ECONOMETRIC ANALYSIS

CGIAR Climate Security Observatory

METHODS PAPERS SERIES

03/2023

CSO Actionable Evidence

CGIAR FOCUS Climate Security

CGIAR INITIATIVE ON Climate Resilience

PURPOSE

ThThe latest IPCC report indicates food insecurity and nutrition among the most plausible mechanisms linking climate variability and conflict (IPCC, 2022). However, quantitative research has yet to provide evidence on the relationship between climate and conflict through food and nutrition insecurity (see, for instance, Buhaug et al., 2015). This analysis aims at filling this gap in the literature, and carries out a mediation analysis to test whether and how climate variability and conflict risk are indirectly correlated through highly localized food and nutrition insecurity dynamics.

In doing so, this approach makes three important contributions to the existing econometric literature. First, it moves beyond approaches that assume unique transmission channels when investigating the climate-conflict nexus. Instead, it employs a Structural Equation Model that allows for the existence of a multiplicity of mediating factors to get more accurate estimates of indirect impacts. Second, this approach stands out as one of the first of its kind to employ food and nutritional indicators to identify the indirect contribute of climate shocks to conflicts through the food and nutrition security pathway. Third, this analysis moves beyond standard country-year approaches by employing quarterly socioeconomic survey data combined with climate and conflict information at a fine subnational level to better capture the local and intertemporal dynamics scales.

The data used and the approach developed, as described below, are consistent across the countries selected for the ClimBeR initiative, with minor variations due to data and context specificities.

.

DATA

The econometric analysis relies on three major sources of data that respectively contain conflict, climate and socio-economic information. In this section we describe each of those.

Conflict data

Geo-localized event-based data on conflict events are extracted from the Armed Conflict Location and Event Dataset (ACLED) - or the Uppsala Conflict Data Program (UCDP). When using ACLED, this analysis focuses on two types of conflicts: (i) violent conflicts (including battles, violence against civilians, and remote violence); and (ii) civil unrest (including protests, riots, and strategic developments). While civil unrest assesses the grievances channel of the climate-security nexus, violent conflicts should be able to capture state and non-state communal, ethnic, clan, and religious conflicts that might be linked to resource-based conflict outcomes. When using UCDP, this analysis is primarily focused on one-sided, non-state, and state-based violence, where each episode resulted in at least 25 battle-related deaths. In a few cases, self-reported information on perceptions of insecurity and episodes of violence are employed instead of using ACLED or UCDP data. This is only possible with datasets like the Afrobarometer or the Latinobarometro.

Climate data

Climate data are extracted from TerraClimate, a high-spatial-resolution dataset (1/24°4km). This dataset provides information on maximum temperature, precipitation accumulation, climate water deficit, Palmer Drought Severity Index (PDSI), soil moisture and reference evapotranspiration, from January 1958 to December 2021. These monthly climate data are mainly used to create climate anomalies, namely variables that account for spatial and temporal deviations from the long-run means (Maystadt et al., 2014). The anomalies are computed as the difference between the monthly climate data and the long-term monthly mean divided by the long-run standard deviation (Helman et al., 2020; Maystadt et al., 2014).

Socio-economic data
Socio-economic data come from Demographic and Health Surveys (DHSs), Afrobarometer and Latinobarometro surveys, which provide information on food and nutrition security, minimum acceptable dietary conditions, regular food and water access, general living conditions, and malnutrition (stunting, wasting, and underweight) – depending on the data source. The surveys also provide a wide set of information on households' living conditions and social status, such as employment, wealth index, rurality, ethnicity, access to resources, livestock, and agricultural land ownership. Furthermore, we employ commodity price data coming from the World Food Programme Price Database.[1]

DATABASE CONSTRUCTION

The dataset's construction entails several steps. First, using the geographical coordinates of the DHS clusters[2], we create buffers to merge household survey data with georeferenced quarterly conflict information registered during and after DHS interviews. Conflict buffers cover approximately a 50-kilometer-radius from the DHS cluster locations and provide a relatively accurate proxy for conflict incidence by accounting for any quarterly event in the proximity of the clusters, including cross-border episodes (Figure 1). Second, we integrate monthly climate data and price data with the survey data at the DHS cluster level. Finally, we rasterize the combined dataset into 20 squared km grid-cells and create an unbalanced panel at grid-cell level covering the study period.

