Policy Research Working Paper 9071

Adaptive Safety Nets for Rural Africa

Drought-Sensitive Targeting with Sparse Data

Javier E. Baez
Varun Kshirsagar
Emmanuel Skoufias

WORLD BANK GROUP
Poverty and Equity Global Practice
December 2019

Policy Research Working Paper 9071

Abstract

This paper combines remote-sensed data and individual child-, mother-, and household-level data from the Demographic and Health Surveys for five countries in Sub-Saharan Africa (Malawi, Tanzania, Mozambique, Zambia, and Zimbabwe) to design a prototype drought-contingent targeting framework that may be used 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 and a fall to a lower wealth quintile; and (iii) shows that, in this context, decision trees and logistic regressions predict stunting as accurately (out-of-sample) as machine learning methods that are not interpretable. Taken together, the analysis lends support to the idea that a data-driven approach may contribute to the design of policies that mitigate the impact of climate change on the world's most vulnerable populations.

This paper is a product of the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at eskoufias@worldbank.org.

Adaptive Safety Nets for Rural Africa: Drought-Sensitive Targeting with Sparse Data*

Javier E. Baez[†] **Varun Kshirsagar**[‡] **Emmanuel Skoufias**[§]

JEL CLASSIFICATION: I32, Q54, N57

KEY WORDS: Poverty, Child Welfare, Climate Change, Poverty Targeting, Social Protection, Child Malnourishment, Stunting

*This paper was funded by the World Bank's Poverty and Equity Global Practice under its global knowledge program "Welfare Implications of Climate Change, Fragility and Conflict Risks". We would like to thank John Baffes, Molly Brown, Frank Davenport, Gary Eilerts, Chris Funk, Madhur Gautam, Sharada Ramanathan, and especially David M. Johnson for invaluable conversations. In addition, we benefited from comments from participants at presentations at the World Bank and the 2019 American Agricultural Economics Association Meetings.

[†]Senior Economist at the World Bank. Email: jbaez@worldbank.org

[‡]Independent. Email: varun.kshirsagar@gmail.com

[§]Lead Economist at the World Bank. Email: eskoufias@worldbank.org

1 Introduction

Rural Sub-Saharan Africa has the largest concentration of poor and deprived households in the world (World Bank (2018), Skoufias et al. (2019)). These households are particularly vulnerable to anomalous weather events.[1] These events are projected to increase in frequency and become more severe in a warmer world. However, policy responses to these events are often slow and large funding gaps are the norm. A transparent and data-driven early warning and monitoring framework may engender the requisite budget flexibility and responsiveness.[2] Yet, survey data that would inform such a framework are collected infrequently and at a considerable cost (Beegle et al. (2016), Castañeda et al. (2018)). Consequently, existing surveys need to be complemented with current data and additional analysis. Our objective is to contribute to the design of a transparent and comparable early warning framework that identifies deprived children in areas that are both data-constrained and vulnerable to the impacts of climate change.

In what follows, we develop a drought-contingent framework, for five countries in Southern Africa, that identifies children at risk of being stunted. This paper makes four contributions that may inform policies designed to mitigate the impacts that climate change has on the world's most vulnerable populations. First, we provide a prototype of a drought-contingent targeting system for data-scarce contexts.[3] While several countries (c.f. Devereux and Nzabamwita (2018)) and economists (e.g. Alderman (2009), Barrett (2010), Clarke and Dercon (2016), and Hill et al. (2019)) have advocated for a standard and data-driven approach, there are few studies that provide a detailed empirical analysis of a drought-contingent targeting strategy. In this paper we build prototype drought-contingent targeting frameworks that are comparable across countries because they employ standard, objective and verifiable measures of child and maternal anthropometrics, in addition to standard measures of household deprivation such as those related to housing, sanitation and education.

Second, our design is intended to assist programs in reaching beneficiaries based on

[1]See, for example, Kudamatsu et al. (2012), Alfani et al. (2019), Damania et al. (2017, Ch. 3), Wineman et al. (2017), Lentz et al. (2019), Hill and Mejia-Mantilla (2017), Baez et al. (Forthcoming), and Cooper et al. (2019). For context, see the thoughtful analysis aimed at practitioners by Hallegatte et al. (2016), and the evocative reportage by Sengupta (2018).

[2]While a consensus exists on the need for a timely, transparent and objective framework — see, for example, Alderman (2009), Barrett (2010), Clarke and Dercon (2016), Hill et al. (2019), Maxwell and Gelsdorf (2019), and especially Lentz et al. (2019) — there is little agreement on its design or implementation. de Waal (2018) argues that famines are rare and caused by conflict, which is tenuously related to weather events. We examine episodes of food insecurity, rather than famines.

[3]Our focus here is on the empirical analysis, rather than the implementation details. The latter includes a host of critical issues, including the nature of the support provided, and is outside of the scope of this paper, but is arguably more important. Interested readers may choose to begin with Del Ninno and Mills (2015) or Devereux et al. (2017).

non-monetary measures of child well-being. This is in contrast to the large poverty targeting literature (e.g. Grosh and Baker (1995), Coady et al. (2004) and Brown et al. (2018)) that has traditionally focused on proxy-means tests that use *poverty scores* — proxies for household income, consumption or wealth — or community-based assessments that use *subjective ranks*.[4] Monetary measures of poverty, while objective, are especially noisy in agrarian economies characterized by subsistence agriculture and volatile incomes (Deaton and Zaidi (2002)).[5] In addition, adverse impacts on children are likely to have the most serious long run consequences early in life. These impacts on children's human capital have been found to have long-run consequences on growth at both macro and micro levels (Alderman et al. (2006), Victora et al. (2008), and Maccini and Yang (2009)). Consequently, we focus on internationally comparable measures of child malnourishment. However, as we discuss below, non-monetary measures are also associated with significant measurement errors. Therefore, we view our analysis as complementary to studies that focus on monetary measures of poverty.

Third, we use remote-sensed data to automate our categorization of both harvest cycles and drought shocks. Our measure of drought shocks — two to four successive below-average 10-day rainfall spells during the growing season — while not typically used by economists, is standard in the agronomy literature (e.g. Sivakumar (1992), Barron et al. (2003)). Lentz et al. (2019) use a related measure — the length of the longest dry spell — which is also commonly used in the agronomy literature, but is harder to interpret and standardize across regions and countries with differing lengths of their respective growing seasons. We show that our simple measure is consistent with adverse impacts on stunting in all five countries that we examine. Especially given the concerns around comparability and timeliness, and the endogeneity of growing season lengths to short-term shocks (Sivakumar (1988)) and climate change (e.g. Linderholm (2006)), it is critical to use measures of growing season disturbances that are robust and standard across different countries in a region.

Fourth, we show that interpretable (white box) methods such as logistic regressions and decision trees are comparable to, in terms of out-of-sample predictive accuracy, black box machine learning methods such as random forests and gradient boosting. Further, we

[4]Community-Based Assessments are subjective and therefore susceptible to elite capture (Del Ninno and Mills (2015)). However, they may be a useful source of local knowledge that is not observable or quantifiable (Alatas et al. (2012), Karlan and Thuysbaert (2016)). Therefore, there is the potential to usefully triangulate this local knowledge with other assessments, as well as more objective information (Premand and Schnitzer (2018)). This perspective is also gaining currency in the humanitarian sector (Maxwell and Gelsdorf (2019)).

[5]In addition, because food prices are particularly volatile in remote rural regions that are likely to be the most affected by weather shocks (e.g. Brown and Kshirsagar (2015), Baffes et al. (2017), Hill and Fuje (2017), and Baez et al. (Forthcoming)), spatial and temporal price adjustments may not be sufficiently granular to adequately capture changes in the real values of household income and expenditure.

show that conditional inference decision trees (Hothorn et al. (2006)) are almost as robust (to the introduction of additional noise in the stunting measure) as logistic regressions and random forests. In other contexts, decision trees are widely considered to be both less robust to small perturbations of the data and more prone to over-fitting, with the attendant poor out-of-sample performance (Friedman et al. (2001), Murphy (2012)).[6] While we show that the relative disadvantages of using decision trees are less pronounced in our context, their advantages are clear — traditional proxy-means tests based on regressions are harder to interpret and more difficult to implement in the field.

Further, regression-based proxy-means tests that are augmented to include variables that capture droughts, typically assume that droughts influence welfare in a (log) linear manner (e.g. Del Ninno and Mills (2015), Hill et al. (2019)) or with one-level of interactions (e.g. Dercon et al. (2005)). Regression-based approaches can, of course, involve complicated sets of interactions (e.g. Hill and Porter (2016) and Pape and Wollburg (2019)). However, the interpretation gets considerably more involved. We show that decision trees, which involve 4-6 levels of interactions, but with much fewer variables, predict child stunting with comparable out-of-sample accuracy.

A data-driven framework has the potential to complement expert panels, such as those that set *Integrated Food Security Phase Classification* levels — the current global standard (Maxwell and Gelsdorf (2019)). 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 (e.g. Lentz et al. (2019)). Further, expert panels are not always immune to political influence and other sources of local and national bias. This problem is particularly acute in those contexts in which the incentives of local, national and international actors are misaligned (e.g. Alderman (2009), Clarke and Dercon (2016)).

We conduct this analysis for five southern African countries — Malawi, Mozambique, Tanzania, Zambia and Zimbabwe (Figure 1). We have six reasons for selecting these countries. First, all five countries have some segments of their rural populations engaged in subsistence agriculture. Second, stunting rates, particularly for the rural populations, are high in all five countries. More broadly, these countries have relatively low scores on the Human Development Index (see Appendix Figure 2). Third, rural economies and livelihoods in these countries may be characterized as primarily agrarian or agro-pastoral, rather than pastoral. Consequently, harvest cycles determine seasonal patterns. Fourth, in contrast to the countries closer to the equator, there is typically just one harvest cycle every

[6] These weaknesses are particularly relevant to decision trees that have small terminal nodes. The trees we estimate typically have terminal nodes that contain more than a 30 children, and often more than a 100 children.

year in all sub-national areas of all five countries (Northern Tanzania is an exception). This simplifies the analysis. Fifth, the countries border each other and are sufficiently similar so that some pooled analysis — i.e. analysis that combines data from all five countries — is reasonable. Sixth, unlike, for example Kenya and Ethiopia, drought-contingent targeting and adaptive safety nets had not been systematically implemented during the period in which the surveys were administered.

The rest of the paper proceeds as follows. In addition to describing the main variables, the next section describes the methodology we use to estimate harvest cycles and droughts. Section 3 provides estimates of the average impacts of droughts across the five countries we study, and then shows that these impacts are more pronounced for the poorest households. Section 4 describes our targeting framework and provides estimates for inclusion and exclusion errors for logistic regressions, decision trees, and decision tree ensembles (random forests and gradient boosted trees). We conclude by summarizing our results and discussing future directions.

2 Measuring Shocks and Outcomes in Agrarian Contexts

2.1 Child Anthropometrics as a Proxy for Child Well-Being

The first two Sustainable Development Goals (SDGs) focus on ending poverty and eliminating hunger across the world.[7] Cognizant of the limitations of defining poverty exclusively in monetary terms (Sen (1999), Deaton (2016)), the first goal also targets poverty in a broader sense — including non-monetary and dynamic measures. While goal 1.5 specifically emphasizes the need to "*build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events*", precisely how this target may be measured in practice remains unclear.

Closely related, the second goal concerns ending hunger and malnutrition. Objective measures of hunger rely on caloric intake which may not be strongly correlated with nutrition. Therefore, Deaton (2010, pg. 41) recommends measuring anthropometrics, but highlights a key data limitation.

> The obvious alternative is to use the anthropometric measures directly. Here there has been enormous progress, through the spread of the Demographic and Health Surveys...These surveys are as close to a gold standard as we are going to get in this area, although the irregularity of the DHS surveys makes it difficult to use them for monitoring, for example for assessing the effects of the food price crisis on the heights and weights of children.

Our study may be viewed as an attempt to address this limitation by combining existing Demographic and Health Survey (DHS) data with remote-sensed information. The

[7] https://www.un.org/sustainabledevelopment/sustainable-development-goals/

DHS data are used to estimate the World Health Organization (WHO) measures of malnourishment (WHO Multicentre Growth Reference Study Group and others (2006)).[8] Male and female growth curves and distributions of height/length -for-age, weight-for-age, and weight-for-height, are estimated by the WHO using samples of 882 children (longitudinal survey) and 6669 children (cross-section survey) from six countries (Brazil, Ghana, India, Norway, Oman and the United States). The medians and standard deviations from these distributions are used by the DHS to calculate z-scores for every child.[9] Children under 5, with a z-score below -2 standard deviations (with respect to their age) for a given metric are considered to be malnourished with respect to that metric.

The measurement of child anthropometrics is challenging and several caveats are worth noting.[10] First, it is assumed that children with similar rearing, health care, and exposure to environmental hazards will grow in exactly the same manner. This, in turn, assumes that ethnicity is not related to anthropometrics. For obvious reasons, this is complicated to delve into, but remains an assumption, rather than an established empirical fact.[11] However, there is some reassuring evidence that differences (between African and Indian populations) in anthropometrics that cannot be explained by income, may be explained by the disease environment (Coffey et al. (2013)) and culture (Jayachandran and Pande (2017)). Second, the measurement of a child's age is challenging (Larsen et al. (2019)). Third, the surveys are implemented by national statistical offices with differing degrees of expertise. Taken together, these facts suggest that we should expect that any empirical analysis using these measures will involve a relatively low signal-to-noise ratio.

Our focus here is on designing a prototype drought-contingent targeting system, rather than providing a systematic analysis of the causes of stunting.[12] However, it is worth noting that other studies have found that stunting is not responsive to single-sector interventions.[13] Multi-sectoral approaches that address food availability, health and sanitation *are* associated with reductions in stunting rates (Skoufias et al. (2019)).

In this paper, we view the prevalence of stunting as a time-varying, but lagged, sig-

[8]We use a harmonized version of the DHS Surveys constructed by Boyle et al. (2019).

[9]In the case of stunting, this entails subtracting the WHO median height from the child's height and then dividing by the WHO standard deviation — all with reference to a particular age.

[10]See Assaf et al. (2015) for an analysis of measurement challenges and data quality concerns.

[11]If ethnicity *does* have a role to play in influencing anthropometrics, policies that target stunted children will have hidden biases that favor some ethnicities over others.

[12]Our results speak to concerns raised by Deaton and Cartwright (2018) regarding the causal attribution of outcomes associated with interventions that are randomly implemented at different locations. We show that weather shocks are frequent and impact stunting. Therefore, shocks that occur *after* randomization may result in the misattribution of the impacts of interventions that aim to reduce child malnourishment.

[13]For example, using sub-national data from 59 countries, Headey and Palloni (2019) found no impact of water, sanitation and hygiene variables on stunting.

nal of child deprivation — that may usefully, when combined with remote-sensed data, inform the design of social protection and other government assistance. We do not, however, take a position regarding the precise nature of the support required to alleviate the underlying deprivation(s). Further, a weather disturbance in an agrarian context may be thought of as an additional deprivation (in the sense used by Sen (1999)), because it constrains choices and livelihood opportunities. Support triggered via an adaptive safety net should therefore not be limited to maternal and child health, but instead involve sectors most relevant to that location and country.

2.2 Measuring Shocks to Crop Production

There is no consensus on the best approach to monitoring crop production in Sub-Saharan Africa. At the same time, there is broad agreement on both the challenges and the relevant sources of information. There are two related challenges to measuring crop production in these data-constrained contexts (e.g. Burke and Lobell (2017), Jin et al. (2017)). First, there is a dearth of information on actual crop yields, particularly for small-holder farmers — farmers that constitute the majority of poor households. Second, without adequate ground-truthing, analyses of high-resolution satellite data provide noisy predictions. This is exacerbated by the fact that areas under crop production are constantly changing and the mix of crops produced together is hard to classify across contexts. Consequently, it is also challenging to transfer knowledge gained in a particular context to other locations.

