fraction of shocks is not statistically different from zero. The average marginal effect of fraction of shocks for male child is -0.108 and nonsignificane at the 5 percent significance level. A better soil quality provides resilience in response to rainfall shocks for female child. We further explore the effect of monsoon rainfall shocks on female child health in the next section.

5.5 Heterogeneity by female sibling

In rural households, due to inadequate endowments to lean on in response to shock, female children may have to compete with their siblings for already scarce resources. Table 7 presents the effects of monsoon rainfall shocks and soil health on female child health outcomes[24]. The focus is on female children. I restricted the sample data to female children with at least one female sibling and at least one male sibling to estimate the sibling rivalry. The results suggest that monsoon rainfall shocks severely affects WHZ rates of female children, and the effect is even larger for a female child in the presence of at least one other female sibling. The point estimate is -0.320 and significant at the 1 percent significance level. The average marginal effect of fraction of shocks is -0.202 and significant at the 1 percent significance level. It is evident that a move from a low to high soil organic carbon content provides resilience in response to rainfall shocks to female child.

5.6 Heterogeneity by climate zone

The effect of soil organic carbon content on child health in response to monsoon rainfall shock might differ by different climatic zones in India. Following Dimitrova and Bora (2020), I constructed six major climate zones at the district level based on Köppen Geiger climate classification.[25]: Tropical wet, tropical wet and dry, Arid, Semi-arid, Humid sub-tropical, and Mountain. We would expect Arid and Semi-arid zones to be less resilient to rainfall shocks while,

[24] Simply put, a male child is preferred over a female child and hence the resources available at the household favors a male child.

[25] I thank Anna Dimitrova for sharing the data and code with me.

Tropical wet and Humid sub-tropical to be relatively more resilient to rainfall shocks.

Table 8 presents the heterogeneous effects of rainfall shocks on WHZ rates by selected climate zones. The fraction of shocks negatively affects the child health outcomes residing in the semi-arid climate zones more than other climatic zones. The point estimate is -0.269 and significant at the 5 percent significance level. The average marginal effect of fraction of shocks is -0.245 and significant at the 5 percent level.

5.7 Extended results

5.7.1 Logistic estimate

I estimate Logistic regression model to predict whether a move from low to high SOC reduce the probability of child being stunted or wasted in response to rainfall shocks. Table 9 report the effect of monsoon rainfall shocks on likelihood of child stunting and wasting. The dependent variable is binary for child stunting that is height-for-age below -2 ($HAZ < -2$) is 1; 0 otherwise. Similarly, binary for child wasting cases that is weight-for-height below -2 ($WHZ < -2$) is 1; 0 otherwise. The results suggest that children exposed to monsoon rainfall shocks *in-utero* through age 4 are more likely to be stunted and wasted. The magnitude on the child wasting is substantially larger than child stunting. Columns 3 and 4 present the results after applying coarsened exact matching weights to logit model to estimate the causal effect of high soil organic carbon content on child health in response to rainfall shocks.

The odds of child wasting are 53 percent higher when exposed to cumulative rainfall shock. A higher soil organic content can moderate the negative effect of rainfall shock on child wasting by 26 percent (1-0.736=0.264). The average marginal effect of fraction of shocks is -0.030 for stunted and is 0.045 for wasted. Both estimates are significant at the 1 percent significance level. Child stunting cases are chronic in nature and difficult to explain through agricultural mechanism between rainfall, soil heath, and child nutrition.

Figure A5 and A6 shows the average marginal effects of high SOC on child stunting and wasting respectively. The left panel in the figure A6 shows us the difference between two probabilities that is child wasting with low SOC and high SOC. The right panel shows the average marginal effects of high SOC on child wasting indicating that as the fraction of shocks increases a high SOC has larger effect on reducing the child wasting, the confidence interval grows larger with the fraction of shocks. In other words, a move from low to high soil organic carbon content reduces the incidence of child wasting.

6 Conclusion

This paper estimates the linkage between soil quality and child health in response to monsoon rainfall shock. I used soil organic carbon content at 250 m spatial resolution as a measure of soil quality. I employed coarsened exact matching methods to estimate the casual effects of soil organic carbon content on child health in response to monsoon rainfall shock. The result suggest that soil organic carbon content moderates the effect of monsoon rainfall shock on children as the intensity increases. However, the results are sensitive to different threshold values of soil organic carbon content.

