WORLD BANK GROUP
Poverty & Equity

POVERTY & EQUITY NOTES

JULY 2020 · NUMBER 26

Adaptive Safety Nets for Rural Africa: Drought-Sensitive Targeting with Sparse Data[1]

Javier E. Baez, Varun Kshirsagar, Emmanuel Skoufias

This paper combines remote-sensed data and individual child, mother, and household-level data from the Demographic and Health Surveys for 5 countries in Sub-Saharan Africa to design a prototype drought-contingent targeting framework for use in scarce-data contexts. To accomplish this, the paper: (i) develops simple and easy-to-communicate measures of drought shocks; (ii) shows that droughts have a large impact on child stunting in these five countries—comparable, in size, to the effects of mother's illiteracy or a fall to a lower wealth quintile; and (iii) shows that, in this context, decision trees and regressions predict stunting as accurately as complex machine learning methods that are not interpretable.[2] Taken together, the analysis lends support to the idea that a data-driven approach may contribute to the design of a transparent and easy-to-use drought-contingent targeting framework.

Households in rural Sub-Saharan Africa (SSA), with the largest concentration of poor and deprived households in the world, are particularly vulnerable to weather events.[3] These events are projected to increase in frequency and become more severe. Transparent data-driven approaches are needed to both predict adverse outcomes of weather events such as droughts and to identify the most vulnerable populations to target with assistance when these outcomes occur. While machine learning applications for early warning frameworks hold considerable promise ,they tend to be difficult to understand (that is, "not interpretable"), which serves as an impediment to their use in designing policies to mitigate the impact of shocks on the poor and vulnerable.

Our contribution here is to design a transparent early warning framework to identify deprived children in areas that are both data-constrained and vulnerable to climate change. **Such frameworks** may help achieve the first two Sustainable Development Goals (SDGs)—ending poverty and eliminating hunger across the world. Further, a transparent and data-driven early warning and monitoring framework may engender the requisite budget flexibility and policy-maker responsiveness, and complement expert panels, such as those that set Integrated Food Security Phase Classification levels—the current global standard. While USAID's FEWS NET and the UN's FAO and WFP have achieved remarkable progress in making sophisticated remote-sensed detection and analyses of weather anomalies publicly accessible, standard approaches to estimating the human impacts of these anomalies are less mature.

We tested a relatively simple drought-contingent framework for 5 SSA countries—Malawi,

[1] This note summarizes the full report, Adaptive Safety Nets for Rural Africa: Drought-Sensitive Targeting with Sparse Data. See the full report for the list of references, detailed results and in more in-depth analysis.

[2] (Supervised) Machine-learning models predict outcomes for which data is unavailable by identifying relationships in existing data sets that contain information on these outcomes. Interpretability is the ability to explain to a human being how a machine learning algorithm came to its decision. Most machine learning algorithms cannot be interpreted in a transparent manner (e.g. https://hdsr.mitpress.mit.edu/pub/f9kuryi8/release/5) .

[3] See the full report for references to related studies.

Mozambique, Tanzania, Zambia and Zimbabwe—to identify children at risk of being stunted. The report summarized in this brief makes four contributions that may inform policies designed to mitigate climate change impacts on the most vulnerable:

- Demonstrate use of a prototype, drought-contingent targeting system for data-scarce contexts.
- Assist programs to reach beneficiaries based on non-monetary measures of child well-being.
- Use remote-sensed data to automate categorization of both harvest cycles and drought shocks.[4]
- Show that interpretable (white box) methods, such as logistic regressions and decision trees, are comparable, in terms of predictive accuracy to opaque machine learning methods.

Methodology and Data

Our work attempts to address current data limitations by combining existing Demographic and Health Survey (DHS)[5] data with remote-sensed information. Our focus here is on designing a prototype drought-contingent targeting system, rather than providing a systematic analysis of the causes of, and solutions to, stunting.

Child stunting: Objective measures of hunger rely on caloric intake, which may not be strongly correlated with nutrition. We follow the more standard approach of using anthropometric measures[6] (child height, weight) as indicators of child malnutrition.[7]

Crop production: While there is no consensus on the best approach to monitoring crop production, there is broad agreement on challenges and sources of information. Monitoring crop production in SSA relies on remote-sensed information, but averaged at broad spatial scales—traditionally focusing on vegetation and rainfall. Unlike rainfall, the Normalized Difference Vegetation Index (NDVI) is a cumulative measure. With available ground data, remote-sensed vegetation anomalies can be calibrated to historic crop yields to build accurate models. However, when ground data are scarce, vegetation indices cannot be properly calibrated with crop-productivity. Rainfall, during the growing season, is unambiguously correlated with crop productivity; however, the influence of rainfall on yields depends on timing and distribution, and aggregating rainfall estimates over a growing season can be misleading.

Estimation of Harvest Cycles: Given these challenges, we employ the following strategy. First, we focus on shocks at a country's first administrative level. Second, we use NDVI—averaged at the province x month level—to determine harvest cycles for each province (i.e. the first administrative level for each country). Third, we use rainfall estimates, averaged for each province, to estimate sequences of dry spells. NDVI data we used were collected by the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor aboard the NASA polar orbiting satellite Terra from February 2000 to April 2019. MODIS provided an NDVI estimate every 8 days at an approximately 250-meter ground sample (pixel) resolution across the entire study area. The values were spatially and temporally averaged to each province x month. We identify harvest cycles—more specifically growing months, the month that a harvest starts, and the lean season—using a simple algorithm: (a) calculate the percentage change in monthly NDVI. A month is labeled as a growing month if the percent increase is greater than 20 percent; (b) the start of the harvest season is defined as the first month in which the NDVI falls by at least 5 percent; (c) the lean season is defined as the 3 months preceding start of the harvest season.