Therefore, organizations charged with monitoring crop production in Sub-Saharan Africa rely on remote-sensed information averaged at broader spatial scales (e.g. Funk et al. (2019)) — traditionally focusing on vegetation (e.g. Becker-Reshef et al. (2010), Brown (2014)) and rainfall (Funk et al. (2015)), both of which have longer time series than the newer measures. Unlike rainfall, the Normalized Difference Vegetation Index (Tucker (1979)) is a cumulative measure. With available ground data, as is the case in the United States, remote-sensed vegetation anomalies can be calibrated to historic crop yields to build models that provide accurate out-of-sample estimates (Johnson (2014)). However, when ground data are scarce, vegetation indices cannot be properly calibrated and less is known about the relationship with crop-productivity in particular contexts (Burke and Lobell (2017)). Rainfall, during the growing season, is unambiguously correlated with crop productivity. However, the influence of rainfall on yields depends on its timing and distribution. Therefore, aggregating rainfall estimates over a growing season can be misleading (Sivakumar (1992), Barron et al. (2003)).

Given these challenges, we employ the following strategy. First, we focus on shocks at a particular spatial scale: a country's first administrative level (Hijmans et al. (2019)) —

which we refer to as a province for the rest of this paper.[14] Adverse weather events are typically spatially correlated, and even expected long-run changes in weather patterns, in Southern Africa and elsewhere, are expected to occur at broad spatial scales (e.g. Shongwe et al. (2009)). This choice is consistent with interpreting our shocks as being those that affect the entire local economy. Second we use Normalized Difference Vegetation Indices (NDVI) — averaged at the province x month level — to determine harvest cycles for each province. Third, we use rainfall estimates, averaged for each province, to estimate sequences of dry spells. In the rest of this section we describe our methodology and results in more detail.

A. *Estimation of Harvest Cycles*

The NDVI data 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 eight days at an approximately 250 meter ground sample (pixel) resolution across the entire study area. To isolate the NDVI signal to only those areas where crops are expected the time series was 'masked' by the Croplands and Cropland/Natural Vegetation Mosaic categories found within the IGBP land cover classification (Loveland et al. (2000)). Next, to reduce potential pixel-level noise (both in time and space) and simplify the modeling efforts, the measurements were aggregated to monthly values at a 5 kilometer grid cell size. And finally, the values were spatially and temporally averaged further to each province x month.

We identify harvest cycles — more specifically growing months, the month that a harvest starts, and the lean season — in the following algorithmic manner.[15] First, we calculate the percentage change in monthly NDVI. A month is labeled as a growing month if the percent increase is greater than 20 percent. Second, the start of the harvest season is defined as the first month in which the NDVI falls by at least 5 percent. Third, the lean season is defined as the three months preceding the start of the harvest season.

For a graphical illustration of the use of monthly NDVI averages to estimate harvest cycles, please refer to Figures A-1a to A-1e in the appendix. There are subtle differences in harvest cycle patterns even within a country. For example, in Mozambique (Figure A-1b), the South-Western province of Manica has a slightly earlier harvest cycle than the North-Eastern province of Nampula. In Manica, the average November NDVI is greater than

[14]These are districts in Malawi, regions in Tanzania, and provinces in Mozambique, Zambia and Zimbabwe.

[15]Our purpose is to standardize these definitions so that they can be used to build an automated (i.e. rules-based) definition of a drought shock. This is a prototype; in practice we could use NDVI data at more disaggregate temporal frequencies, allow for differences in harvest cycles at more refined spatial scales, and incorporate other data (e.g. soil moisture) into this algorithm.

its average for October, while for Nampula the two months are similar. Consistent with this, NDVI peaks in March in Manica, while it peaks a month later in Nampula. Unlike in temperate countries, tropical Africa has a green harvest and some local food insecurity is alleviated just after the NDVI has reached its peak.[16] More pronounced within-country differences are evident closer to the equator. In Tanzania (Figure A 1-c), Kilimanjaro, in the Northern Highlands, exhibits a bimodal pattern. In contrast, Iringa and Mbeya, in the Southern Highlands, have unimodal harvest cycles.[17]

We have chosen not to use deviations from NDVI averages as a measure of weather disturbances. This is because NDVI deviations may reflect deviations in vegetation unrelated to food. However, it is essential to have a broad understanding of what harvest cycles typically look like for each province, to more accurately estimate the impacts of weather disturbances during a child's birth season. In what follows, we use this understanding, along with rainfall data, to develop measures of drought shocks.

B. *Identification of Dry Spells*

A consensus is yet to emerge on best-practice approaches to measuring droughts (Trenberth et al. (2014)). Soil moisture, evapotranspiration and water levels are all important. However, there are significant challenges to developing comprehensive measures for tropical Sub-Saharan Africa and other data-scarce contexts. At the same time, there is little disagreement that anomalous rainfall during the growing season adversely affects crop productivity (Funk et al. (2019)). Consistent with our overall objective, here we develop a simple and transparent approach to measuring droughts. Other than the use of the NDVI to estimate harvest cycles, the only data we use are the CHIRPS Rainfall dataset (Funk et al. (2015)).

Our algorithm involves the following steps. First, we calculate the average and standard deviations of total rainfall (in mm) for every 10-day period or decad in a given month in any given province (i.e. decad x month x province).[18] Second, we calculate the standardized anomalies for every decad x month x province x year. Third, we exclude all decads that are not in the growing season. Fourth, we define a dry spell of length n for a given decad, as having below average rainfall in that decad, as well as the $n-1$ preceding decads. In addition, at least half the decads in a dry spell are required to receive less than

[16]Expert, on the ground opinion (e.g. http://fews.net/southern-africa/mozambique) is broadly consistent with these observations (scroll down to look at the seasonal calendars for South-Central and Northern Mozambique).

[17]See, for example, http://fews.net/east-africa/tanzania.

[18]The length of the last decad in a month adjusts so that it just covers the month. Consequently, decads may be compared across years.

-0.5 standard deviations rainfall. Fifth, we define a harvest as being affected by drought if the preceding growing season contains at least one dry spell.

Our measure is designed to capture the idea that large, but infrequent, amounts of rainfall during a growing season may be consistent with close to average total rainfall during a growing season, but will still cause stress and consequently engender below-average harvests. In addition, this approach ensures that a dry spell is not characterized by a sequence of decads with below-average rainfall, each of which is close to the average.

Sivakumar (1992) has previously employed a very similar approach. Instead of assuming that, at least, half of the decads in a dry spell have a -0.5 standard deviation rainfall, he sets a threshold in mm (and defines a shock for different definitions of the threshold). Our approach is better suited to be used across regions in a country and across countries because we employ standardized measures.

Table 1 describes the incidence of dry spells for the five countries we study. There are differences across countries. Mozambique and Tanzania had the lowest incidences of 1-decad and 2-decad dry spells. In Mozambique a 1-decad dry spell occurred only 55 percent of the time between 2001 and 2018, while a 2-decad dry spell occurred 43 percent of the time. In contrast, Malawi exhibited the greatest occurrence of dry spells, with a 1-decad dry spell occurring 89 percent of the time and a 2-decad dry spell occurring 71 percent of the time. A 4-decad dry spell occurs rarely — less than 10 percent of the time — in 3 of the countries. Therefore, we choose a 3-decad dry spell as our main measure of drought.[19]

When a drought is defined 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 between 2001 and 2018, while Zambia experienced droughts 32 percent of the time, and once again Malawi is the outlier with droughts occurring 45 percent of the time. Further, Mozambique and Malawi experienced greater incidences of droughts in the last decade in comparison to the previous one.

In this section, we have explained our approach to, and rationale behind, measuring child well-being and drought shocks in contexts in which data are scarce. In the next section, we establish a causal relationship between droughts and child stunting, quantify the average magnitudes of the drought impacts, and show that these impacts vary across countries and households.

[19]In the appendix (Tables A3 and A4) we show that other lengths are also consistent with significant impacts of droughts on stunting.

3 Quantifying the Impact of a Drought on Stunting

3.1 Relationship between Droughts and Stunting

In what follows, we document two important characteristics of the relationship between drought shocks and child stunting. First, we show that our measures of drought shocks are causally related to stunting and other child malnourishment outcomes. Second, we show that the average magnitudes of these impacts are comparable to the impacts associated with differences in wealth status, as well as with differences in a mother's literacy level.[20]

It is possible that droughts occur more frequently in areas that are less developed and that are characterized by a greater prevalence of child malnourishment. The most disaggregate spatial unit in the Demographic and Health Surveys (DHS) is the cluster. This is a village or a small collection of villages. By employing cluster x month fixed effects, we control for time-invariant influences at the cluster level (such as variation in remoteness, local governance and social structures, and in agro-ecological factors such as altitude and soil quality) interacted with the factors related to the month (harvest cycles, festivals, and school calendars). The drought coefficients are identified by variation in the incidence of shocks across years.

When we discuss controlling for month effects, we refer to the month that a child was born. This is measured with error (Larsen et al. (2019)). In the appendix (see Tables A1 and A2), we show that our results are robust to removing children born at the start of the year (January and February), the end of the year (November and December), children with birth months corresponding to whole years, and including *age* and age^2 as explanatory variables.

Table 2 describes results when child anthropometric outcomes are regressed against droughts, with cluster x month fixed effects. The sample includes rural populations of five Southern African countries (Mozambique, Malawi, Tanzania, Zimbabwe and Zambia). For the stunting regression, we control for 2,934 cluster x month fixed effects. As a first-pass estimate of the magnitude of the impact of drought shocks on the likelihood of stunting, it is useful to start with this basic set of regressions that uses wealth quintiles as the only household-level controls (Equation 1). S_{ict} measures whether a child i born in cluster c in year t is stunted at the time of measurement. D_{ct} is a dummy variable that

[20] These are likely to be under-estimates. Responses to droughts — both systematic and ad-hoc — from national governments, international organizations, and even large non-profits, may have mitigated the impacts of previous shocks in all of the countries that we examine. In addition, there are other confounding factors that we have not adjusted for — including those related to migration, and infant and child mortality. As a robustness check, we exclude children whose families moved in the preceding 5 years before they were surveyed and show that our results are robust to this exclusion (Table A5). Further, wealth is measured at the time of the survey and not at the time of the occurrence of a drought.

captures whether a drought occurred in cluster c during the growing months leading up to the harvest season (t) that coincided with the child's birth.[21] $D_{c(t+1)}$ is similar except it refers to the year after the birth year. W_i is a vector of four dummy variables indicating whether the child was born in a household in one of the top four wealth quintiles (the lowest wealth quintile is the contrast variable). The month of birth x cluster fixed effect is captured by θ_{cm}. α is the constant. The error is clustered at the cluster level.

$$S_{ict} = \alpha + \beta_0 D_{ct} + \beta_1 D_{c(t+1)} + \gamma W_i + \theta_{cm} + \varepsilon_{ic} \tag{1}$$

Wealth quintiles are relative measures and cannot be compared across countries. However, after controlling for cluster fixed effects, they may be used to estimate the relationship between wealth status and the likelihood of stunting for households in the same cluster. Two points are worth noting. First, as expected, after controlling for geographic and spatial factors, differences in wealth are associated with differences in child anthropometric measures. When compared with the poorest quintile, every other quintile is significantly less likely to be stunted. Second, drought impacts are comparable to a household falling from the second wealthiest quintile to the poorest quintile. On average (for the five countries taken together), for stunting outcomes, the impact of a drought during a child's birth season (10 percentage points) is similar to the difference between being born to a household in the poorest quintile instead of the second wealthiest quintile (8 percentage points).

While these numbers are large, three important caveats need to be kept in mind. First, there is limited variation in wealth status within a cluster. Second, the wealth quintiles are constructed (c.f. Rutstein (2015)) in a manner that suggests that this DHS wealth index is probably a very rough approximation of our idea of household wealth — in particular, no attempt is made to account for differences in the quality and price of the underlying assets, especially the two most important rural assets (land and livestock). Third, these impacts are averages across countries. However, what we would like to stress is that the magnitude of the impact of a drought on stunting is substantial, especially given that droughts (based on our definition) occur more than a fifth of the time in all five countries.

Table 3, which reports results from Equation 2 below, uses a more complete set of controls, including maternal biological factors M_i — which include the mother's height, weight and age and the birth order and the birth interval —- as well as all the assets (A_i) that are used to construct wealth indices, instead of the quintiles of the wealth index.

$$S_{ict} = \alpha + \beta_0 D_{ct} + \beta_1 D_{c(t+1)} + \beta_m M_i + \beta_a A_i + \theta_{cm} + \varepsilon_{ic} \tag{2}$$

[21] As discussed above, this varies by province, we use the cluster here to make the notation simpler.

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. Forty-one percent of children under five, in this rural population, are stunted. Therefore, 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 a child being underweight 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.[22]

We also report results for wasting in Table 3, however, just 5.1 percent of the children in this population are wasted, and we find no impact. A child could be stunted *and* underweight and not wasted. 12.6 percent of the children are stunted and underweight. A drought is associated with a 3 percentage point increase in children that are both stunted and underweight. 45.2 percent of the children are malnourished based on any of the three metrics. Droughts are associated with a 6 percentage point increase for this category.[23]

While employing cluster x month fixed effects provides evidence in favor of a causal relationship between droughts and malnourishment outcomes, the method requires a child to have at least one counter-part, meaning another child under 5 born in the same cluster and month. For the stunting regressions, this approach forces us to eliminate almost a third of our sample (31 percent) — 12,800 children out of 40,817 — leaving us with 28,017 observations. To increase our sample size, as well as connect our results to drought-contingent targeting (discussed in the next section), we use province (θ_p) and year (θ_y) fixed effects for our main regressions.

$$S_{ict} = \alpha + \beta_0 D_{ct} + \beta_1 D_{c(t+1)} + \gamma W_i + \beta_p \theta_p + \beta_y \theta_y + \varepsilon_{ic} \tag{3}$$

Table 4 describes results from estimating equation 3. Coefficients are estimated using variation within provinces rather than within clusters. The magnitudes of the drought impacts are comparable to those obtained from the more stringent specifications that use cluster x month fixed effects. Similarly, they are quantitatively similar to the effects of wealth on stunting. The adverse impact of a drought during the birth season (6 percentage points) is comparable to a child being born to a household in the poorest wealth quintile instead of the middle wealth quintile (5 percentage points). These impacts are

[22] While it may appear that variables that have the greatest influence are the biological ones involving a mother's height and age, the mother's height is correlated with wealth, even within a cluster. Therefore it is difficult to separate biological and economic causes.

[23] A child is stunted (underweight) if her/his height-for-age (weight-for-age) is two standard deviations below the WHO median. A child is wasted if her/his weight-for-height is two standard deviations below the WHO median.

even larger for droughts that occur the year after a child's birth. The impact of a drought during the birth season is more muted for a child's weight (measured at the time of the survey), but it is still statistically significant. Being underweight is likely to be influenced by more contemporaneous factors, while impacts on stunting are more lasting.

This method also allows us to estimate the impacts on a child being born during the lean season. The impact, on stunting, of a child born during the lean season (3 percentage points) is as consequential as the difference in a child being born into a household in the poorest quintile instead of the second poorest quintile. Once again, the likelihood of being underweight, if a child is born during the lean season, is muted (but still statistically significant).

Table 5, comparable to Table 3, uses a full set of controls, but employs province and survey year fixed effects (equation 4 below).

$$S_{ict} = \alpha + \beta_0 D_{ct} + \beta_1 D_{c(t+1)} + \beta_m M_i + \beta_a A_i + \beta_p \theta_p + \beta_y \theta_y + \varepsilon_{ic} \tag{4}$$

Once again, it is clear that the height of a child's mother is a major influence on the likelihood of a child being stunted, even after controlling for factors correlated with household wealth and income. The magnitudes of the drought impacts are substantial, when compared to the other factors. A child has a 4 percentage point greater likelihood of being stunted if born during a poor harvest than a normal harvest. This is double the magnitude of the impact associated with having a mother that is not literate. A child born to a partially literate or illiterate mother has a 2 percentage point greater likelihood of being stunted (controlling for other wealth correlates). Similarly, other important deprivations — such as not owning livestock and not treating drinking water — are comparable to drought impacts.

Having established that average drought impacts are causal and significant, both in terms of magnitudes and p-values, it is worth examining how these impacts may vary across countries. This is possible because we can use variation in birth years. If we were trying to understand drought impacts on the diversity of a child's diet, for example, the available data would limit us to exploiting cross-section variation.[24] Given that weather anomalies are spatially correlated, at different spatial scales, this would necessitate combining data from several countries and regions.