The results from this study indicate that a higher soil quality can moderate the negative effects of monsoon rainfall shock on child health. The predicted probability of wasting reduces sharply with a move from moderate to high soil quality in response to monsoon rainfall shocks. Poor households are more vulnerable to shock relative to non-poor households; female children are severely disadvantaged relative to male children; and the effect of shock on female children with at least one female sibling is even larger. The results also suggest that the semi-arid climatic zone is most affected by the monsoon rainfall shock compared to other climatic zones.

It is common to witness large weather volatilities, characterized by both drought conditions in some seasons and flooding in other seasons. These extreme weather events can adversely impact the soil with varying degrees. One limitation is that the soil organic carbon content variable is time invariant. Some regions may experience greater declines in soil organic carbon content overtime than others, which may lead to measurement error. For instance, fields that grow wheat, they also experience stubble burning, which can reduce soil organic carbon by disrupting the process of nitrogen cycle of soil. However, because of the time invariant measure of soil, I am unable to capture that variation. Given this, my estimate is an under-estimate. Nonetheless, 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. Integrated child development programs must be strengthened to incorporate a critical component of the effect of climate change on child health. Food nutrients and soil conditions are interlinked through agriculture, and better soil quality helps reduce malnutrition.

References

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

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

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

Heckman JJ (2007) The economics, technology, and neuroscience of human capability formation. Proceedings of the national Academy of Sciences 104(33):13,250–13,255

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

Letta M, Montalbano P, Pierre G (2022) Weather shocks, traders' expectations, and food prices. American Journal of Agricultural Economics 104(3):1100–1119

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

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

Rosales-Rueda M (2016) Impact of early life shocks on human capital formation: Evidence from el nino floods in ecuador. University of California Irvine

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

Taraz V (2018) Can farmers adapt to higher temperatures? evidence from india. World Development 112:205–219

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

Zhao C, Liu B, Piao S, et al (2017) Temperature increase reduces global yields of major crops in four independent estimates. Proceedings of the National Academy of Sciences 114(35):9326–9331

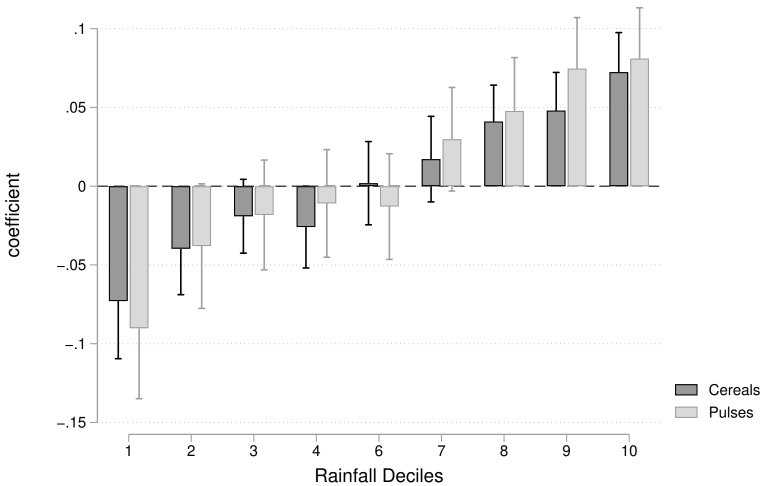


Fig. 2: Coefficient for rainfall deciles and 95% CI in India. The dependent variable is the natural logarithm of annual crop yield (kg per hectare) from 2001 to 2015. The specification includes district and year fixed effects. The 5th decile is selected as reference.

Appendix A Additional Figures and Tables

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 content (%)	169,897	0.945	0.675

Note: The rainfall shock for the first to the fourth year have different observations to adjust for child's age. For example, 33,951 children of age 4 years are present in the sample. Source: DHS and CHIRPS data.