[4] Our measure of drought shocks—two to four successive below average10-day rainfall spells during the growing season — while not typically used by economists, is standard in the agronomy literature.
[5] DHS data are used to estimate the World Health Organization (WHO) measures of malnourishment (WHO Multicentre Growth Reference Study Group and others, 2006).
[6] Anthropometrics refers to the measurement of the human individual.
[7] But cautions that the irregularity of the DHS surveys makes it difficult to use them for monitoring, for example for assessing the effects of a food price crisis on the heights and weights of children.

Identifying Dry Spells: We develop a simple approach to measuring droughts. Our rainfall measure comes from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset and the algorithm used involves the following steps: (a) calculate the average and standard deviations of total rainfall (in mm) for every 10-day period or decad in a given month in a province (i.e. decad x month x province); (b) calculate the standardized anomalies for every decad x month x province x year; (c) exclude all decads that are not in the growing season; (d) define a dry spell of length n for a given decad, as having below average rainfall in that decad, as well as the preceding decads. In addition, at least half the decads in a dry spell are required to receive less than -0.5 standard deviations rainfall; (e) define a harvest as being affected by drought if the preceding growing season contains at least one dry spell.[8]

We calculate the incidence of dry spells and droughts for the 5 countries. When defining a drought as occurring during a growing season with a 3-decad dry spell, 3 of the countries (Tanzania, Mozambique and Zimbabwe) experienced droughts 22 percent of the time (once every five years) between 2001 and 2018, while Zambia experienced droughts 32 percent of the time, and Malawi experienced droughts 45 percent of the time. Further, Mozambique and Malawi experienced greater incidences of droughts in the last decade in comparison to the previous one.

We estimate that the incidence of droughts increases the likelihood of a child being stunted by 5 to 10 percentage points. In base case analysis, we find that across the 5 countries, a poor harvest season at the time of a child's birth (defined here as at least one 3-decad dry spell during the preceding growing season) is associated with a 7 percentage-point increase in stunting (see Figure 1). Since 41 percent of children under 5 in this rural population are stunted, on average, stunting rates are approximately 16 percent higher for children born during a poor harvest year. If the poor harvest occurs in the year after a child's birth, it is associated with a 4 percentage-point increase in stunting. The effects of a drought on child weight are also large. On average, 15.1 percent of children in this population are underweight. A drought causes a 3 percentage-point, or 20 percent increase in the likelihood of a child being underweight.

To demonstrate the applicability of our approach, we also develop decision trees – that involve simple decision rules – and show that they are as accurate as complex machine learning methods.[9]

Conclusions

Taken together, our analysis supports the idea that a simple data-driven approach may contribute to policies that mitigate climate change impacts on the world's most vulnerable populations. We have constructed a prototype drought-contingent targeting framework to inform the design of social safety nets in data-scarce contexts. We addressed several concerns related to other predictive and beneficiary targeting approaches. First, we developed measures of drought shocks that are simple, transparent and easy-to-communicate. Second, we demonstrated that droughts have a large impact on stunting in Southern Africa. Third, we show that interpretable methods such as decision trees and logistic regressions predict stunting as accurately as "black-box" machine learning methods.

Next Steps

We have focused on one particular spatial scale — a first administrative level (or province or region). Weather disturbances occur at different spatial scales, including at the household or village, regional, or national levels. A better understanding of the impacts of shocks at different spatial scales requires knowledge of how food markets operate. Weather disturbances in areas that are connected to national and international markets may have a much more muted impact compared to disturbances that occur in

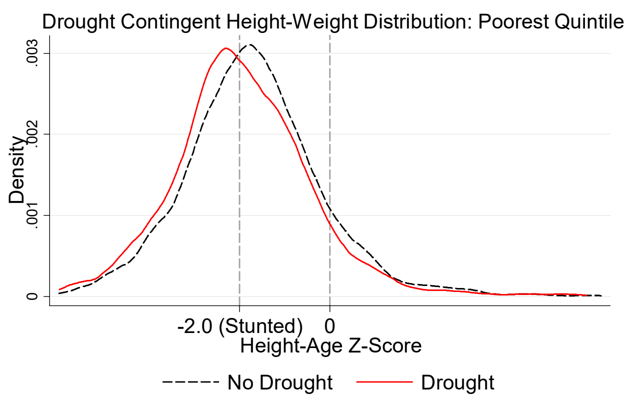
[8] A dry spell is defined as below average rainfall (for that province x decad) for every decad in the sequence, with at least half the decads having less than -0.5 standard deviations of rainfall.

[9] See full report, *Adaptive Safety Nets for Rural Africa: Drought-Sensitive Targeting with Sparse Data*, for full results and in more in-depth analysis.

remote rural areas. Further, weather disturbances in food-surplus areas may engender food scarcity in other parts of a country.

While the data we have used in this study are freely available, collecting these data requires substantial public investment and capacity. Our overarching objective is to leverage nationally representative, comparable and publicly available data to construct frameworks that are transparent and straight-forward to communicate. We believe that these are essential prerequisites to designing shock-sensitive government systems that are both credible and sustainable.

Figure 1: Drought and Stunting for the Poorest



ABOUT THE AUTHORS

Javier E. Baez is Senior Economist at the Poverty and Equity Global Practice and Global Lead of the Global Solutions Group "Welfare Implications of Risk, Fragility and Conflict" of the World Bank.

Varun Kshirsagar an independent consultant working of food security and early warning systems.

Emmanuel Skoufias is a Lead Economist at the Poverty and Equity Global Practice of the World Bank.