3.2 Impacts across Countries

Droughts may exert differing stunting impacts in different countries, even in the case in which all the countries are in the same region. First, countries have very different so-

[24]The DHS reports nutrition intake for children between 6 and 24 months at the time of the survey.

cial protection, nutrition and food policies. Second, countries have different historical relationships with international organizations that provide drought-contingent support. Third, countries have different social structures, which result in differences in how societies help the most vulnerable diversify risks associated with shocks. In what follows, we restrict our analysis to stunting outcomes, but it is worth keeping in mind that droughts also impact a child's weight.

Table 6a describes stunting impacts across the five countries, in a manner that is analogous to Table 3. Specifically, we begin with the first regression in Table 3 and repeat the exercise for each of the five countries separately. With the exception of Malawi, the other four countries exhibit drought impacts that are between 5 and 12 percentage points, all of which are statistically significant. Zambia (12 percentage points) and Zimbabwe (11 percentage points) have the largest impacts. Mozambique is close to the average of the five countries (7 percentage points), and Tanzania is the smallest (5 percent points), but still large.[25]

In contrast, a drought during the birth season in Malawi is not associated with a greater likelihood of stunting. We conjecture that there are three possible reasons for this. First, our algorithm may not capture harvest cycles as accurately for Malawi. Second, Malawi may have policies that protect children during the birth year, especially in drought years. Third, Malawi also has recurring floods. Therefore, it is possible that the drought dummy is being compared to both flood years and normal years. However, even in Malawi, droughts do exert a large and significant impact on stunting (9 percentage points) if they occur the season after a child is born.

Table 6b shows results for the same exercise in which a child's Height for Age Z-Score (HAZ, multiplied by 100) is the dependent variable. The results are very similar. On average a drought during the birth season is associated with a 0.19 standard deviation decrease in the height for age z score. Mozambique, Zambia and Zimbabwe have impacts that are larger than the average and Tanzania and Malawi have impacts that are more muted. But once again, even for height for age, the drought impacts (for either or both the birth season or the season following the birth) are statistically significant for all five countries, even after controlling for cluster x month fixed effects.

Taken together, the results in Tables 6a and 6b support the arguments we made above regarding the causal attribution of the impact that drought shocks have on child stunting. This is encouraging from the perspective of the design of an internationally comparable drought-contingent system, because Tables 6a and 6b provide evidence in support of a causal relationship for each of the five countries. Future work will need to repeat this

[25]All five countries show similar relationships between a mother's height and stunting, as well as a child's age and stunting. This speaks to reasonable data quality standards across the surveys.

exercise for other parts of the world. The Sahel region and the Horn of Africa are important regions, that are different than the countries we have examined here, in terms of both agro-ecology and governance.

3.3 Heterogeneous Impacts across Household Types

Given that budgets are limited, it would be very useful for interventions that aim to mitigate the impacts of droughts to be able to identify ex-ante (at a low-cost) children in households that are most likely to be affected. While this may seem straightforward, there are both conceptual as well as measurement challenges involved. From a conceptual perspective, should interventions target children that are in the poorest households, or children close to the stunting threshold, or children with profiles that have historically exhibited the largest drought impacts?[26] If children that are close to the stunting threshold are also more vulnerable to drought shocks, and in the poorest households, then supporting this group of children would address all three concerns. However, as we show below, these categories may not fully overlap.

From a measurement perspective, the set of surveys that capture child anthropometrics, as well as monetary poverty measures, both of acceptable quality, is very small. The DHS data on child anthropometrics are of reasonable quality, but the surveys do not have any economic measures. We are left with using the wealth indices, which are imperfect, for reasons discussed earlier.

Brown et al. (2017) show that around half of all undernourished children and underweight females, in their sample of 30 countries in Sub-Saharan Africa, are found in the top three DHS wealth quintiles. They use this fact to infer that a large fraction of poor individuals do not live in poor households, as defined by the wealth index. While this is certainly plausible, in our view, at least part of what they attribute to intra-household inequality or local health effects could merely reflect the error associated with measuring a household's wealth status.[27] In any event, the facts they document, rather than their inferences regarding those facts, are important for our purposes. Figure 2 (which excludes outliers) shows that there is significant overlap in height for age z-scores across the five wealth quintiles for all the five countries in our study. Consequently, the DHS wealth quintiles may not be a good proxy for stunting.

Wealth quintiles are, however, the best available proxy for household wealth in the

[26] The incentives of policy makers and other organizations may be consistent with approaches that meet particular targets, for example, a given reduction in stunting rates. However, the threshold of a negative 2 standard deviation difference from the median is not based on biology or economics.

[27] The data that are needed to address these issues are not available. In the future, there may be surveys that provide good quality measures of anthropometrics and economic variables.

DHS data. Further, while they may not predict stunting accurately, the impact of a drought may vary across wealth quintiles. Figure 3-a shows the distribution of HAZ scores for the lowest quintile, contingent on a drought. Compare this with Figure 3-b, which shows the drought-contingent HAZ distributions for the top four wealth quintiles. First, the density close to the stunting threshold is higher for the lowest wealth quintile. Second, the shift leftward appears to be more pronounced for the lowest wealth quintile. Tables 7 and 8 confirm this observation.

Table 7 starts with the regression reported in Table 5 (specification 1) and estimates this regression for different sub-samples of wealth quintiles and countries. Since sample sizes are limited, we group the second and third lowest quintiles and the wealthiest two quintiles. For all the five countries taken together, the impacts are muted for the top two wealth quintiles. The top two quintiles exhibit a 1 percentage point increase in stunting, considerably smaller than the 6 percentage points impact for the bottom quintile. In Mozambique, Malawi (for the season after birth) and Tanzania, the differences in impacts across wealth quintiles are large. The magnitudes of the drought impact for the lowest quintile is more than double the impact for the wealthiest two quintiles. However, in Zimbabwe and Zambia, the drought impacts are comparable across wealth quintiles.

It may be that in Tanzania, Malawi and Mozambique, children in the wealthier quintiles have z-scores significantly above the stunting threshold, and are still vulnerable to drought shocks, even though they do not fall below the threshold. Figure 2 shows that the z-scores for children in the wealthiest quintile in these three countries are, in fact, lower than the z-scores for the comparable children in Zambia and Zimbabwe. There are smaller proportions of children in the wealthiest quintiles in Zambia and Zimbabwe below the -2 standard deviation cut-off.

Table 8, repeats the exercise reported in Table 7, except using height for age z-scores, rather than stunting outcomes as the dependent variable. The results are very similar. In Mozambique, Malawi (once again for the season after birth) and Tanzania, drought shocks are associated with larger impacts for the poorest wealth quintile compared to the wealthiest. The results are most pronounced for Mozambique — a country with large differences in agro-ecological conditions, wealth and connectivity. For the poorest quintile, a drought shock in the birth season is associated with almost half a standard deviation lower z-score (-0.465 standard deviations). In contrast, the wealthiest quintiles have no decline in z-scores (a 0.041 standard deviation increase that is not significantly different from zero). At the opposite extreme, in Zimbabwe, impacts are close to the national average (-0.29 standard deviations) regardless of wealth status.

We draw three conclusions from this exercise. First, although the data are noisy, we show that poor harvests (either during the birth season or the following season) exert a strong adverse impact on children born to households in the lowest wealth quintile in all five countries. This holds true for both stunting as well as for height-for-age outcomes. Second, in four of the countries (Tanzania, a food surplus country, is the exception), children in households in the second and third lowest wealth quintile are also significantly affected by droughts. Third, we show that for the top two quintiles, the impacts vary by country. Mozambique, Tanzania and Malawi show considerably muted impacts, while Zambia and Zimbabwe have impacts that are comparable across the DHS wealth quintiles.[28] Taken together, this suggests that while the impacts of droughts are most pronounced for the lowest wealth quintile (which corresponds to what practitioners refer to as the ultra-poor population), broader segments of the rural populations of all five countries are affected by droughts.

4 Pragmatic Drought-Contingent Targeting Strategy

While machine learning applications to early warning frameworks, and poverty targeting more generally, hold considerable promise — particularly when timeliness and accuracy are critical — the majority of these approaches are not interpretable. Our contribution here is to emphasize the value of simple approaches in contexts that are both controversial and uncertain. Baumol and Quandt (1964) provide a useful characterization of four necessary attributes for *optimally imperfect decisions*: i) objectively measurable inputs and outputs; ii) objectively communicable decision criteria and an absence of human judgment; iii) a unique and deterministic mapping of inputs to decisions; and iv) decision algorithms that involve inexpensive and verifiable calculations.

It is worth examining a poverty targeting framework from this perspective. McBride and Nichols (2016) seminal application of machine learning to poverty targeting is the first to use an approach that is explicitly designed to improve out-of-sample prediction. However, along with Sohnesen and Stender (2017), they employ a black-box random forest methodology.[29]

Mark Schreiner's insightful and widely-used approach to designing proxy-means tests (e.g. Schreiner (2007), Schreiner (2008)) emphasizes interpretability and ease-of-use, in addition to out-of-sample performance. However, as Diamond et al. (2016) argue, his approach is not based on well-understood statistical methods and, while it does well on

[28] These results may be complemented by surveys that contain more information on wealth, income and consumption to better understand the reasons for these differences. For example, we are unable to separate the mechanisms through which drought shocks impact stunting. Possible channels involve crop income, farm labor and market prices. Further, there are social and economic channels through which shocks could be mitigated. In Zambia, different from the other countries, the mining industry provides a critical source of income and employment for several local economies.

[29] These exercises are done in the context of monetary poverty targeting, which involves a larger number of explanatory variables and a higher signal-to-noise ratio than our context.

average, its out-of-sample performance is less satisfactory for different parts of a country. Kshirsagar et al. (2017) combine penalized (regularized) regressions with sub-sampling (stability selection) to develop monetary poverty measurement tools that are both interpretable and relevant across a diverse country. This is more useful for constructing monetary poverty PMTs because there are a larger number of candidate explanatory variables.

In what follows, we train commonly used black box machine learning methods (Random Forests and Gradient Boosting) and compare these with logistic regressions and decision trees. The black box methods we use are among the most successful prediction methods for tabular data (Friedman et al. (2001)). We train the models by optimizing hyper-parameters based on the out-of-sample log-loss (cross-entropy) metric using 10-fold cross-validation.[30] We test these models using a validation data set that was not used to train the model.

Table 9 summarizes results for the four predictions models. The results are similar for all five countries. Logistic Regressions, Random Forests and Gradient Boosting have almost identical out-of-sample performance. In the context of monetary poverty, Caire and Schreiner (2013) and Brown et al. (2018) argue that objective functions are flat for a range of parameters around the optimum, and therefore more sophisticated prediction methods may not improve the out-of-sample performance of a proxy-means test. With respect to our particular context, the results summarized in Table 9 (and Table 10, discussed below) provide unambiguous support for this hypothesis.[31]

Table 9 shows that classification trees are only slightly less accurate than the other methods.[32] Interpreting differences in log-loss measures is less useful than directly estimating out-of-sample targeting errors. However, these errors cannot be summarized by one number, and depend on a government's (or an organization's) budget. At one extreme, a government may be able to afford to reach every child. In which case, the proportion of poor children that are excluded from the intervention (i.e. the exclusion error) will be zero. However, the proportion of non-poor children that are included in the intervention (i.e. the inclusion error) will be 1. At the other extreme, if the government does not reach a single child, the exclusion error will be 1 and the inclusion error will be zero. Typically governments and organizations do not have the requisite budgets to reach every single child.

The out-of sample targeting errors are estimated using these four steps. First, data on children in the training set are used to build the model and children in the test set are then

[30]For our binary classification problem this is $-\frac{1}{N}\sum[S_i log(p_i) + (1-S_i) log(1-p_i)]$. Where S_i is the dummy capturing whether child i is stunted and p_i is the predicted probability that the child is stunted.

[31]In the appendix we show that this is also true for predictions of height for age z-scores (see Table A7).

[32]In the appendix (Table A6) we show that this remains true even after introducing a zero mean and one standard deviation noise term into the height-age z-scores for children in the training set.

ranked based on the predicted (out-of-sample) probability of not being stunted. Second, for a budget of b, a child is included if the percentile is less than b. Third, this inclusion status is compared with the actual stunting dummy. Fourth, we calculate inclusion and exclusion errors. This exercise is repeated for different budgets.

Table 10 reports out-of-sample inclusion and exclusion errors for different budgets - in which 20, 30, 40, 50 and 60 percent of the children are reached. We draw three conclusions. First, the results are almost identical for logistic regressions, boosting and random forests. This is consistent with the almost identical log-loss measures. At the same time, the predictions from the three models are not perfectly correlated, with correlations typically ranging from 0.85 to 0.95. This suggests the possibility of combining these into an ensemble model that may improve out-of-sample predictive accuracy (Friedman et al. (2001)). We do not pursue this direction in this paper, because our focus is on interpretability.

Second, the methods have explanatory power, although there are errors in exclusion and inclusion. A PMT constructed using a logistic regression, with a budget consistent with reaching 60 percent of the children, will exclude between 24 and 29 percent of stunted children, while also including between 50 and 53 percent of non-stunted children. Therefore, when combined with geographic targeting or a community-based assessment, a quantitative framework designed using these methods will add substantial value to the targeting of an intervention.

Third, and perhaps most usefully for a practitioner, decision trees have comparable out-of-sample performance. Decision trees are easy to interpret and typically involve fewer variables. As such, these trees provide data-driven rules of thumb that are free from the cognitive biases typically associated with heuristics (Tversky and Kahneman (1974)). Further, even compared to logistic regressions, they are much easier to implement in the field. However, practitioners would be limited to a certain number of budget choices, because the predictions depend on the number of terminal nodes, and are therefore discontinuous.

For example, in Mozambique, a budget consistent with reaching 62 percent of the children is consistent with excluding 26 percent of stunted children and including 52 percent of non-stunted children (Table 10). This is comparable to a logistic regression used to reach 60 percent of the children. However, if the government has a smaller budget, it would need to reach just 28 percent of the households. The exclusion errors (63 vs 59) and inclusion errors (21 vs 22) are comparable to a method informed by a logistic regression that reaches 30 percent of the children.

Continuing with Mozambique, Figure 4 illustrates an example decision tree. Zones (i.e. Provinces) 1-6 are in Central and Northern Mozambique.[33] The tree in Figure 4

[33]These are Cabdo Delgado, Niassa, Nampula, Zambezia, Tete and Manica.

shows that droughts most affect rural children who live in provinces in Northern and Central Mozambique, are above 6 months old and have mothers under 5 ft. These are the provinces that have a greater proportion of households engaged in subsistence agriculture.

Figure 5, for rural Zambia, shows that geography is less important in predicting stunting. Once again, the most vulnerable children are above 6 months old. For children with mothers who are of intermediate height (between 5 ft and 5ft 3 inches), the drought impacts depend on household assets. In the case of children born to tall mothers (i.e. above 5ft 3 inches in height), the likelihood of stunting is, in fact, high if the mother is underweight and lives in a house with a primitive roof.

Our objective here is to demonstrate that decision trees provide several advantages over less interpretable and harder to implement targeting tools — at least in the context that we have examined. In practice, considerably more work needs to go into incorporating local information, program contexts, and geographic data into the design of a targeting tool. We believe that incorporating objective and feasible targeting tools into program design will allow resources to more efficiently reach their intended beneficiaries.

The shock-sensitive programs in both Kenya and Ethiopia do attempt to use designs informed by regressions — although successful in serving the poor, these programs have experienced challenges expanding and contracting in a state-contingent manner.[34]

5 Conclusions

We have constructed a prototype drought-contingent targeting framework that may be used to inform the design of social safety nets in contexts in which data are scarce. To do this we address three concerns. First, we have developed measures of drought shocks that are simple, transparent and easy-to-communicate. Second, we have shown that droughts have a large impact on stunting in Southern Africa — comparable, in size, to the effects of mother's illiteracy and in a fall to a lower wealth quintile. Third, we show that interpretable methods such as decision trees and logistic regressions predict stunting as accurately as black-box machine learning methods. Taken together, our analysis lends support to the idea that a data-driven approach may contribute to the design of policies that mitigate the impact that climate change has on the world's most vulnerable populations.

We have focused on one particular spatial scale — a first administrative level (or province or region). Weather disturbances occur at different spatial scales. At one extreme,

[34] While several studies have documented the positive impacts of Ethiopia's Productive Safety Net Programme (e.g. Gilligan et al. (2009)), we are not aware of a study documenting its impacts on shock mitigation. Further, Berhane et al. (2017) find that the program did not improve child nutrition outcomes — although there are no controls for weather events or other shocks.

shocks may occur at the household or village level. At another, shocks may occur at a national or region level. 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 remote rural areas.[35] Further, weather disturbances in food-surplus areas may engender food scarcity in other parts of a country.[36] A drought-contingent targeting framework needs to incorporate this information.