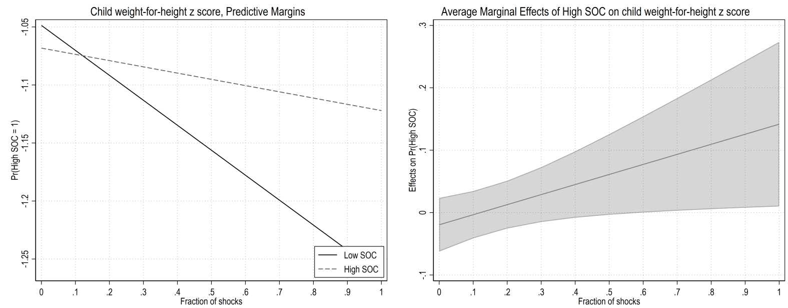


Fig. 3: Average Marginal Effects of high SOC on child weight-for-height z score

Table 2: Effect of monsoon rainfall shocks on child weight-for-height z score

	(1) Full	(2) Full	(3) Full	(4) Matched
Fraction of shocks	-0.222***	-0.219***	-0.234***	-0.215***
	(0.054)	(0.053)	(0.055)	(0.059)
High SOC (%)	-0.022	-0.027	-0.026	-0.020
	(0.018)	(0.018)	(0.018)	(0.022)
High SOC × Fraction of shocks	0.161***	0.182***	0.186***	0.161**
	(0.061)	(0.061)	(0.061)	(0.075)
AME of fraction of shocks	-0.101***	-0.083***	-0.094***	-0.134***
	(0.033)	(0.033)	(0.033)	(0.046)
AME of high SOC	-0.0001	-0.003	-0.001	0.005
	(0.016)	(0.016)	(0.016)	(0.019)
DHS controls	No	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes
Mean dependent variable		-0.991		-1.081
Average years of exposure		0.134		0.151
Observations	169,904	169,904	169,904	80,256
Adjusted R^2	0.067	0.086	0.086	0.069

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 soil organic carbon content is set at the threshold above 25 percentile. AME refers to the average marginal effect. 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: Average marginal effects of monsoon rainfall shocks on timing of child outcomes

	Height-for age z score	Weight-for-age z score	Weight-for-height z score
Rainfall shock - *in-utero*	0.053	0.062*	0.050
	(0.049)	(0.037)	(0.044)
Rainfall shock - year of birth	0.018	-0.044	-0.084
	(0.063)	(0.048)	(0.057)
Rainfall shock - 1st year	0.177**	0.064	-0.070
	(0.084)	(0.064)	(0.090)
Rainfall shock - 2nd year	0.105*	0.037	-0.048
	(0.054)	(0.041)	(0.047)
Rainfall shock - 3rd year	0.052	0.038	0.011
	(0.044)	(0.034)	(0.039)
Rainfall shock - 4th year	-0.020	0.032	0.070
	(0.044)	(0.032)	(0.039)
Observations	15,900	15,900	15,900

Levels of significance: $p< 0.01^{***}$, $p< 0.05^{**}$, $p< 0.1^{*}$. Robust standard errors are presented in parentheses. Appendix table A3 report the full results.

Table 4: Heterogeneity by higher and lower incidence region

	Weight-for-height z score					
	North	North central	Central plains	East	West	South
Fraction of shocks	-0.184	-0.048	-0.142	-1.087**	-0.239	-0.180
	(0.340)	(0.079)	(0.130)	(0.452)	(0.556)	(0.427)
High SOC (%)	-0.112	0.080**	-0.013	-0.397***	-0.109	-0.038
	(0.077)	(0.033)	(0.043)	(0.137)	(0.073)	(0.072)
High SOC × Fraction of shocks	-1.045**	-0.020	-0.002	1.069**	0.836	0.191
	(0.483)	(0.105)	(0.170)	(0.457)	(0.533)	(0.460)
AME of fraction of shocks	-0.380	-0.057	-0.142	-0.085	0.149	-0.043
	(0.330)	(0.062)	(0.113)	(0.168)	(0.364)	(0.198)
Mean dep. var.	-0.888	-1.069	-1.198	-1.046	-1.277	-1.021
Avg. years of exposure	0.094	0.231	0.093	0.134	0.055	0.090
Observations	4556	34,041	21,596	5597	5238	5458
Adjusted R^2	0.059	0.072	0.060	0.086	0.041	0.046

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 matched weights, *cem*, are applied in all regressions. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects. Hills and North East are limited by very small sample to proved meaningful estimates and hence excluded.