[1] Commodity prices are included when DHS is the socio-economic source of data, considering the limited amount of controls available as well as the nature of the dependent variables selected (usually malnutrition variables). For Afrobarometer and Latinobarometro, prices are not added.
[2] A cluster is defined as a number of households that participated in the DHS survey and are scattered over one or more inhabited areas. To maintain confidentiality, the locations of the DHS clusters are randomly shifted, up to 5 km for rural locations, and up to 2 km for urban locations. This approach is applicable only when coordinates at the household or cluster level are available.

CSO
Actionable Evidence

The way the dataset is constructed overcomes several limitations of the DHS data. It allows us for conducting a battery of important econometric exercises at a much granular resolution. Moreover, the buffer zone naturally accounts for spatial clustering of conflicts and cross-border conflicts (shown in Figure 1), while reducing measurement error stemming from the imprecise location of DHS clusters (Wagner et al., 2018).

In the final dataset, the conflict variables are computed as the average number of conflicts occurred in the buffers of the households' location in the grid.[3] The food and nutrition variables are expressed as shares or percentages of food insecure or malnourished households in the grid. Similarly, socio-economic indicators are computed as shares or percentages at grid level. The commodity prices are computed as the average price of the closest market to households' location. Lastly, climate variables are averaged at the grid level to identify climate anomalies.

[3] In the Afrobarometer, perceptions of insecurity and episodes of violence are captured as households' shares at the grid level.