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.

[35] See, for example, Baffes et al. (2017), Hill and Fuje (2017), and Baez et al. (Forthcoming).

[36] Baffes and Kshirsagar (forthcoming) develop a quantitative approach to understanding *market systems* in Tanzania — in particular, they identify markets that are the main suppliers of maize and rice, and markets from which demand shocks originate.

Table 1: The Frequency of Droughts by Country and Decade

	Average Incidence of Dry Spells				
Period	**1 Decad**	**2 Decads**	**3 Decads**	**4 Decads**	**5 Decads**
Mozambique					
2001-2009	0.52	0.39	0.16	0.07	0.04
2010-2018	0.61	0.50	0.32	0.13	0.00
2001-2018	0.55	0.43	0.22	0.09	0.03
Malawi					
2001-2009	0.85	0.64	0.39	0.20	0.15
2010-2018	0.95	0.82	0.54	0.23	0.10
2001-2018	0.89	0.71	0.45	0.21	0.13
Tanzania					
2001-2009	0.60	0.49	0.26	0.09	0.05
2010-2018	0.52	0.35	0.16	0.06	0.01
2001-2018	0.57	0.43	0.22	0.07	0.04
Zambia					
2001-2009	0.71	0.58	0.32	0.11	0.03
2010-2018	0.78	0.61	0.31	0.20	0.09
2001-2018	0.74	0.59	0.32	0.14	0.06
Zimbabwe					
2001-2009	0.61	0.41	0.20	0.03	0.01
2010-2018	0.79	0.61	0.25	0.11	0.01
2001-2018	0.68	0.49	0.22	0.06	0.01

Notes: 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.

Table 2: Drought Impacts on Child Anthropometrics
OLS with Cluster x Month Fixed Effects and Wealth Quintile Controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Stunted**	**Underweight**	**Wasted**	**Stunted & Underwt**	**Any**
Drought: Birth Season	0.10***	0.04***	0.00	0.04***	0.09***
	[8.16]	[4.01]	[0.01]	[5.11]	[7.20]
Drought: Season after Birth	0.10***	0.01	-0.03***	0.03***	0.07***
	[7.41]	[1.26]	[-4.39]	[3.14]	[5.00]
Second Poorest	-0.03**	-0.03**	-0.01	-0.02**	-0.04***
	[-2.31]	[-2.55]	[-0.90]	[-2.12]	[-2.99]
Middle	-0.03**	-0.05***	-0.01	-0.04***	-0.04***
	[-2.20]	[-4.16]	[-0.81]	[-3.72]	[-2.63]
Second Richest	-0.08***	-0.07***	-0.01*	-0.07***	-0.09***
	[-4.46]	[-5.22]	[-1.69]	[-5.25]	[-4.97]
Richest Quintile	-0.14***	-0.08***	-0.01	-0.07***	-0.14***
	[-4.16]	[-4.04]	[-0.54]	[-4.11]	[-4.03]
Constant	0.39***	0.17***	0.06***	0.14***	0.45***
	[41.12]	[24.47]	[15.17]	[21.39]	[46.53]
Observations	28,162	28,144	28,062	28,162	28,162
R-squared	0.43	0.40	0.40	0.40	0.42

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. All regressions use all available data on rural households from 5 Southern African countries (Mozambique, Malawi, Tanzania, Zambia and Zimbabwe). All metrics are based on children being measured below -2 standard deviation thresholds. 41.0 percent of the children are stunted, 15.1 percent are underweight, 5.1 percent are wasted, 12.6 percent are stunted and underweight, and 45.2 percent are malnourished based on any of the three metrics.

Table 3: Drought Impacts on Child Anthropometrics (OLS with Cluster x Month Fixed Effects)

VARIABLES	(1) Stunted	(2) Underweight	(3) Wasted	(4) Stunted & Underwt	(5) Any
Drought: Birth Season	0.07***	0.03***	0.00	0.03***	0.06***
	[5.43]	[2.95]	[0.40]	[3.96]	[4.52]
Drought: Season after Birth	0.04***	0.01	-0.01	0.02*	0.03**
	[3.07]	[0.57]	[-1.28]	[1.77]	[2.43]
Mother ht between 152 and 160 cm	0.13***	0.06***	0.01**	0.06***	0.13***
	[10.75]	[7.79]	[2.39]	[7.97]	[10.83]
Mother ht below 152 cm	0.25***	0.13***	0.01*	0.12***	0.24***
	[17.56]	[11.00]	[1.86]	[11.53]	[16.94]
Mother under 19 yrs	0.05***	0.03*	-0.01	0.02*	0.05**
	[2.72]	[1.87]	[-0.76]	[1.83]	[2.46]
Age:7-24 months	0.19***	0.06***	-0.01	0.07***	0.17***
	[11.98]	[5.32]	[-1.01]	[6.44]	[10.62]
Age:25-59 months	0.23***	0.05***	-0.05***	0.07***	0.18***
	[15.38]	[4.67]	[-5.86]	[6.82]	[12.27]
No Livestock	0.03***	0.00	0.00	0.01	0.02*
	[2.77]	[0.33]	[0.26]	[1.19]	[1.85]
Does not treat water	0.03**	0.01	-0.01	0.01	0.02*
	[2.18]	[0.88]	[-1.62]	[1.55]	[1.78]
Mother underweight	0.04**	0.10***	0.03***	0.08***	0.07***
	[2.19]	[5.97]	[3.00]	[5.73]	[3.79]
Mother able to read part of sentence	0.01	-0.01	-0.01	-0.01	0.00
	[0.70]	[-0.91]	[-1.30]	[-0.96]	[0.23]
Mother cannot read at all/Unknown	0.01	0.01	0.01**	0.00	0.02
	[0.69]	[1.00]	[2.14]	[0.10]	[1.51]
Constant	-0.09**	-0.14***	0.06***	-0.16***	-0.02
	[-2.25]	[-5.26]	[3.34]	[-6.65]	[-0.59]
Observations	28,017	28,152	27,987	30,057	30,057
R-squared	0.46	0.42	0.41	0.42	0.45

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. All regressions use all available data on rural households from 5 Southern African countries (Mozambique, Malawi, Tanzania, Zambia and Zimbabwe). Additional controls are for roof, floor and wall quality, cooking fuel, employment in agriculture, source of water, type of toilet, ownership of radio, tv, bicyles, refridgerators, cars, trucks, birth order, birth interval, gender of child and the gender of the household head. All metrics are based on children being measured below -2 standard deviation thresholds. 41.0 percent of the children are stunted, 15.1 percent are underweight, 5.1 percent are wasted, 12.6 percent are stunted and underweight, and 45.2 percent are malnourished based on any of the three metrics.

Table 4: Drought Impacts on Child Anthropometrics
OLS with Province and Year of Survey Fixed Effects and Wealth Quintile Controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Stunted**	**Underweight**	**Wasted**	**Stunted & Underwt**	**Any**
Drought: Birth Season	0.06***	0.02***	-0.00	0.02***	0.06***
	[7.64]	[3.40]	[-0.39]	[4.16]	[7.19]
Drought: Season after Birth	0.10***	0.01**	-0.02***	0.02***	0.08***
	[10.66]	[2.07]	[-4.50]	[3.67]	[8.77]
Lean Season	0.03***	0.01***	-0.00	0.02***	0.03***
	[5.05]	[2.80]	[-1.27]	[3.21]	[4.44]
Second Poorest	-0.03***	-0.03***	-0.01***	-0.03***	-0.04***
	[-3.73]	[-4.21]	[-3.33]	[-3.90]	[-4.61]
Middle	-0.05***	-0.05***	-0.01**	-0.05***	-0.06***
	[-5.02]	[-7.02]	[-2.44]	[-6.71]	[-5.84]
Second Richest	-0.10***	-0.08***	-0.02***	-0.07***	-0.11***
	[-8.79]	[-9.61]	[-4.01]	[-9.76]	[-9.72]
Richest Quintile	-0.16***	-0.10***	-0.03***	-0.09***	-0.18***
	[-8.81]	[-8.89]	[-4.30]	[-8.36]	[-9.67]
Constant	0.38***	0.12***	0.05***	0.10***	0.41***
	[9.98]	[6.34]	[3.52]	[5.48]	[11.15]
Observations	40,977	40,961	40,870	40,977	40,977
R-squared	0.04	0.02	0.01	0.02	0.04

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ province and year of survey fixed effects. The weights used are person weights adjusted for the national populations. All regressions use data on rural households from 5 Southern African countries (Mozambique, Malawi, Tanzania, Zambia and Zimbabwe). All metrics are based on children being measured below -2 standard deviation thresholds. 41.0 percent of the children are stunted, 15.1 percent are underweight, 5.1 percent are wasted, 12.6 percent are stunted and underweight, and 45.2 percent are malnourished based on any of the three metrics.

Table 5: Drought Impacts on Child Anthropometrics
OLS with Province and Year of Survey Fixed Effects and Full Set of Controls

VARIABLES	(1) Stunted	(2) Underweight	(3) Wasted	(4) Stunted & Underwt	(5) Any
3 Decad Dry Spell: Birth Season	0.04***	0.01*	-0.00	0.01***	0.04***
	[4.53]	[1.95]	[-0.20]	[2.64]	[4.43]
3 Decad Dry Spell: Season after Birth	0.04***	0.00	-0.00	0.01	0.04***
	[4.33]	[0.29]	[-0.23]	[0.86]	[4.00]
Mother ht between 152 and 160 cm	0.12***	0.05***	0.01***	0.05***	0.12***
	[16.21]	[10.26]	[2.63]	[10.98]	[15.99]
Mother ht below 152 cm	0.25***	0.11***	0.01***	0.11***	0.25***
	[25.99]	[14.26]	[2.84]	[15.32]	[25.46]
Mother under 19 yrs	0.05***	0.01	-0.00	0.02*	0.05***
	[4.28]	[1.57]	[-0.84]	[1.77]	[3.76]
Age:7-24 months	0.21***	0.07***	-0.01	0.07***	0.18***
	[21.23]	[9.54]	[-1.39]	[10.44]	[17.40]
Age:25-59 months	0.25***	0.06***	-0.05***	0.08***	0.19***
	[26.35]	[8.65]	[-8.91]	[11.91]	[19.52]
Lean Season	0.03***	0.01**	-0.00	0.01**	0.02***
	[3.94]	[2.24]	[-0.87]	[2.55]	[3.60]
No Livestock	0.02**	0.00	0.00	0.00	0.02**
	[2.53]	[0.53]	[0.83]	[0.69]	[2.35]
Does not treat water	0.02**	0.01*	0.00	0.01*	0.02***
	[2.51]	[1.93]	[1.05]	[1.87]	[2.78]
Mother Underweight	0.07***	0.11***	0.03***	0.09***	0.09***
	[5.80]	[9.76]	[4.39]	[8.85]	[7.39]
No Car, Motor, Fridge	0.04***	0.03***	0.01**	0.02***	0.05***
	[3.56]	[3.96]	[2.43]	[3.17]	[4.47]
Mother able to read only part of sentence	0.02	0.01	0.00	0.01	0.02
	[1.46]	[0.82]	[0.19]	[0.87]	[1.56]
Mother cannot read at all/Unknown	0.02***	0.02***	0.01**	0.01***	0.03***
	[2.90]	[2.74]	[1.99]	[2.64]	[3.46]
Constant	-0.03	-0.07**	0.05***	-0.08***	0.04
	[-0.74]	[-2.16]	[2.93]	[-2.66]	[1.00]
Observations	40,817	40,802	40,714	40,817	40,817
R-squared	0.10	0.05	0.03	0.05	0.09

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ province and survey year fixed effects. Refer to the Table 3 notes for all other details.

Table 6a: Drought Impacts on Child Stunting (OLS with Cluster x Month Fixed Effects)

VARIABLES	(1) All	(2) Mozambique	(3) Malawi	(4) Zimbabwe	(5) Tanzania	(6) Zambia
Drought: Birth Season	0.07***	0.07***	-0.03	0.11***	0.05***	0.12***
	[5.43]	[2.64]	[-0.87]	[4.08]	[2.59]	[4.96]
Drought: Season after Birth	0.04***	0.02	0.09***	0.04	0.04*	0.05**
	[3.07]	[0.47]	[2.62]	[1.21]	[1.84]	[2.01]
Mother ht between 152 and 160 cm	0.13***	0.13***	0.13***	0.11***	0.14***	0.09***
	[10.75]	[4.66]	[3.81]	[5.40]	[7.57]	[4.12]
Mother ht below 152 cm	0.25***	0.24***	0.29***	0.25***	0.27***	0.21***
	[17.56]	[7.45]	[7.49]	[7.19]	[12.08]	[7.44]
Mother under 19 yrs	0.05***	-0.00	-0.02	0.05	0.08**	0.02
	[2.72]	[-0.12]	[-0.45]	[1.30]	[2.45]	[0.68]
Age:7-24 months	0.19***	0.15***	0.14***	0.18***	0.18***	0.25***
	[11.98]	[3.95]	[2.63]	[5.82]	[7.33]	[8.35]
Age:25-59 months	0.23***	0.19***	0.18***	0.23***	0.24***	0.22***
	[15.38]	[5.34]	[4.13]	[8.39]	[10.51]	[7.06]
No Livestock	0.03***	-0.01	0.02	-0.04*	0.06***	0.07***
	[2.77]	[-0.37]	[0.58]	[-1.74]	[2.67]	[2.90]
Does not treat water	0.03**	0.02	0.00	0.06*	0.03**	0.00
	[2.18]	[0.44]	[0.00]	[1.77]	[1.98]	[0.14]
Mother underweight	0.04**	0.11**	0.06	0.07*	0.01	0.06*
	[2.19]	[2.43]	[1.11]	[1.71]	[0.36]	[1.78]
Mother able to read part of sentence	0.01	0.04	0.01	-0.00	-0.01	0.05
	[0.70]	[0.78]	[0.20]	[-0.08]	[-0.23]	[1.43]
Mother cannot read at all/Unknown	0.01	0.03	0.05	0.04	-0.00	0.01
	[0.69]	[1.04]	[1.43]	[1.16]	[-0.03]	[0.57]
Constant	-0.09**	-0.00	-0.22*	-0.16*	-0.12*	0.02
	[-2.25]	[-0.01]	[-1.65]	[-1.86]	[-1.95]	[0.26]
Observations	28,017	5,002	3,618	4,400	9,358	5,639
R-squared	0.46	0.44	0.52	0.47	0.47	0.46

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. The first regression is the same as the first regression in table 3. Additional controls are for roof, floor and wall quality, cooking fuel, employment in agriculture, source of water, type of toilet, ownership of radio, tv, bicyles, refridgerators, cars, trucks, birth order, birth interval, gender of child and the gender of the household head. All metrics are based on children being measured below -2 standard deviation thresholds.