Table 5: Heterogeneous effects of rainfall shocks on child weight-for-height z-score by wealth index

	All	Poor	Non-poor
Fraction of shocks	-0.215***	-0.198**	-0.188**
	(0.059)	(0.080)	(0.084)
High SOC (%)	-0.020	0.033	-0.069**
	(0.022)	(0.031)	(0.030)
High SOC × Fraction of shocks	0.161**	0.123	0.185*
	(0.075)	(0.101)	(0.108)
AME of fraction of shocks	-0.134***	-0.132***	-0.101**
	(0.046)	(0.062)	(0.066)
Mean dep. var.	-1.081	-1.199	-0.971
Avg. years of exposure	0.151	0.167	0.136
Observations	80,256	38,645	41,588
Adjusted R^2	0.069	0.069	0.063

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 matched weights, *cem*, are applied in all regressions. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects.

Table 6: Heterogeneous effects of rainfall shocks on child weight-for-height z score by gender

	All	Female	Male
Fraction of shocks	-0.215***	-0.219***	-0.188**
	(0.059)	(0.078)	(0.082)
High SOC (%)	-0.020	-0.019	-0.017
	(0.022)	(0.029)	(0.029)
High SOC × Fraction of shocks	0.161**	0.159	0.160
	(0.075)	(0.102)	(0.104)
AME of fraction of shocks	-0.134***	-0.140***	-0.108*
	(0.046)	(0.062)	(0.064)
Mean dep. var.	-1.081	-1.052	-1.107
Avg. years of exposure	0.151	0.151	0.151
Observations	80,256	38,223	42,020
Adjusted R^2	0.069	0.069	0.070
p-val[Female=Male]		0.995	

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 matched weights, *cem*, are applied in all regressions. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects. The row 'p-val[Female=Male]' reports the p-value of the test of difference in the coefficient across female and male groups for the interaction terms between High SOC and fraction of shocks.

Table 7: Heterogeneous effects of rainfall shocks on female child weight-for-height z-score

	Female	Female sibling[1]	Male sibling[2]
Fraction of shocks	-0.219***	-0.320***	-0.210*
	(0.078)	(0.110)	(0.111)
High SOC (%)	-0.019	-0.037	-0.030
	(0.029)	(0.040)	(0.039)
High SOC × Fraction of shocks	0.159	0.237	0.219
	(0.102)	(0.145)	(0.146)
AME of fraction of shocks	-0.134***	-0.202***	-0.103
	(0.046)	(0.088)	(0.089)
Mean dep. var.	-1.081	-1.076	-1.076
Avg. years of exposure	0.151	0.152	0.152
Observations	38,223	20,495	20,293
Adjusted R^2	0.069	0.075	0.080
p-val[Female sibling=Male sibling]		0.893	

[1]Female child with at least one female sibling.

[2]Female child with at least one male sibling.

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 matched weights, *cem*, are applied in all regressions. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects. The row 'p-val[Female sibling=Male sibling]' reports the p-value of the test of difference in the coefficient across female sibling and male sibling groups for the interaction terms between High SOC and fraction of shocks.

Table 8: Heterogeneous effects of rainfall shocks on child weight-for-height z-score by selected climate zones

	Tropical wet and dry	Semi-arid	Humid sub-tropical
Fraction of shocks	-0.014	-0.269**	-0.069
	(0.370)	(0.124)	(0.076)
High SOC (%)	-0.035	-0.044	0.015
	(0.048)	(0.046)	(0.031)
High SOC × Fraction of shocks	0.034	0.147	0.031
	(0.375)	(0.264)	(0.097)
AME of fraction of shocks	0.012	-0.245**	-0.052
	(0.142)	(0.120)	(0.058)
Mean dep. var.	-1.236	-1.013	-1.057
Avg. years of exposure	0.061	0.127	0.216
Observations	17,500	18,802	38,123
Adjusted R^2	0.066	0.065	0.071