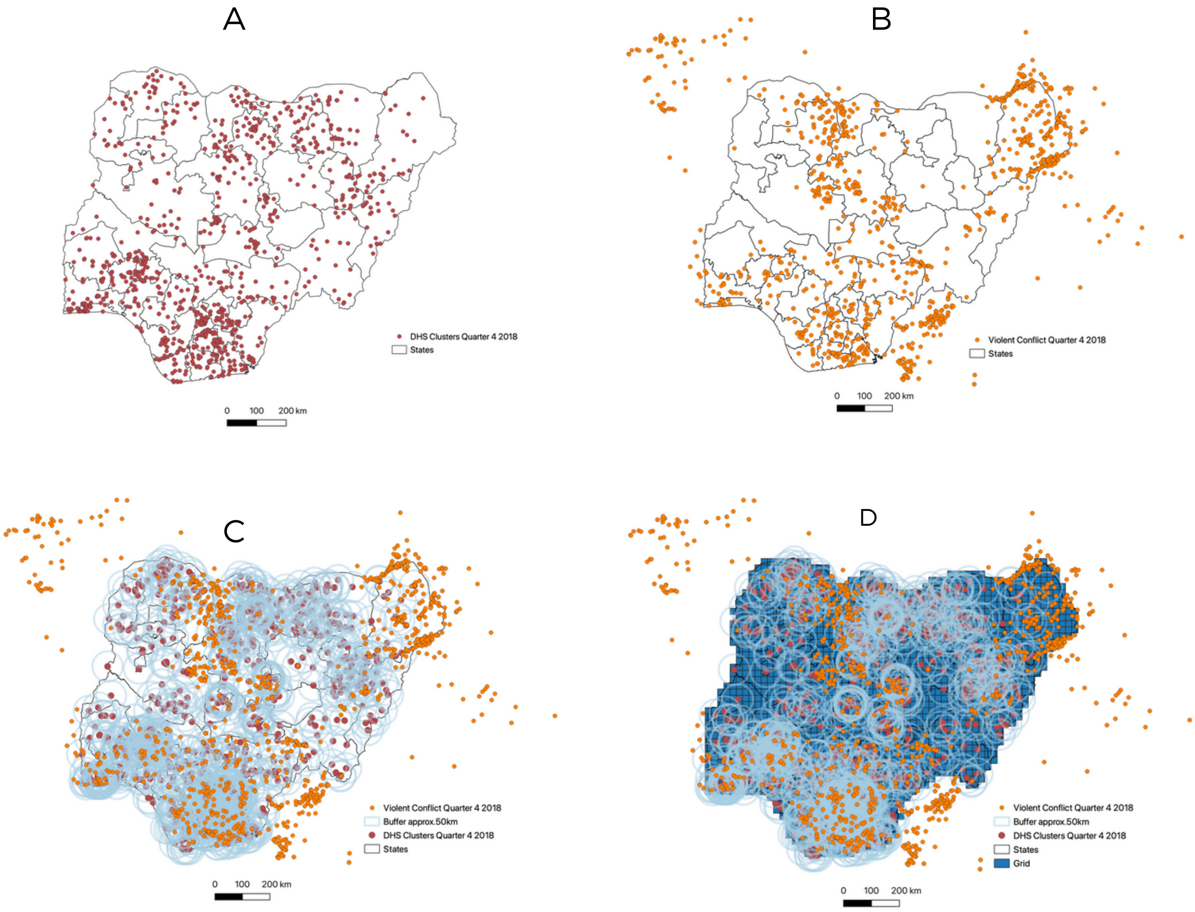

FIGURE 1. PROCEDURE OF DATABASE CONSTRUCTION FOR THE CASE OF NIGERIA. THE MAPS SHOW: (A) DHS CLUSTER COORDINATES, (B) VIOLENT CONFLICT COORDINATES, (C) BUFFERS THAT ASSOCIATE CONFLICT TO HOUSEHOLD CLUSTERS, (D) GRIDS.

METHODS

This analysis investigates the role of food insecurity and malnutrition as a mediator that channels changing climate conditions into more frequent conflicts. However, even if we expect food insecurity and malnutrition to be key channels, they are unlikely to be the only ones. For instance, other factors such as price changes, migratory patterns, or ethnicity, can play a role that disregards malnutrition.

We, therefore, employ a Structural Equation Model (SEM) that explicitly identifies: (i) the indirect effect that links climate to conflict through malnutrition (A*B in Figure 2), and (ii) the residual effect that links climate to conflict through any other channel that is not correlated with nutrition insecurity (C in Figure 2). To test these effects, we follow Baron & Kenny (1986) and estimate a system of three [4] simultaneous equations, as follows:

$$\log\left(C_{i,[q+1,q+3]}\right) = \gamma_0 + \gamma_1 TA_{i,t} + \gamma_n \boldsymbol{X}_{i,q} + \phi_s + \psi_{s,q} + \varpi_q + \delta_i + \varepsilon_{s,i,t,q} \quad (1)$$

$$M_{i,q} = \beta_0 + \beta_1 TA_{i,t} + \beta_n \boldsymbol{X}_{i,q} + \phi_s + \psi_{s,q} + \varpi_q + \delta_i + \varepsilon_{s,i,t,q} \quad (2)$$

$$\log\left(C_{i,[q+1,q+3]}\right) = \mu_0 + \mu_1 TA_{i,t_{-1}} + \mu_2 M_{i,q} + \mu_n \boldsymbol{X}_{i,q} + \phi_s + \psi_{s,q} + \varpi_q + \delta_i + \varepsilon_{i,t,q} \quad (3)$$

where $\left(C_{i,[q+1,q+3]}\right)$ s the average number of violent conflicts at the grid i and quarter-year q evel that occurred in Nigeria one-to-three quarters $(q+1; q+3)$ after the DHS interviews; $TA_{i,t}$ are the climate anomalies averaged over the ***n*** months before the DHS interviews (included), where ***n*** varies from 3 to 12 months; $M_{i,q}$ is the food security and nutrition mediator, mainly captured by the prevalence of stunting, wasting or underweight children in the households, or the average commodity prices to proxy for food insecurity;

[4] The first equation (1) is not reflected in the empirical SEM specification, being incorporated in the third one (3) due to SEM features in Stata.