Table 6b: Drought Impacts on Child HAZ*100 (OLS with Cluster x Month Fixed Effects)

	(1)	(2)	(3)	(4)	(5)	(6)
	All	Mozambique	Malawi	Zimbabwe	Tanzania	Zambia
Drought: Birth Season	-19.83***	-33.31***	8.89	-24.82***	-13.59**	-28.70***
	[-5.13]	[-3.81]	[0.96]	[-2.80]	[-2.18]	[-3.80]
Drought: Season after Birth	-10.63**	5.88	-21.51**	-13.75	-11.29*	-21.15***
	[-2.57]	[0.48]	[-2.15]	[-1.55]	[-1.79]	[-2.59]
Mother ht between 152 and 160 cm	-39.77***	-39.95***	-37.50***	-40.78***	-42.84***	-30.06***
	[-11.33]	[-3.95]	[-3.69]	[-7.35]	[-8.13]	[-4.33]
Mother ht below 152 cm	-75.07***	-64.73***	-79.33***	-82.95***	-81.98***	-69.08***
	[-17.79]	[-5.69]	[-7.08]	[-8.28]	[-13.29]	[-7.65]
Mother under 19 yrs	-21.58***	-10.72	-14.01	-8.80	-30.48***	-6.58
	[-3.85]	[-0.84]	[-1.00]	[-0.81]	[-3.33]	[-0.59]
Age:7-24 months	-80.25***	-60.92***	-47.24***	-95.11***	-80.60***	-106.15***
	[-15.05]	[-4.41]	[-3.13]	[-8.95]	[-10.32]	[-9.73]
Age:25-59 months	-104.15***	-88.39***	-77.60***	-119.63***	-106.90***	-112.55***
	[-20.02]	[-6.72]	[-5.78]	[-12.35]	[-13.87]	[-10.68]
No Livestock	-5.55	0.21	-8.17	10.40	-1.99	-20.15***
	[-1.43]	[0.03]	[-0.87]	[1.40]	[-0.25]	[-2.76]
Does not treat water	-4.07	-1.38	1.87	-22.23**	-4.64	2.66
	[-1.16]	[-0.08]	[0.22]	[-2.46]	[-1.02]	[0.32]
Mother underweight	-14.36***	-35.45**	-15.26	-12.08	-8.95	-13.31
	[-2.72]	[-2.47]	[-1.10]	[-0.86]	[-1.13]	[-1.53]
Mother able to read part of sentence	1.04	-7.36	-1.27	-4.08	10.68	-13.25
	[0.20]	[-0.54]	[-0.08]	[-0.42]	[1.31]	[-1.18]
Mother cannot read at all/Unknown	-0.15	-6.33	-18.93*	3.45	4.48	-6.06
	[-0.04]	[-0.64]	[-1.84]	[0.27]	[0.90]	[-0.96]
Constant	29.17**	22.04	9.21	29.49	48.16***	-8.48
	[2.52]	[0.69]	[0.27]	[1.17]	[2.64]	[-0.33]
Observations	28,017	5,002	3,618	4,400	9,358	5,639
R-squared	0.49	0.47	0.56	0.51	0.49	0.48

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. The first regression is the same as the first regression in table 3. Additional controls are for roof, floor and wall quality, cooking fuel, employment in agriculture, source of water, type of toilet, ownership of radio, tv, bicyles, refridgerators, cars, trucks, birth order, birth interval, gender of child and the gender of the household head. All metrics are based on children being measured below -2 standard deviation thresholds.

Table 7: Drought Impacts on Child Stunting by Country and Wealth Status
OLS with Province and Year of Survey Fixed Effects and Full Set of Controls

	Probability of Stunting (OLS with Province and Survey Year FE)					
	All	**Mozambique**	**Malawi**	**Zimbabwe**	**Tanzania**	**Zambia**
	Lowest (Poorest) Wealth Quintile					
Drought: Birth Season	0.06***	0.11**	-0.00	0.12***	0.06**	0.09***
	[4.79]	[2.46]	[-0.07]	[3.58]	[2.34]	[3.65]
Drought: Season after Birth	0.04***	0.07	0.07**	0.04	0.07**	0.00
	[2.97]	[1.32]	[2.34]	[1.20]	[2.58]	[0.13]
Observations	11,605	1,517	1,941	2,179	3,302	2,666
	Second and Third Wealth Quintiles					
Drought: Birth Season	0.04***	0.08***	-0.00	0.09***	-0.02	0.11***
	[3.70]	[2.84]	[-0.25]	[3.50]	[-1.16]	[5.13]
Drought: Season after Birth	0.04***	0.04	0.06***	0.03	0.02	0.09***
	[4.04]	[1.09]	[3.07]	[1.27]	[0.80]	[3.55]
Observations	20,895	3,238	4,097	3,282	6,328	3,950
	Fourth and Richest Wealth Quintile					
Drought: Birth Season	0.01	-0.01	-0.02	0.10**	-0.00	-0.00
	[0.31]	[-0.17]	[-0.93]	[2.05]	[-0.03]	[-0.08]
Drought: Season after Birth	0.03	0.04	0.03	-0.02	-0.01	0.14**
	[1.51]	[0.70]	[1.24]	[-0.39]	[-0.29]	[2.24]
Observations	8,317	1,601	2,338	877	2,859	642
	Full Sample: All Wealth Quintiles					
Drought: Birth Season	0.04***	0.08***	-0.01	0.10***	0.01	0.10***
	[4.53]	[3.32]	[-0.80]	[5.33]	[0.57]	[6.56]
Drought: Season after Birth	0.04***	0.05*	0.06***	0.03	0.03*	0.07***
	[4.33]	[1.74]	[3.97]	[1.38]	[1.89]	[3.95]
Observations	40,817	6,356	8,376	6,338	12,489	7,258

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. This table uses the first specification in Table 5 and applies it different countries and wealth quintiles. Refer to the notes in Table 3 for all the other details.

Table 8: Drought Impacts on Child Height-Age Z-Score* 100 by Country and Wealth Status
OLS with Province and Year of Survey Fixed Effects and Full Set of Controls

	Height-Age Z-Score*100 (OLS with Province and Survey Year FE)					
	All	**Mozambique**	**Malawi**	**Zimbabwe**	**Tanzania**	**Zambia**
	Lowest (Poorest) Wealth Quintile					
Drought: Birth Season	-19.62***	-46.55***	-1.27	-28.23***	-11.33	-25.00***
	[-4.63]	[-3.33]	[-0.13]	[-2.71]	[-1.54]	[-3.18]
Drought: Season after Birth	-13.06***	-28.49**	-24.93***	-16.91	-17.60***	-0.81
	[-3.32]	[-2.03]	[-2.67]	[-1.61]	[-2.64]	[-0.10]
Observations	11,605	1,517	1,941	2,179	3,302	2,666
	Second and Third Wealth Quintiles					
Drought: Birth Season	-10.92***	-29.02***	6.07	-28.42***	5.05	-28.70***
	[-3.46]	[-3.50]	[0.95]	[-3.59]	[0.90]	[-3.97]
Drought: Season after Birth	-6.72**	7.66	-18.96***	-3.80	1.40	-23.08***
	[-1.96]	[0.58]	[-3.06]	[-0.54]	[0.23]	[-3.26]
Observations	20,895	3,238	4,097	3,282	6,328	3,950
	Fourth and Richest Wealth Quintile					
Drought: Birth Season	0.42	4.15	12.55*	-35.32**	-2.92	2.72
	[0.08]	[0.25]	[1.72]	[-2.32]	[-0.32]	[0.15]
Drought: Season after Birth	-14.40***	-17.93	-15.38*	-6.48	-11.66	-35.52*
	[-2.76]	[-1.02]	[-1.91]	[-0.44]	[-1.22]	[-1.80]
Observations	8,317	1,601	2,338	877	2,859	642
	Full Sample: All Wealth Quintiles					
Drought: Birth Season	-10.73***	-31.29***	6.62	-29.02***	-0.75	-27.75***
	[-4.22]	[-4.09]	[1.49]	[-5.28]	[-0.18]	[-5.54]
Drought: Season after Birth	-9.08***	-6.59	-19.23***	-7.31	-6.39	-18.34***
	[-3.50]	[-0.70]	[-4.31]	[-1.33]	[-1.46]	[-3.38]
Observations	40,817	6,356	8,376	6,338	12,489	7,258

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. This table uses the same specification as Table 7, except the dependent variable is the height-age Z-score.

Table 9: A Comparision of Out-of Sample Predictive Accuracy of Different Methods

Method	Optimal Hyper-Parameter Values (via cross-validation)	Log-Loss
	Mozambique	
Logistic Regression	None	0.640
Conditional Inference Decision Trees	Level of Significance = 0.95	0.653
Random Forest	Min. Node Size = 200 , Number of Var Tried = 6	0.642
Gradient Boosted Trees	eta = 0.05 , max depth = 3, num rounds = 150, min child wt = 20	0.643
	Malawi	
Logistic Regression	None	0.635
Conditional Inference Decision Trees	Level of Significance = 0.90	0.646
Random Forest	Min. Node Size = 200 , Number of Var Tried = 9	0.636
Gradient Boosted Trees	eta = 0.05 , max depth =2, num rounds = 300, min child wt = 60	0.633
	Zimbabwe	
Logistic Regression	None	0.592
Conditional Inference Decision Trees	Level of Significance = 0.95	0.597
Random Forest	Min. Node Size = 300 , Number of Var Tried = 10	0.592
Gradient Boosted Trees	eta = 0.05 , max depth =2, num rounds = 150, min child wt = 10	0.591
	Tanzania	
Logistic Regression	None	0.616
Conditional Inference Decision Trees	Level of Significance = 0.95	0.644
Random Forest	Min. Node Size = 150 , Number of Var Tried = 8	0.617
Gradient Boosted Trees	eta = 0.4 , max depth = 1, num rounds = 300, min child wt = 20	0.616
	Zambia	
Logistic Regression	None	0.645
Conditional Inference Decision Trees	Level of Significance = 0.90	0.655
Random Forest	Min. Node Size = 100 , Number of Var Tried = 5	0.640
Gradient Boosted Trees	eta = 0.1 , max depth = 2, num rounds = 100, min child wt = 10	0.644

Notes: The models were trained on 70 percent of each set of surveys and tested on the remaining 30 percent. The hyper-parameters were optimized using 10-fold cross-validation (log-loss metric). The table reports the optimized (i.e. minimum) log-loss for each model. The hyperparameter candidate sets are as follows. Conditional Inference Decision Trees: Bonferroni level of significance (0.90, 0.95). Random Forests (R package Ranger): minimum node size (50, 100, 150, 200, 250, 300, 350), number of variables tried at each node (4, 5, 6, 7, 8, 9,10, 11, 12, 13, 14), and number of trees (1000). Gradient Boosted Trees (R package Xgboost): shrinkage (0.05, 0.1, 0.2, 0.3, 0.4), max depth (1, 2, 3, 4), number of rounds (50, 100, 150, 200, 250, 300) and min. child weight (10, 20, 30, 40, 50, 60).

Table 10: Out-of Sample Inclusion and Exclusion Targeting Errors

Classification Trees			Logistic Regression			Random Forest			Boosted Trees		
Incl.	Excl.	Budget	Incl.	Excl.	Budget	Incl.	Excl.	Budget	Incl.	Excl.	Budget
					Mozambique						
0.21	0.63	0.28	0.13	0.71	0.20	0.13	0.71	0.20	0.13	0.71	0.20
0.21	0.63	0.28	0.22	0.59	0.30	0.22	0.59	0.30	0.21	0.58	0.30
0.21	0.63	0.28	0.30	0.47	0.40	0.30	0.47	0.40	0.30	0.46	0.40
0.52	0.26	0.62	0.39	0.35	0.50	0.39	0.36	0.50	0.39	0.35	0.50
0.52	0.26	0.62	0.50	0.27	0.60	0.50	0.27	0.60	0.51	0.28	0.60
					Malawi						
0.13	0.76	0.18	0.13	0.70	0.20	0.13	0.70	0.20	0.13	0.69	0.20
0.16	0.68	0.23	0.21	0.57	0.30	0.21	0.57	0.30	0.21	0.57	0.30
0.32	0.47	0.41	0.30	0.46	0.40	0.31	0.46	0.40	0.30	0.46	0.40
0.39	0.38	0.49	0.40	0.36	0.50	0.41	0.36	0.50	0.40	0.36	0.50
0.43	0.34	0.53	0.51	0.27	0.60	0.50	0.26	0.60	0.50	0.26	0.60
					Zimbabwe						
0.19	0.67	0.23	0.15	0.69	0.20	0.15	0.70	0.20	0.15	0.68	0.20
0.19	0.67	0.23	0.24	0.57	0.30	0.24	0.58	0.30	0.24	0.57	0.30
0.35	0.42	0.43	0.33	0.45	0.40	0.34	0.46	0.40	0.33	0.46	0.40
0.40	0.38	0.47	0.43	0.34	0.50	0.43	0.34	0.50	0.43	0.34	0.50
0.40	0.38	0.47	0.53	0.24	0.60	0.53	0.25	0.60	0.53	0.25	0.60
					Tanzania						
0.07	0.84	0.10	0.06	0.83	0.10	0.06	0.83	0.10	0.06	0.83	0.10
0.15	0.71	0.20	0.13	0.69	0.20	0.13	0.69	0.20	0.13	0.69	0.20
0.27	0.53	0.35	0.22	0.56	0.30	0.21	0.56	0.30	0.22	0.57	0.30
0.36	0.43	0.44	0.30	0.44	0.40	0.31	0.45	0.40	0.30	0.44	0.40
0.42	0.35	0.51	0.40	0.34	0.50	0.40	0.34	0.50	0.40	0.34	0.50
0.48	0.30	0.56	0.50	0.24	0.60	0.51	0.25	0.60	0.51	0.25	0.60
					Zambia						
0.15	0.74	0.19	0.14	0.72	0.20	0.14	0.72	0.20	0.15	0.72	0.20
0.28	0.57	0.34	0.23	0.59	0.30	0.23	0.60	0.30	0.23	0.60	0.30
0.28	0.57	0.34	0.32	0.48	0.40	0.32	0.49	0.40	0.32	0.49	0.40
0.28	0.57	0.34	0.42	0.39	0.50	0.42	0.39	0.50	0.43	0.40	0.50
0.60	0.23	0.67	0.52	0.29	0.60	0.52	0.28	0.60	0.53	0.29	0.60

Notes: The table reports inclusion and exclusion out-of-sample errors based on the optimal models from Table 9. Inclusion errors are defined as the proportion of non-stunted children labelled as being stunted. Exclusion errors are defined as the proportion of stunted children labelled as being non-stunted. The budget is the proportion of total children reached. The assumption is that every child is either given some fixed amount of support or is not provided anything.