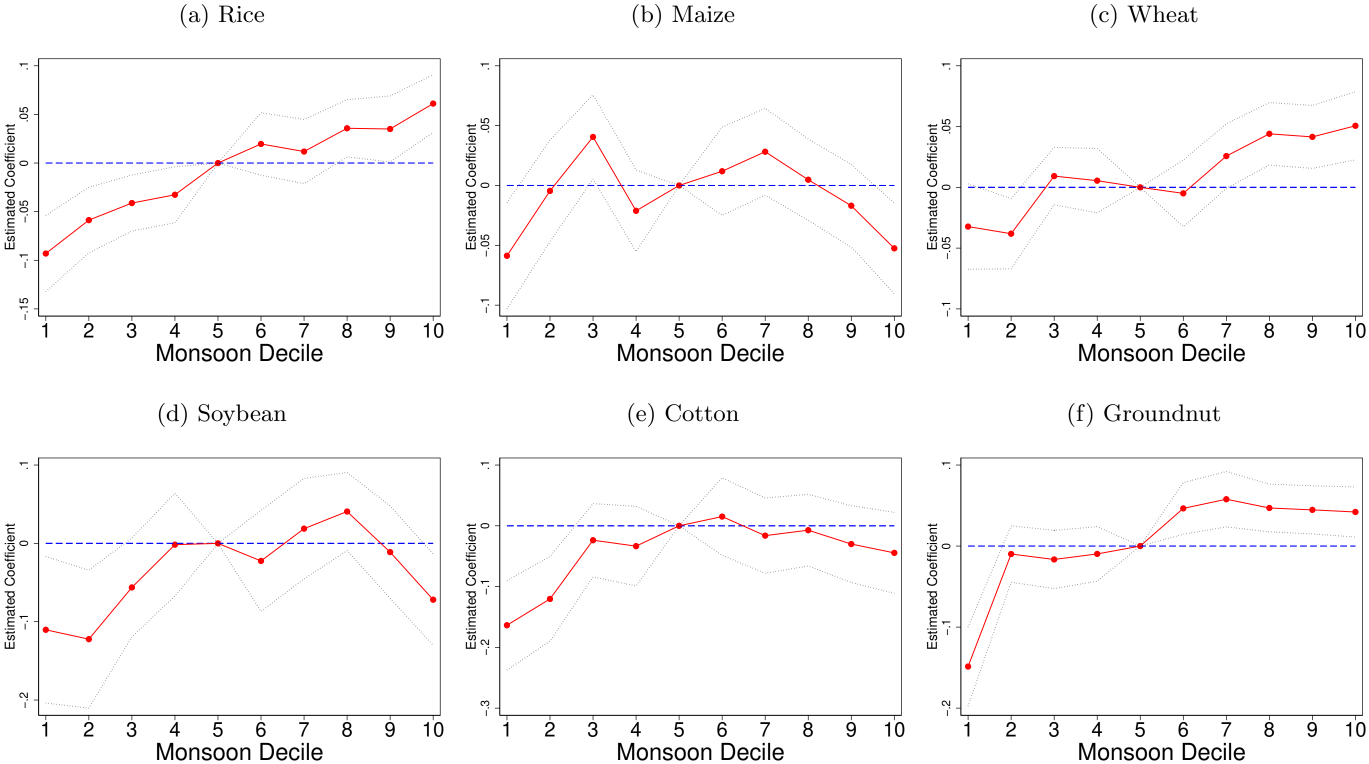
Levels of significance: p< 0.01***, p< 0.05**, p< 0.1*. Robust standard errors in parentheses are clustered at the DHS cluster level. The matched weights, *cem*, are applied in all regressions. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects. Tropical wet, Arid, and Mountain are limited by very small sample to provide meaningful estimates and hence excluded.

Table 9: Effect of rainfall shocks on likelihood of child stunting and wasting: Logit estimates.

	Full		Matched	
	Stunted	Wasted	Stunted	Wasted
Fraction of shocks	0.723***	1.483***	0.748***	1.532***
	(0.059)	(0.131)	(0.067)	(0.147)
High SOC (%)	0.984	1.008	0.970	1.001
	(0.028)	(0.032)	(0.034)	(0.038)
High SOC × Fraction of shocks	1.298***	0.764***	1.362***	0.736**
	(0.118)	(0.075)	(0.154)	(0.090)
AME of fraction of shocks	-0.028***	0.030***	-0.030***	0.045***
	(0.011)	(0.008)	(0.015)	(0.012)
Mean dep. var.	0.405	0.209	0.432	0.222
Avg. years of exposure	0.134	0.134	0.151	0.151
Observations	169,898	169,879	80,190	79,989

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. All regressions include demographic controls such as child, mother, and household level characteristics, and weather controls. All regressions include district and month-birth year specific fixed effects.