CSO
Actionable Evidence

$X_{i,q}$ s a set of control variables; ϕ_s and ϖ_q are the location (state) and time (quarter-year) fixed effects, to respectively control for unobserved time-invariant state heterogeneity, as well as temporal changes, and common shocks in a given quarter; δ_i are the random effects at grid level employed to capture systematic unobserved heterogeneity at grid level through a random intercept term. The choice of a combination of state-quarter fixed effect and grid-level random intercept is likely to capture most of the unobserved heterogeneity while avoiding the estimation of large standard errors that are anticipated in a grid-level-quarter fixed effects framework, where grid fixed-effects are likely to absorb almost all the variation in the weather variables (Fisher et al., 2012). Lastly, the error term is $\varepsilon_{i,t,q}$,clustered at grid level.

In the above system of equations, the indirect effect is captured by $\beta_1 * \mu_2$ and the residual effect is given by μ_1 The rational of the model is that the mediator: i) varies significantly if the treatment variable (climate anomalies) change consistently; ii) affects the variation of the dependent variable (conflict) if the mediator itself significantly varies; and iii) indicates the presence of a strong mediation if the residual effect is zero (Baron & Kenny, 1986). However, we do not necessarily expect a null residual effect because we recognize that climate can influence conflict through a variety of channels. Building on this rational, the model relies on three major assumptions: (i) the treatment is exogenous to the dependent variable; (ii) the treatment is exogenous to the mediator; and (iii) the mediator is exogenous to the dependent, conditional on the treatment. While the first two assumptions are easily met, the third one may be difficult to be fully respected, as there may be other mechanisms that, as a consequence of the climate stressors, could influence both the mediator and conflict variable. We relax these concerns by controlling for potential observed and unobserved confounders. The use of a combination of state-quarter fixed effect and grid-level random effects allows to control for the state and time invariant unobserved heterogeneity as well as the grid-level systematic unobserved factors. In addition, we include a number of grid-level characteristics.

Among them, we control for the presence of ongoing violent conflicts, to relax concerns about reverse causality, and to control for the potential spatial autocorrelation and path dependency traditionally characterizing the conflicts. Other controls frequently employed include the commodity prices, female employment and the presence of rural households. Nonetheless, the inclusion of such a comprehensive set of controls does not allow to fully exclude the existence of other unobserved (time-state variant and/or intra-grid subgroup specific) characteristics that may confound the estimation of the indirect association.