Figure 1

Geographic Scope of the Analysis

Figure 2

Height-Age Z-Score*100 Distributions Across Wealth Quintiles

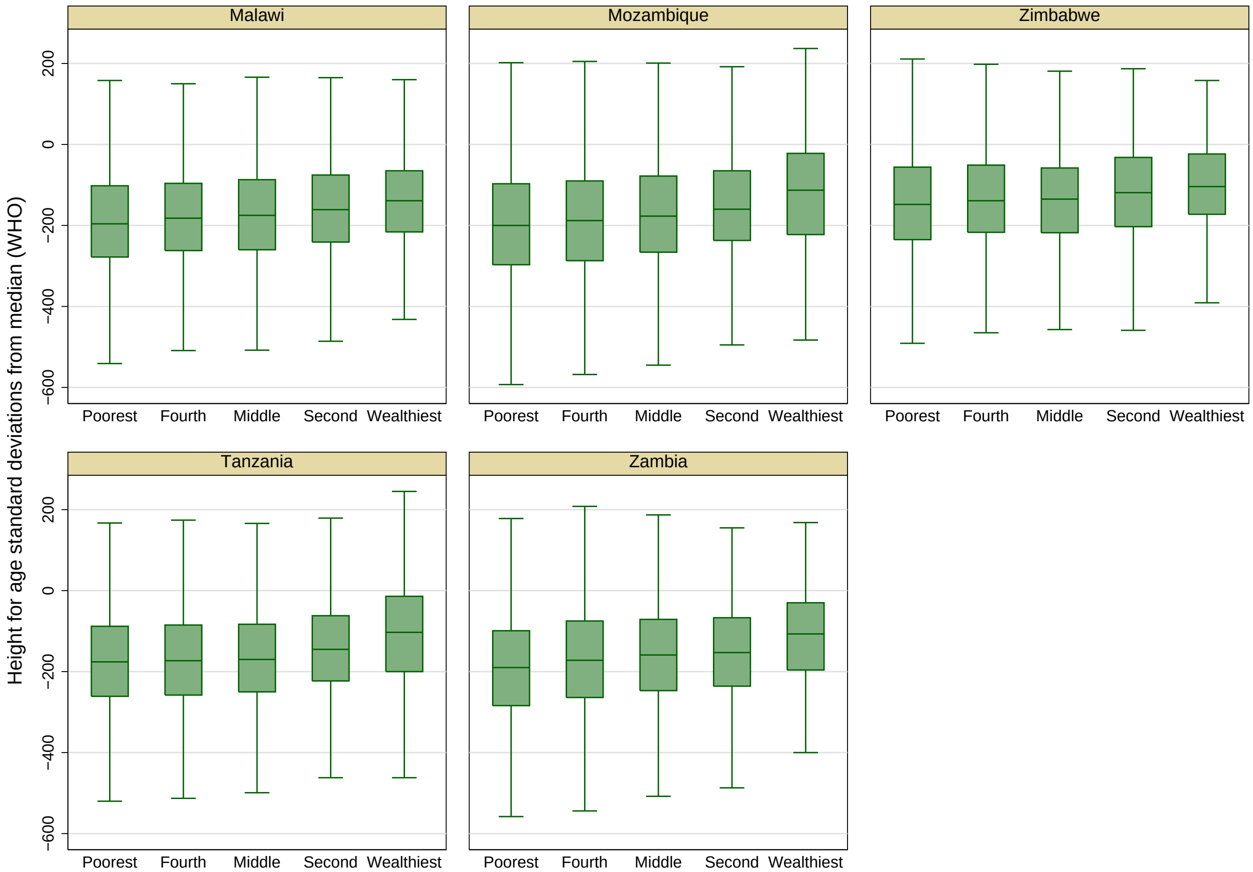


Figure 3-a

Drought-Contingent Height-Age Distributions: Lowest Wealth Quintile

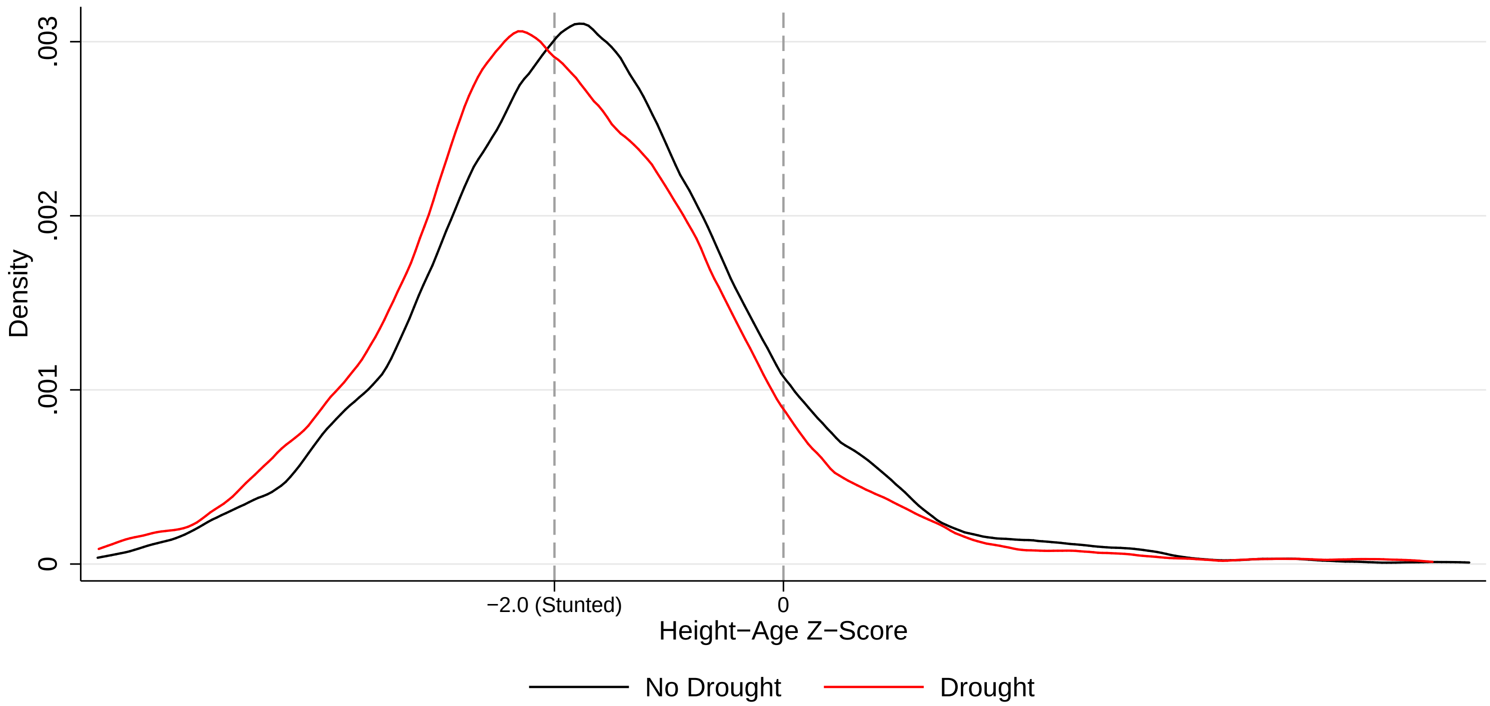


Figure 3-b

Drought-Contingent Height-Age Distributions: Top 4 Wealth Quintiles

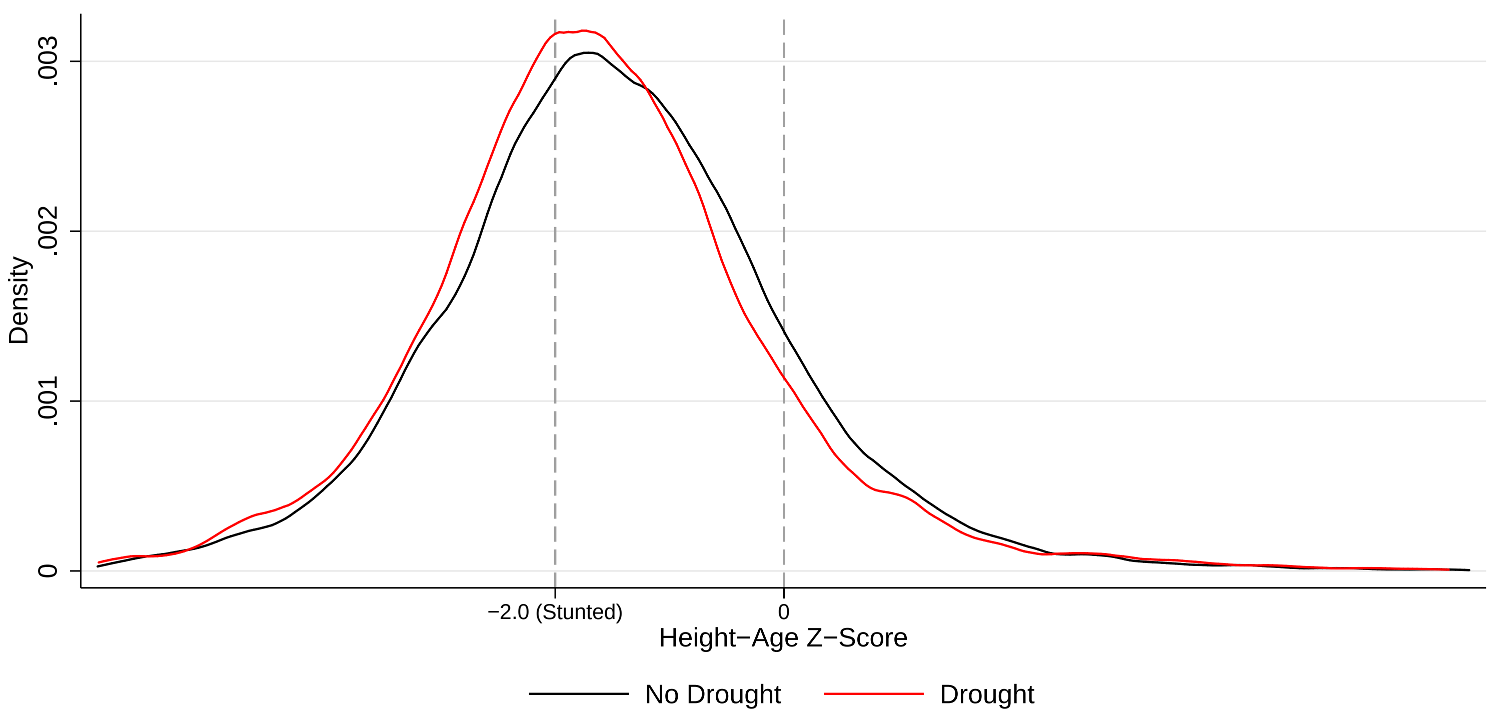


Figure 4: Probability of Child Stunting in Rural Mozambique

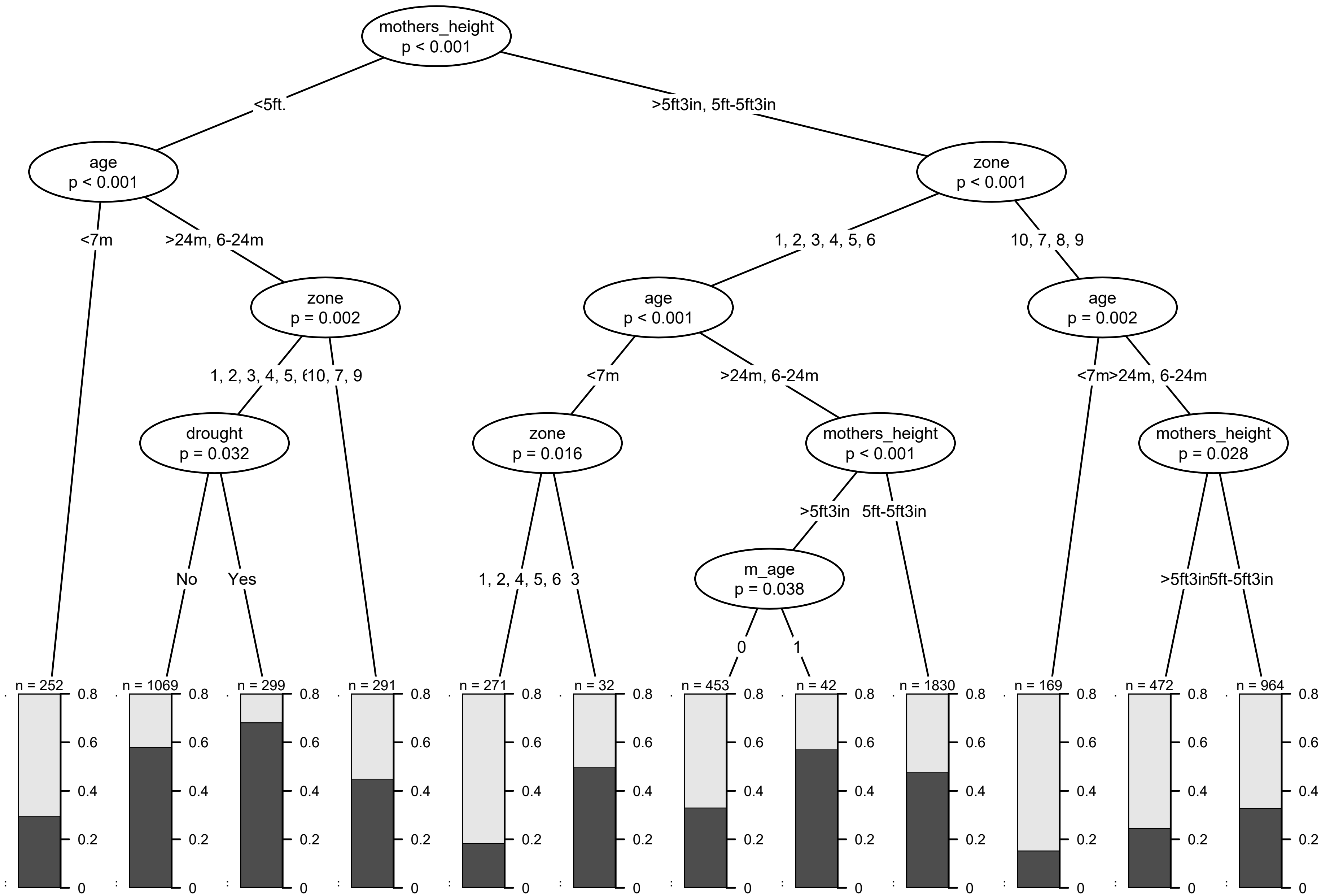
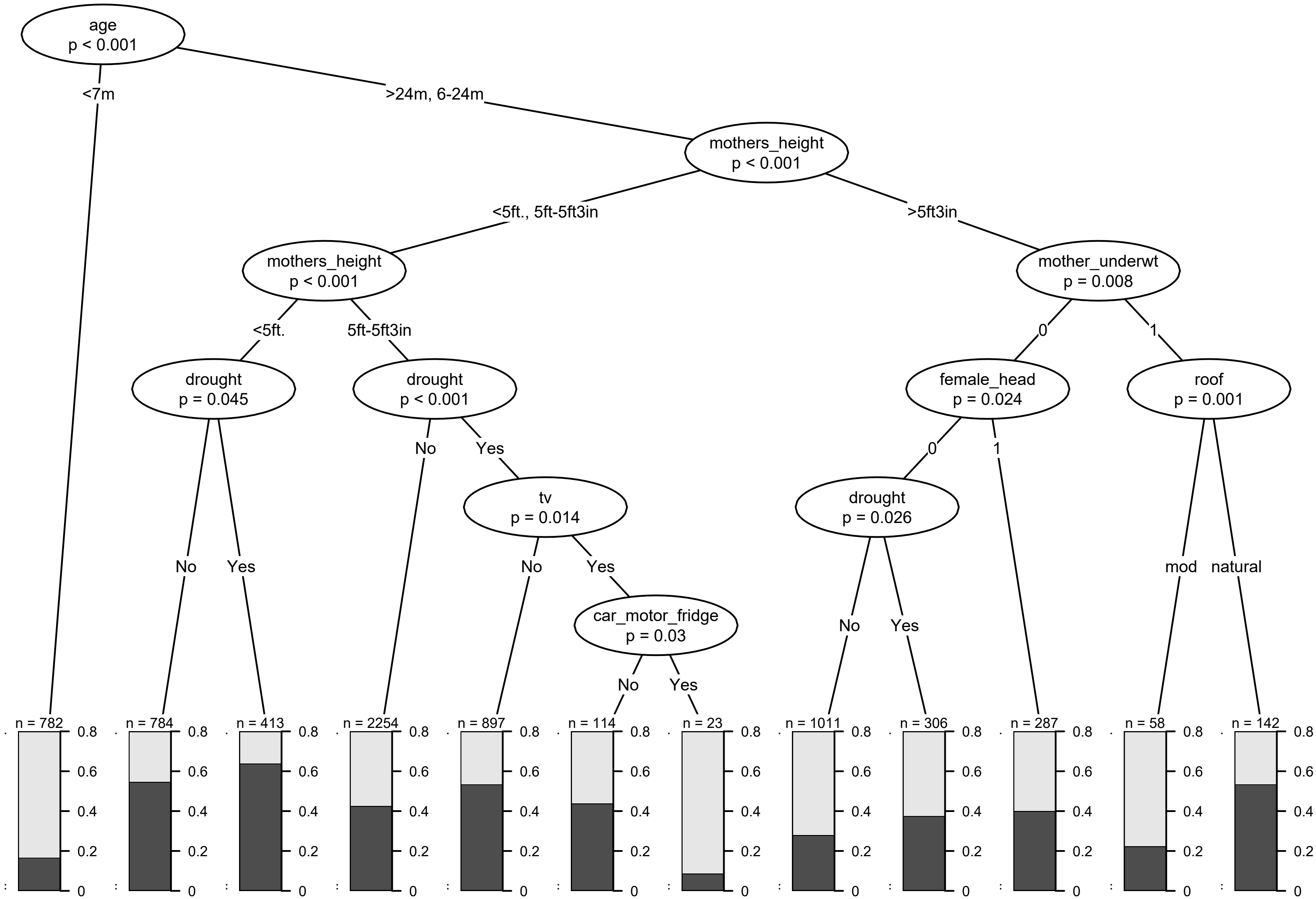


Figure 5: Probability of Child Stunting in Rural Zambia



References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias, "Targeting the poor: evidence from a field experiment in Indonesia," *The American economic review*, 2012, *102* (4), 1206–1240.

Alderman, Harold, "Safety nets can help address the risks to nutrition from increasing climate variability," *the Journal of Nutrition*, 2009, *140* (1), 148S–152S.

_, John Hoddinott, and Bill Kinsey, "Long term consequences of early childhood malnutrition," *Oxford economic papers*, 2006, *58* (3), 450–474.

Alfani, Federica, Andrew Dabalen, Peter Fisker, and Vasco Molini, "Vulnerability to stunting in the West African Sahel," *Food Policy*, 2019, *83*, 39–47.

Assaf, Shireen, Monica T Kothari, and Thomas W Pullum, *An assessment of the quality of DHS anthropometric data, 2005-2014*, ICF International, 2015.