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).

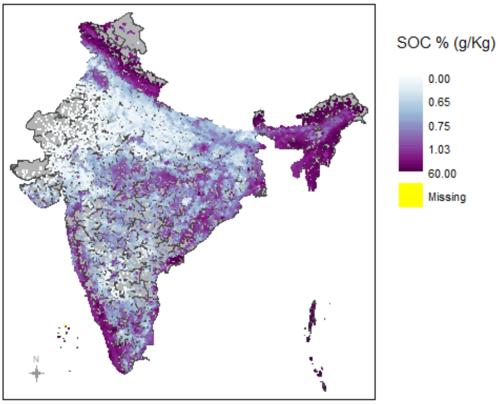


Fig. A2: Soil organic carbon content in rural areas. The missing in the map indicates the null values for union territory Lakshadweep.

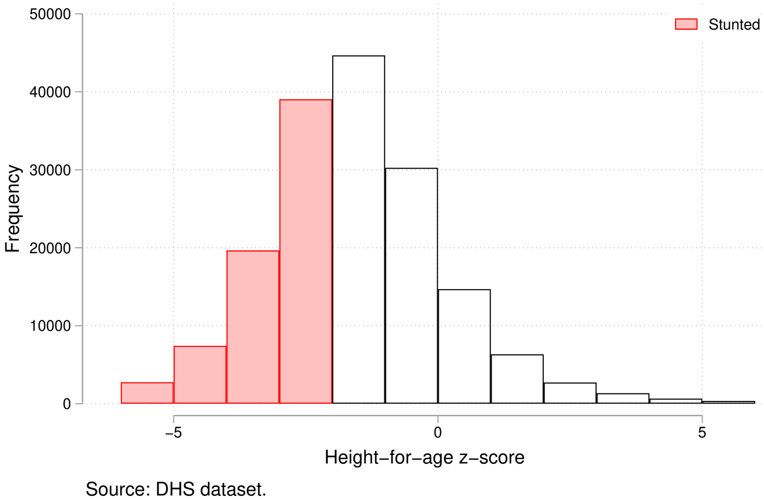


Fig. A3: Stunted (HAZ< -2)

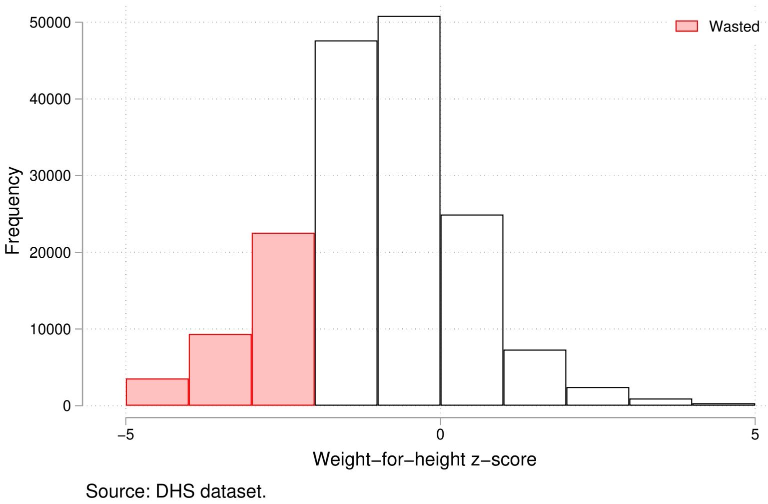


Fig. A4: Wasted (WHZ< -2)