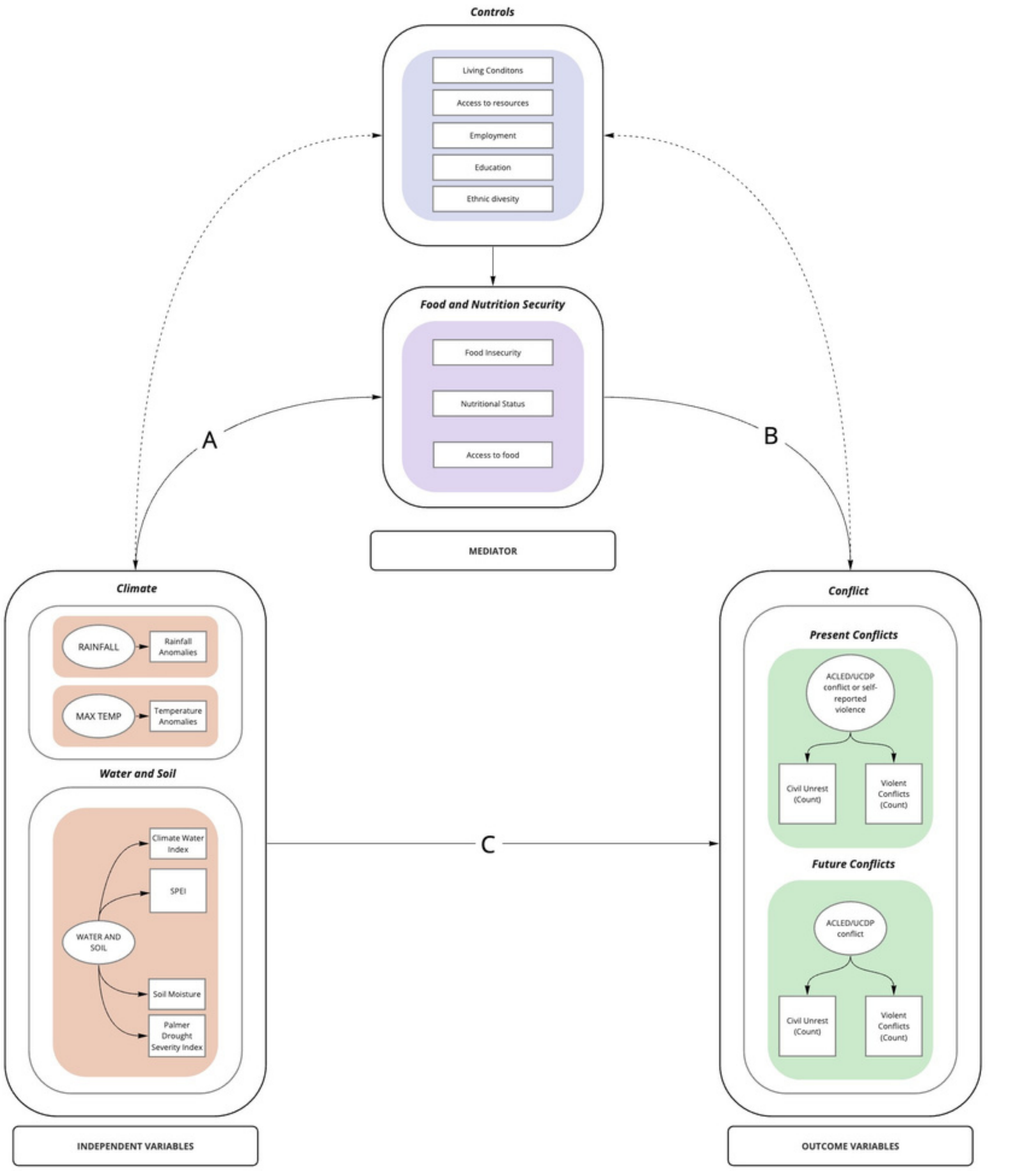

FIGURE 2. MEDIATION ANALYSIS APPROACH.

CSO
Actionable Evidence

REFERENCES

1. Baron R.M., Kenny D.A., (1986), The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. Journal of Personality and Social Psychology, Vol.51, No.6, 1173-1182
2. Helman, D., Zaitchik B. F., Funk. C., (2020), Climate Has Contrasting Direct and Indirect Effects on Armed Conflicts. Environmental Research Letters 15(10):104017.
3. Maystadt, J.-F., Calderone M., You L., (2015), Local warming and violent conflict in North and South Sudan. Journal of Economic Geography 15 (2015) pp. 649–671.
4. Maystadt, J.-F., Ecker, O., (2014), Extreme weather and civil war: does drought fuel conflict in Somalia through livestock price shocks? American Journal of Agricultural Economics, 96: 1157–1181.

Suggested citation

Belli, A., Mastrorillo, M., and Villa, V. Econometric Analysis. Climate Security Observatory Methods papers series (2023)