Baez, Javier E, German Caruso, and Chiyu Niu, *Extreme Weather and Poverty Risk: Evidence from Multiple Shocks in Mozambique*, Economics of Disasters and Climate Change, Forthcoming.

Baffes, John and Varun Kshirsagar, "Shocks to Food Market Systems: A Network Approach," *Agricultural Economics*, forthcoming.

_, _, and Donald Mitchell, "What drives local food prices? Evidence from the Tanzanian maize market," *The World Bank Economic Review*, 2017, *33* (1), 160–184.

Barrett, Christopher B, "Measuring Food Insecurity," *Science*, 2010, *327* (5967), 825–828.

Barron, Jennie, Johan Rockström, Francis Gichuki, and Nuhu Hatibu, "Dry spell analysis and maize yields for two semi-arid locations in east Africa," *Agricultural and forest meteorology*, 2003, *117* (1-2), 23–37.

Baumol, William J and Richard E Quandt, "Rules of thumb and optimally imperfect decisions," *The American economic review*, 1964, *54* (2), 23–46.

Becker-Reshef, Inbal, Chris Justice, Mark Sullivan, Eric Vermote, Compton Tucker, Assaf Anyamba, Jen Small, Ed Pak, Ed Masuoka, and Jeff Schmaltz, "Monitoring global croplands with coarse resolution earth observations: The Global Agriculture Monitoring (GLAM) project," *Remote Sensing*, 2010, *2* (6), 1589–1609.

Beegle, Kathleen, Luc Christiaensen, Andrew Dabalen, and Isis Gaddis, *Poverty in a rising Africa*, The World Bank, 2016.

Berhane, Guush, John F Hoddinott, and Neha Kumar, *The impact of EthiopiaâĂŹs Productive Safety Net Programme on the nutritional status of children: 2008–2012*, Vol. 1604, Intl Food Policy Res Inst, 2017.

Boyle, Elizabeth Heger, Miriam King, and Matthew Sobek, *IPUMS-Demographic and Health Surveys: Version 7 [dataset]*, Minnesota Population Center and ICF International, 2019.

Brown, Caitlin, Martin Ravallion, and Dominique van de Walle, "A poor means test? Econometric targeting in Africa," *Journal of Development Economics*, 2018, *134*, 109–124.

Brown, Caitlin S, Martin Ravallion, and Dominique van de Walle, "Are Poor Individuals Mainly Found in Poor Households? Evidence using Nutrition Data for Africa," Technical Report, National Bureau of Economic Research 2017.

Brown, Molly E, *Food security, food prices and climate variability*, Routledge, 2014.

_ **and Varun Kshirsagar**, "Weather and international price shocks on food prices in the developing world," *Global Environmental Change*, 2015, *35*, 31–40.

Burke, Marshall and David B Lobell, "Satellite-based assessment of yield variation and its determinants in smallholder African systems," *Proceedings of the National Academy of Sciences*, 2017, *114* (9), 2189–2194.

Caire, Dean and Mark Schreiner, "Cross-Tab Weighting for Retail and Small-Business Scorecards in Developing Markets," 2013.

Castañeda, Andrés, Dung Doan, David Newhouse, Minh Cong Nguyen, Hiroki Uematsu, João Pedro Azevedo et al., "A new profile of the global poor," *World Development*, 2018, *101*, 250–267.

Clarke, Daniel J and Stefan Dercon, *Dull Disasters? How planning ahead will make a difference*, Oxford University Press, 2016.

Coady, David, Margaret Grosh, and John Hoddinott, *Targeting of transfers in developing countries: Review of lessons and experience*, The World Bank, 2004.

Coffey, Diane, Angus Deaton, Jean Drèze, Dean Spears, and Alessandro Tarozzi, "Stunting among children: Facts and implications," *Economic and Political Weekly*, 2013, *48* (34), 68–70.

Cooper, Matthew W, Molly E Brown, Stefan Hochrainer-Stigler, Georg Pflug, Ian McCallum, Steffen Fritz, Julie Silva, and Alexander Zvoleff, "Mapping the effects of drought on child stunting," *Proceedings of the National Academy of Sciences*, 2019, *116* (35), 17219–17224.

Damania, Richard, Sébastien Desbureaux, Marie Hyland, Asif Islam, Scott Moore, Aude-Sophie Rodella, Jason Russ, and Esha Zaveri, *Uncharted waters: The new economics of water scarcity and variability*, The World Bank, 2017.

de Waal, Alex, *Mass Starvation. The History and Future of Famine*, Cambridge University Press, 2018.

Deaton, Angus, "Measuring Development: Different Data, Different Conclusions?," *Proceedings of the 8th AFD-EUDN Conference, Paris*, 2010.

_ , "Measuring and understanding behavior, welfare, and poverty," *American Economic Review*, 2016, *106* (6), 1221–43.

— **and Nancy Cartwright**, "Understanding and misunderstanding randomized controlled trials," *Social Science & Medicine*, 2018, *210*, 2–21.

— **and Salman Zaidi**, *Guidelines for Constructing Consumption Aggregates for Welfare Analysis* number 135, World Bank Publications, 2002.

Dercon, Stefan, John Hoddinott, and Tassew Woldehanna, "Shocks and Consumption in 15 Ethiopian Villages, 1999–2004," *Journal of African Economies*, 2005, *14* (4), 559–585.

Devereux, Stephen and Jonas Nzabamwita, "Social Protection, Food Security and Nutrition in Six African Countries," 2018.

— **, Edoardo Masset, Rachel Sabates-Wheeler, Michael Samson, Althea-Maria Rivas, and Dolf Te Lintelo**, "The targeting effectiveness of social transfers," *Journal of Development Effectiveness*, 2017, *9* (2), 162–211.

Diamond, Alexis, Michael Gill, Miguel Rebolledo Dellepiane, Emmanuel Skoufias, Katja Vinha, and Yiqing Xu, *Estimating poverty rates in target populations: An assessment of the simple poverty scorecard and alternative approaches*, The World Bank, 2016.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani, *The elements of statistical learning* 2nd Ed., second ed., Springer series in statistics Springer 2012 reprint, Berlin, 2001.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell et al., "The climate hazards infrared precipitation with stationsâĂŤa new environmental record for monitoring extremes," *Scientific data*, 2015, *2*, 150066.

— **, Shraddhanand Shukla, Wassila Mamadou Thiaw, James Rowland, Andrew Hoell, Amy McNally, Gregory Husak, Nicholas Novella, Michael Budde, Christa Peters-Lidard et al.**, "Recognizing the Famine Early Warning Systems Network (FEWS NET): Over 30 Years of Drought Early Warning Science Advances and Partnerships Promoting Global Food Security," *Bulletin of the American Meteorological Society*, 2019, (2019).

Gilligan, Daniel O, John Hoddinott, and Alemayehu Seyoum Taffesse, "The impact of Ethiopia's Productive Safety Net Programme and its linkages," *The journal of development studies*, 2009, *45* (10), 1684–1706.

Grosh, Margaret E and Judy L Baker, *Proxy means tests for targeting social programs: simulations and speculation*, The World Bank, 1995.

Hallegatte, Stephane, Adrien Vogt-Schilb, Mook Bangalore, and Julie Rozenberg, *Unbreakable: building the resilience of the poor in the face of natural disasters*, World Bank Publications, 2016.

Headey, Derek and Giordano Palloni, "Water, sanitation, and child health: evidence from subnational panel data in 59 countries," *Demography*, 2019, *56* (2), 729–752.

Hijmans, Robert, Julian Kapoor, John Wieczorek, Nel Garcia, Aileen Maunahan, Arnel Rala, and Alex Mandel, "GADM database of global administrative areas (v3.6)," *Retrieved April*, 2019.

Hill, Ruth and Carolina Mejia-Mantilla, *With a little help: shocks, agricultural income, and welfare in Uganda*, The World Bank, 2017.

— **and Catherine Porter**, "Vulnerability to Drought and Food Price Shocks," 2016.

— **and Habtamu Fuje**, "What is the impact of drought on prices? Evidence from Ethiopia," Technical Report, Mimeo 2017.

— **, Emmanuel Skoufias, and Barry Maher**, "The Chronology of a Disaster: A Review and Assessment of the Value of Acting Early on Household Welfare," 2019.

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis, "Unbiased recursive partitioning: A conditional inference framework," *Journal of Computational and Graphical statistics*, 2006, *15* (3), 651–674.

Jayachandran, Seema and Rohini Pande, "Why are Indian children so short? The role of birth order and son preference," *American Economic Review*, 2017, *107* (9), 2600–2629.

Jin, Zhenong, George Azzari, Marshall Burke, Stephen Aston, and David Lobell, "Mapping smallholder yield heterogeneity at multiple scales in Eastern Africa," *Remote Sensing*, 2017, *9* (9), 931.

Johnson, David M, "An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States," *Remote Sensing of Environment*, 2014, *141*, 116–128.

Karlan, Dean and Bram Thuysbaert, "Targeting ultra-poor households in Honduras and Peru," *The World Bank Economic Review*, 2016.

Kshirsagar, Varun, Jerzy Wieczorek, Sharada Ramanathan, and Rachel Wells, "Household poverty classification in data-scarce environments: a machine learning approach," *arXiv preprint arXiv:1711.06813*, 2017.

Kudamatsu, Masayuki, Torsten Persson, and David Strömberg, "Weather and infant mortality in Africa," 2012.

Larsen, Anna Folke, Derek Headey, and William A Masters, "Misreporting month of birth: Diagnosis and implications for research on nutrition and early childhood in developing countries," *Demography*, 2019, *56* (2), 707–728.

Lentz, EC, H Michelson, K Baylis, and Y Zhou, "A data-driven approach improves food insecurity crisis prediction," *World Development*, 2019, *122*, 399–409.

Linderholm, Hans W, "Growing season changes in the last century," *Agricultural and forest meteorology*, 2006, *137* (1-2), 1–14.

Loveland, TR, BC Reed, JF Brown, DO Ohlen, Z Zhu, LWMJ Yang, and JW Merchant, "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data," *International Journal of Remote Sensing*, 2000, *21* (6-7), 1303–1330.

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

Maxwell, Daniel G and Kirsten Gelsdorf, *Understanding the Humanitarian World*, Routledge, 2019.

McBride, Linden and Austin Nichols, "Retooling poverty targeting using out-of-sample validation and machine learning," *The World Bank Economic Review*, 2016, p. lhw056.

Murphy, Kevin P, *Machine Learning: a Probabilistic Perspective*, The MIT Press, 2012.

Ninno, Carlo Del and Bradford Mills, *Safety Nets in Africa: Effective Mechanisms to Reach the Poor and Most Vulnerable*, The World Bank, 2015.

Pape, Utz and Philip Wollburg, *Impact of Drought on Poverty in Somalia*, The World Bank, 2019.

Premand, Patrick and Pascale Schnitzer, *Efficiency, legitimacy and impacts of targeting methods: Evidence from an experiment in Niger*, The World Bank, 2018.

Rutstein, Shea O, "Steps to constructing the new DHS Wealth Index," *Rockville, MD: ICF International*, 2015.

Schreiner, Mark, "A simple poverty scorecard for the Philippines," *Philippine Journal of Development*, 2007, *34* (2), 43.

—, "A simple poverty scorecard for India," *report to the Microcredit Summit Campaign, Washington, DC, http://www. microfinance. com/English/Papers/Scoring_Poverty_India. pdf*, 2008.

Sen, Amartya, *Commodities and capabilities*, Oxford University Press, 1999.

Sengupta, Somini, "Hotter, Drier, Hungrier: How Global Warming Punishes the Worlds Poorest," *The New York Times*, 2018, *12.*

Shongwe, Mxolisi E, GJ Van Oldenborgh, BJJM Van Den Hurk, B De Boer, CAS Coelho, and MK Van Aalst, "Projected changes in mean and extreme precipitation in Africa under global warming. Part I: Southern Africa," *Journal of Climate*, 2009, *22* (13), 3819–3837.

Sivakumar, MVK, "Predicting rainy season potential from the onset of rains in Southern Sahelian and Sudanian climatic zones of West Africa," *Agricultural and Forest Meteorology*, 1988, *42* (4), 295–305.

—, "Empirical analysis of dry spells for agricultural applications in West Africa," *Journal of Climate*, 1992, *5* (5), 532–539.

Skoufias, Emmanuel, Katja Vinha, and Ryoko Sato, *All Hands on Deck: Reducing Stunting through Multisectoral Efforts in Sub-Saharan Africa*, World Bank Publications, 2019.

Sohnesen, Thomas Pave and Niels Stender, "Is Random Forest a Superior Methodology for Predicting Poverty? An Empirical Assessment," *Poverty & Public Policy*, 2017, *9* (1), 118–133.

Trenberth, Kevin E, Aiguo Dai, Gerard Van Der Schrier, Philip D Jones, Jonathan Barichivich, Keith R Briffa, and Justin Sheffield, "Global warming and changes in drought," *Nature Climate Change*, 2014, *4* (1), 17.

Tucker, Compton J, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment*, 1979, *8* (2), 127–150.

Tversky, Amos and Daniel Kahneman, "Judgment under uncertainty: Heuristics and biases," *science*, 1974, *185* (4157), 1124–1131.

Victora, Cesar G, Linda Adair, Caroline Fall, Pedro C Hallal, Reynaldo Martorell, Linda Richter, Harshpal Singh Sachdev, Maternal, Child Undernutrition Study Group et al., "Maternal and child undernutrition: consequences for adult health and human capital," *The lancet*, 2008, *371* (9609), 340–357.

WHO Multicentre Growth Reference Study Group and others, "WHO Child Growth Standards based on length/height, weight and age.," *Acta paediatrica (Oslo, Norway: 1992). Supplement*, 2006, *450*, 76.

Wineman, Ayala, Nicole M Mason, Justus Ochieng, and Lilian Kirimi, "Weather extremes and household welfare in rural Kenya," *Food security*, 2017, *9* (2), 281–300.

World Bank, *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*, World Bank, Washington DC, 2018.

APPENDIX

Table A1: Drought Impacts on Child Anthropometrics (Compare with Table 3 in the Main Text)
Robustness Check for Child Age Mismeasurement:
Excluded Children: a) Born in Jan/Feb/Nov/Dec b) 12, 24, 36, 48 months old, Includes Child Age/Change Age Sq.
OLS with Cluster x Month Fixed Effects and Full Controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Stunted**	**Underweight**	**Wasted**	**Stunted & Underwt**	**Any**
Drought: Birth Season	0.05***	0.03***	0.01	0.04***	0.06***
	[3.26]	[2.96]	[1.29]	[4.19]	[4.83]
Drought: Season after Birth	0.00	-0.00	-0.01	0.01	0.01
	[0.01]	[-0.14]	[-0.97]	[0.59]	[0.55]
Mother ht between 152 and 160 cm	0.12***	0.07***	0.01*	0.05***	0.11***
	[7.95]	[6.77]	[1.80]	[6.98]	[9.02]
Mother ht below 152 cm	0.26***	0.13***	0.01	0.11***	0.23***
	[14.29]	[9.14]	[1.15]	[9.49]	[14.39]
Mother under 19 yrs	0.06**	0.03*	-0.01	0.02	0.04**
	[2.56]	[1.94]	[-0.43]	[1.26]	[2.20]
Child Age	0.02***	0.00***	-0.00***	0.00***	0.01***
	[16.51]	[3.18]	[-3.50]	[7.20]	[17.11]
Child Age Squared	-0.00***	-0.00***	0.00**	-0.00***	-0.00***
	[-15.25]	[-3.10]	[2.17]	[-10.47]	[-25.41]
No Livestock	0.03**	0.00	0.00	0.01	0.03**
	[2.08]	[0.13]	[0.24]	[0.83]	[2.10]
Does not treat water	0.04**	0.01	-0.01*	0.02*	0.02*
	[2.42]	[1.19]	[-1.88]	[1.93]	[1.69]
Mother Underweight	0.03	0.09***	0.03**	0.06***	0.07***
	[1.64]	[4.76]	[2.28]	[4.46]	[3.92]
Mother able to read part of sentence	0.03	-0.00	-0.01	0.00	0.02
	[1.23]	[-0.13]	[-0.72]	[0.32]	[0.98]
Mother cannot read at all/Unknown	-0.00	0.01	0.01*	0.00	0.01
	[-0.15]	[0.86]	[1.91]	[0.09]	[1.03]
Constant	-0.15***	-0.13***	0.07***	-0.12***	0.02
	[-3.20]	[-4.06]	[3.10]	[-4.69]	[0.56]
Observations	18,408	18,488	18,372	23,460	23,460
R-squared	0.47	0.42	0.40	0.41	0.48

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. All regressions use all available data on rural households from 5 Southern African countries (Mozambique, Malawi, Tanzania, Zambia and Zimbabwe). Additional controls are for roof, floor and wall quality, cooking fuel, employment in agriculture, source of water, type of toilet, ownership of radio, tv, bicyles, refridgerators, cars, trucks, birth order, birth interval, gender of child and the gender of the household head.