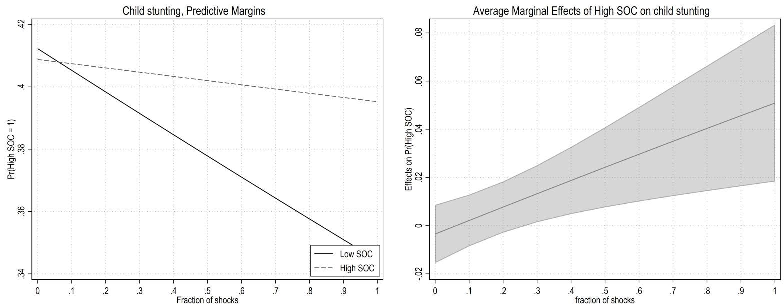


Fig. A5: Average Marginal Effects of high SOC on child stunting

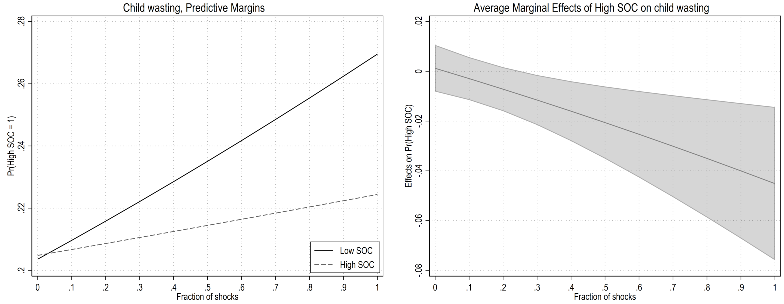


Fig. A6: Average Marginal Effects of high SOC on child wasting

Table A1: 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 A2: Alternative main regression results using population-weighted rain measures

	(1) Full	(2) Full	(3) Full	(4) Matched
Fraction of shocks	-0.208***	-0.208***	-0.212***	-0.196***
	(0.054)	(0.054)	(0.054)	(0.059)
High SOC (%)	-0.018	-0.024	-0.023	-0.017
	(0.018)	(0.018)	(0.018)	(0.022)
High SOC × Fraction of shocks	0.138**	0.160***	0.162***	0.136*
	(0.061)	(0.061)	(0.061)	(0.076)
AME of fraction of shocks	-0.104***	-0.088***	-0.091***	-0.128***
	(0.033)	(0.032)	(0.033)	(0.045)
Mean dependent variable		-0.991		-1.081
Average years of exposure		0.134		0.150
DHS controls	No	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes
Observations	169,904	169,904	169,904	80,256
Adjusted R^2	0.067	0.086	0.086	0.069

Levels of significance: $p< 0.01^{***}$, $p< 0.05^{**}$, $p< 0.1^{*}$. Robust standard errors in parentheses are clustered at the DHS cluster level. 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.

Table A3: Effect of monsoon rainfall shocks on timing of child outcomes

	Height-for age z score	Weight-for-age z score	Weight-for-height z score
Rainfall shock - *in-utero*	0.106	0.105**	0.067
	(0.070)	(0.052)	(0.061)
Rainfall shock - year of birth	0.088	-0.063	-0.179**
	(0.084)	(0.064)	(0.078)
Rainfall shock - 1st year	0.235**	0.091	-0.083
	(0.119)	(0.091)	(0.130)
Rainfall shock - 2nd year	0.080	0.017	-0.055
	(0.072)	(0.054)	(0.061)
Rainfall shock - 3rd year	0.119**	0.082**	0.015
	(0.055)	(0.041)	(0.047)
Rainfall shock - 4th year	0.012	0.099**	0.146***
	(0.053)	(0.038)	(0.048)
High SOC (%)	0.049	0.051	0.035
	(0.051)	(0.039)	(0.047)
High SOC × shock - in-utero	-0.106	-0.085	-0.034
	(0.077)	(0.058)	(0.067)
High SOC × shock - year of birth	-0.140	0.037	0.189**
	(0.102)	(0.079)	(0.095)
High SOC × shock - 1st year	-0.115	-0.054	0.026
	(0.131)	(0.104)	(0.137)
High SOC × shock - 2nd year	0.049	0.039	0.013
	(0.086)	(0.064)	(0.074)
High SOC × shock - 3rd year	-0.135*	-0.087	-0.008
	(0.070)	(0.053)	(0.061)
High SOC × shock - 4th year	-0.064	-0.135***	-0.152**
	(0.068)	(0.052)	(0.063)
DHS controls	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
Mean dependent variable	-1.660	-1.710	-1.081
Observations	15,900	15,900	15,900
Adjusted R^2	0.105	0.135	0.091

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 matched weights, *cem*, are applied in all regressions. The samples only include children of age 4 years to measure all (in-utero through age 4) the cumulative rainfall shock. 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.

Table A4: Sensitivity test for different threshold level: High soil organic carbon above 50th percentile

	(1) Full	(2) Full	(3) Full	(4) Matched
Fraction of shocks	-0.174***	-0.149***	-0.161***	-0.072
	(0.041)	(0.041)	(0.042)	(0.052)
High SOC (%)	-0.030*	-0.023	-0.023	-0.009
	(0.016)	(0.015)	(0.015)	(0.018)
High SOC × Fraction of shocks	0.148**	0.132**	0.136**	0.032
	(0.059)	(0.058)	(0.059)	(0.071)
Marginal effects	-0.100***	-0.083***	-0.094***	-0.056
	(0.033)	(0.033)	(0.034)	(0.041)
Mean dependent variable		-0.991		-1.065
Average years of exposure		0.134		0.130
DHS controls	No	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes
Observations	169,904	169,904	169,904	102,294
Adjusted R^2	0.067	0.086	0.086	0.072

Levels of significance: $p< 0.01^{***}$, $p< 0.05^{**}$, $p< 0.1^{*}$. Robust standard errors in parentheses are clustered at the DHS cluster level. 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: 51,148 matched out of 84,950 observations for control and 51,148 matched out of 84,954 for treated.

Table A5: Sensitivity test for different threshold level: High soil organic carbon above 75th percentile

	(1) Full	(2) Full	(3) Full	(4) Matched
Fraction of shocks	-0.129***	-0.105***	-0.114***	-0.128
	(0.036)	(0.036)	(0.037)	(0.081)
High SOC (%)	-0.045*	-0.017	-0.022	-0.002
	(0.024)	(0.023)	(0.023)	(0.029)
High SOC × Fraction of shocks	0.099	0.069	0.066	0.002
	(0.072)	(0.071)	(0.072)	(0.106)
Marginal effects	-0.104***	-0.087***	-0.097***	-0.127***
	(0.033)	(0.033)	(0.034)	(0.064)
Mean dependent variable		-0.991		-0.922
Average years of exposure		0.134		0.117
DHS controls	No	Yes	Yes	Yes
Weather controls	No	No	Yes	Yes
Observations	169,904	169,904	169,904	45,497
Adjusted R^2	0.067	0.086	0.086	0.082

Levels of significance: $p< 0.01^{***}$, $p< 0.05^{**}$, $p< 0.1^{*}$. Robust standard errors in parentheses are clustered at the DHS cluster level. 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: 22,749 matched out of 127,425 observations for control and 22,749 matched out of 42,479 for treated.

Table A6: 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. 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 A7: 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
*age*510	Dummy for child with sibling between the age 5 and 10 years
*age*1115	Dummy for child with sibling between the age 11 and 15 years
*age*16	Dummy for child with sibling above 16 years
*hw*1	Child's age in months
Woman-specific	
*v*012	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	
*v*104	Years lived in place of residence
*hv*220	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
*hv*244	Dummy for household owns agricultural land
hhirragland	Dummy for household irrigate agricultural land
*sh*52*a*	Dummy for household owns cows/bulls/buffaloes
*sh*52*b*	Dummy for household owns camels
*sh*52*c*	Dummy for household owns horses/donkeys/mules
*sh*52*d*	Dummy for household owns goats
*sh*52*e*	Dummy for household owns sheep
*sh*52*f*	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, *hw*1 was transformed with restricted cubic spline and knots are placed at the interval of 6, 12, 18, 24, 36, and 48.

Table A8: 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 A9: Robustness check: confounding variables included as controls

	(1) Full	(2) Matched
Fraction of shocks	-0.221***	-0.188***
	(0.053)	(0.057)
High SOC (%)	-0.026	-0.015
	(0.018)	(0.022)
High SOC × Fraction of shocks	0.185***	0.148**
	(0.061)	(0.075)
DHS controls	Yes	Yes
Other controls	Yes	Yes
Observations	169,897	80,254
Adjusted R^2	0.086	0.069

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 dependent variable is child weight-for-height z score. The high soil organic carbon content is set at the threshold above 25 percentile. DHS controls include child, mother, and household level characteristics. Other controls include confounding variables such as soil texture, slope, and vegetation.