Table A2: Drought Impacts on Child Anthropometrics (Compare with Table 5 in the Main Text)
Robustness Check for Child Age Mismeasurement:
Excluded Children: a) Born in Jan/Feb/Nov/Dec b) 12, 24, 36, 48 months old, Includes Child Age/Change Age Sq.
OLS with Province and Year of Survey Fixed Effects and Full Set of Controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	**Stunted**	**Underweight**	**Wasted**	**Stunted & Underwt**	**Any**
3 Decad Dry Spell: Birth Season	0.03***	0.02**	0.01	0.02***	0.05***
	[2.99]	[2.34]	[1.20]	[3.24]	[5.49]
3 Decad Dry Spell: Season after Birth	0.01	-0.00	-0.00	0.00	0.03***
	[1.09]	[-0.27]	[-0.01]	[0.17]	[2.80]
Mother ht between 152 and 160 cm	0.11***	0.05***	0.01*	0.04***	0.10***
	[12.42]	[8.22]	[1.79]	[8.90]	[12.62]
Mother ht below 152 cm	0.25***	0.11***	0.01**	0.09***	0.21***
	[21.53]	[11.68]	[2.15]	[12.40]	[20.42]
Mother under 19 yrs	0.05***	0.01	-0.01	0.01	0.03**
	[3.06]	[1.24]	[-0.74]	[1.14]	[2.53]
Child Age	0.02***	0.00***	-0.00***	0.00***	0.01***
	[26.10]	[5.42]	[-5.49]	[12.86]	[24.76]
Child Age Squared	-0.00***	-0.00***	0.00***	-0.00***	-0.00***
	[-23.20]	[-4.50]	[3.20]	[-17.90]	[-37.67]
Lean Season	0.04***	0.01*	-0.00	0.01*	0.03***
	[4.19]	[1.75]	[-1.12]	[1.82]	[3.17]
No Livestock	0.02**	0.01	0.01	0.01	0.02***
	[2.52]	[0.96]	[1.46]	[1.06]	[2.89]
Does not treat water	0.02**	0.01*	0.00	0.01*	0.02**
	[1.97]	[1.82]	[1.22]	[1.69]	[2.51]
Mother Underweight	0.07***	0.10***	0.02***	0.08***	0.08***
	[5.10]	[7.90]	[2.63]	[7.32]	[6.44]
No Car, Motor, Fridge	0.04***	0.03***	0.00	0.02**	0.04***
	[2.63]	[2.86]	[0.71]	[2.15]	[3.05]
Mother able to read only part of sentence	0.03**	0.02	0.00	0.01	0.03**
	[2.18]	[1.42]	[0.37]	[1.51]	[2.20]
Mother cannot read at all/Unknown	0.03***	0.02***	0.01**	0.02***	0.03***
	[3.10]	[2.96]	[2.09]	[2.84]	[3.70]
Constant	-0.19***	-0.13***	0.05**	-0.13***	-0.04
	[-3.91]	[-4.45]	[2.00]	[-5.65]	[-1.00]
Observations	26,912	26,993	26,878	31,977	31,977
R-squared	0.12	0.05	0.03	0.06	0.16

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ province and survey year fixed effects. Refer to the Table 3 (in the main text) notes for all other details.

Table A3: Drought Impacts on Child Anthropometrics (Compare with Table 3 (1) in the Main Text)
Robustness to the Definition of a Dry Spell : Length 2 Decads, 3 Decads (as in main Text), 4 Decads and 5 Decads
OLS with Cluster x Month Fixed Effects and Full Controls

	(1)	(2)	(3)	(4)
VARIABLES	Likelihood of Stunting			
2 Decad Dry Spell: Birth Season	0.04***			
	[2.87]			
2 Decad Dry Spell: Season after Birth	0.07***			
	[5.38]			
3 Decad Dry Spell: Birth Season		0.07***		
		[5.43]		
3 Decad Dry Spell: Season after Birth		0.04***		
		[3.07]		
4 Decad Dry Spell: Birth Season			0.05**	
			[2.57]	
4 Decad Dry Spell: Season after Birth			0.03	
			[1.61]	
5 Decad Dry Spell: Birth Season				0.02
				[0.95]
5 Decad Dry Spell: Season after Birth				0.04
				[1.43]
Constant	-0.10**	-0.09**	-0.08**	-0.08**
	[-2.47]	[-2.25]	[-2.12]	[-2.10]
Observations	28,017	28,017	28,017	28,017
R-squared	0.46	0.46	0.46	0.46

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. Refer to the Table 3 (in the main text) notes for all other details.

Table A4: Drought Impacts on Child Anthropometrics (Compare with Table 5 (1) in the Main Text)
Robustness to the Definition of a Dry Spell : Length 2 Decads, 3 Decads (as in main Text), 4 Decads and 5 Decads
OLS with Province and Survey Year Fixed Effects and Full Controls

	(1)	(2)	(3)	(4)
VARIABLES	Likelihood of Stunting			
2 Decad Dry Spell: Birth Season	0.02***			
	[2.60]			
2 Decad Dry Spell: Season after Birth	0.06***			
	[6.84]			
3 Decad Dry Spell: Birth Season		0.04***		
		[4.53]		
3 Decad Dry Spell: Season after Birth		0.04***		
		[4.33]		
4 Decad Dry Spell: Birth Season			0.03**	
			[2.24]	
4 Decad Dry Spell: Season after Birth			0.04***	
			[3.02]	
5 Decad Dry Spell: Birth Season				0.02
				[1.25]
5 Decad Dry Spell: Season after Birth				0.04**
				[2.27]
Constant	-0.03	-0.03	-0.01	-0.01
	[-0.66]	[-0.74]	[-0.24]	[-0.21]
Observations	40,817	40,817	40,817	40,817
R-squared	0.10	0.10	0.10	0.10

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ province and survey year fixed effects. Refer to the Table 3 (in the main text) notes for all other details.

Table A5: Drought Impacts on Child Stunting (OLS with Cluster x Month Fixed Effects)
Compare with Table 6a in the Main Text
Robustness Test: Omitted children in households that moved residence within the last 6 years

	(1)	(2)	(3)	(4)	(5)	(6)
	All	Mozambique	Malawi	Zimbabwe	Tanzania	Zambia
Drought: Birth Season	0.07***	0.07***	-0.05	0.13***	0.06**	0.14***
	[4.96]	[2.64]	[-1.14]	[4.07]	[2.53]	[4.15]
Drought: Season after Birth	0.03**	0.02	0.10**	0.02	0.03	0.05
	[2.15]	[0.47]	[2.32]	[0.47]	[1.16]	[1.39]
Mother ht between 152 and 160 cm	0.13***	0.13***	0.13***	0.10***	0.14***	0.09***
	[9.13]	[4.66]	[3.27]	[4.73]	[6.53]	[3.22]
Mother ht below 152 cm	0.25***	0.24***	0.27***	0.25***	0.28***	0.19***
	[14.79]	[7.45]	[5.31]	[6.45]	[10.44]	[4.97]
Mother under 19 yrs	0.06**	-0.00	-0.03	0.02	0.10***	-0.02
	[2.48]	[-0.12]	[-0.45]	[0.48]	[2.67]	[-0.27]
Age:7-24 months	0.17***	0.15***	0.09	0.16***	0.17***	0.17***
	[8.99]	[3.95]	[1.47]	[4.57]	[6.15]	[4.09]
Age:25-59 months	0.21***	0.19***	0.13**	0.21***	0.23***	0.13***
	[12.07]	[5.34]	[2.40]	[6.75]	[8.90]	[2.99]
No Livestock	0.01	-0.01	0.02	-0.06*	0.02	0.05
	[0.78]	[-0.37]	[0.40]	[-1.93]	[0.75]	[1.59]
Does not treat water	0.03**	0.02	-0.02	0.03	0.04*	0.01
	[2.01]	[0.44]	[-0.60]	[0.89]	[1.87]	[0.29]
Mother underweight	0.05**	0.11**	0.03	0.10**	0.02	0.08*
	[2.34]	[2.43]	[0.52]	[2.19]	[0.65]	[1.86]
Mother able to read part of sentence	0.00	0.04	0.06	0.02	-0.01	0.04
	[0.24]	[0.78]	[0.92]	[0.39]	[-0.44]	[0.92]
Mother cannot read at all/Unknown	0.00	0.03	0.03	0.03	-0.01	0.05
	[0.17]	[1.04]	[0.63]	[0.72]	[-0.69]	[1.59]
Constant	-0.06	-0.00	-0.17	-0.16	-0.11	0.15
	[-1.26]	[-0.01]	[-1.08]	[-1.51]	[-1.46]	[1.30]
Observations	21,951	5,002	2,486	3,379	7,857	3,227
R-squared	0.47	0.44	0.52	0.49	0.48	0.50

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are clustered at the DHS cluster level. All regresssions employ cluster x month fixed effects. The weights used are person weights adjusted for the national populations. The first regression is the same as the first regression in table 3. Additional controls are for roof, floor and wall quality, cooking fuel, employment in agriculture, source of water, type of toilet, ownership of radio, tv, bicyles, refridgerators, cars, trucks, birth order, birth interval, gender of child and the gender of the household head. All metrics are based on children being measured below -2 standard deviation thresholds.

Table A6: A Comparision of Out-of Sample Predictive Accuracy (Compare with Table 9)

We introduce additive noise (drawn from a 0 mean and 1 standard deviation normal distribution) into the Height-for-Age standardized measure for the training set. This results in 22.6 percent of the stunted households classified as not stunted and 18.2 percent of the non-stunted households classified as stunted.

Method	Optimal Hyper-Parameter Values (via cross-validation)	Log-Loss
	Mozambique	
Logistic Regression	None	0.640
Conditional Inference Decision Trees	Level of Significance = 0.95	0.653
Random Forest	Min. Node Size = 200 , Number of Var Tried = 6	0.642
Gradient Boosted Trees	eta = 0.05 , max depth = 3, num rounds = 150, min child wt = 20	0.643
	Malawi	
Logistic Regression	None	0.635
Conditional Inference Decision Trees	Level of Significance = 0.90	0.646
Random Forest	Min. Node Size = 200 , Number of Var Tried = 9	0.636
Gradient Boosted Trees	eta = 0.05 , max depth =2, num rounds = 300, min child wt = 60	0.633
	Zimbabwe	
Logistic Regression	None	0.592
Conditional Inference Decision Trees	Level of Significance = 0.95	0.597
Random Forest	Min. Node Size = 300 , Number of Var Tried = 10	0.592
Gradient Boosted Trees	eta = 0.05 , max depth =2, num rounds = 150, min child wt = 10	0.591
	Tanzania	
Logistic Regression	None	0.616
Conditional Inference Decision Trees	Level of Significance = 0.95	0.644
Random Forest	Min. Node Size = 150 , Number of Var Tried = 8	0.617
Gradient Boosted Trees	eta = 0.4 , max depth = 1, num rounds = 300, min child wt = 20	0.616
	Zambia	
Logistic Regression	None	0.645
Conditional Inference Decision Trees	Level of Significance = 0.90	0.655
Random Forest	Min. Node Size = 100 , Number of Var Tried = 5	0.640
Gradient Boosted Trees	eta = 0.1 , max depth = 2, num rounds = 100, min child wt = 10	0.644

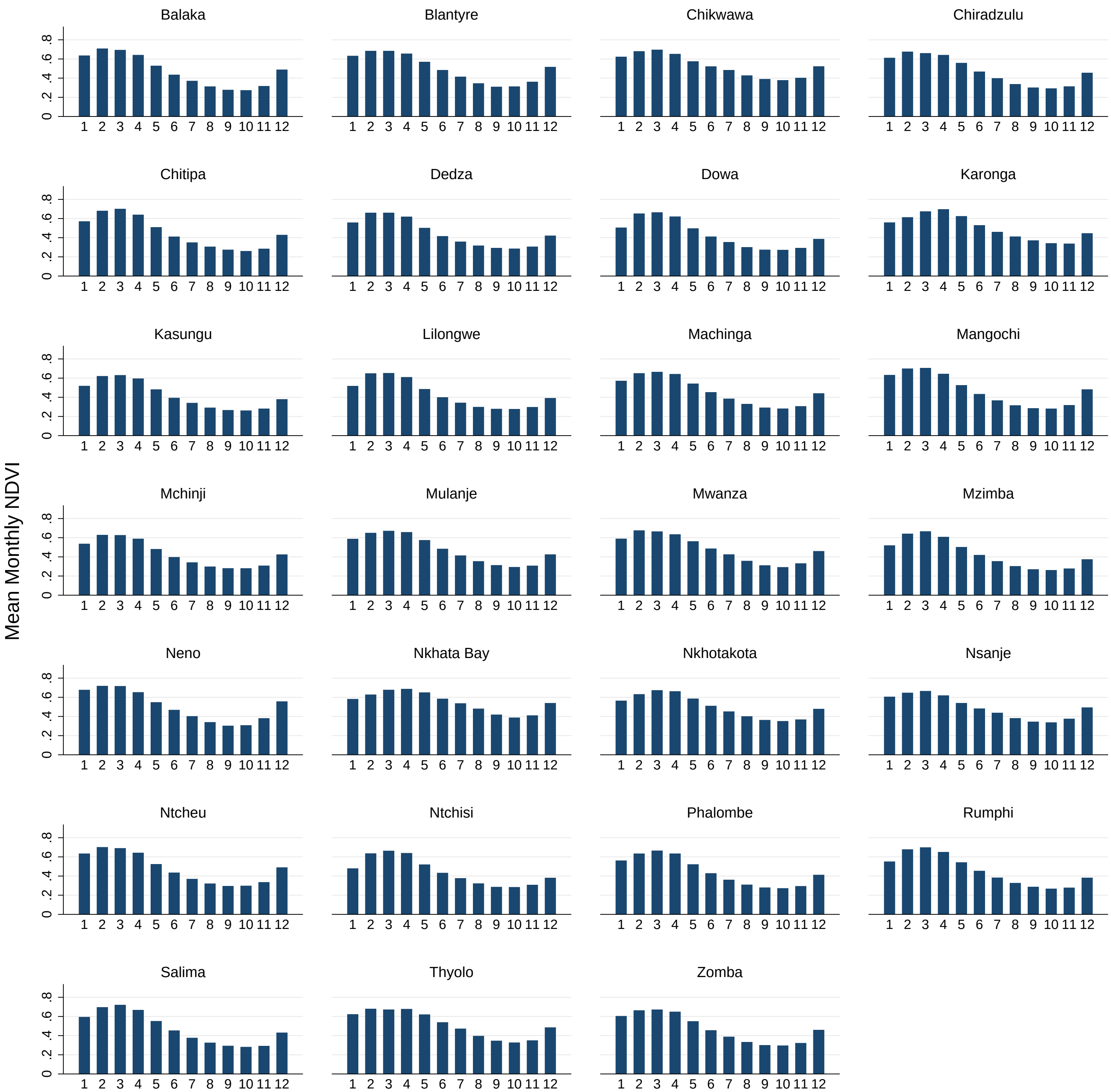
Notes: The models were trained on 70 percent of each set of surveys and tested on the remaining 30 percent. The hyper-parameters were optimized using 10-fold cross-validation (log-loss metric). The table reports the optimized (i.e. minimum) log-loss for each model. The hyperparameter candidate sets are as follows. Conditional Inference Decision Trees: Bonferroni level of significance (0.90, 0.95). Random Forests (R package Ranger): minimum node size (50, 100, 150, 200, 250, 300, 350), number of variables tried at each node (4, 5, 6, 7, 8, 9,10, 11, 12, 13, 14), and number of trees (1000). Gradient Boosted Trees (R package Xgboost): shrinkage (0.05, 0.1, 0.2, 0.3, 0.4), max depth (1, 2, 3, 4), number of rounds (50, 100, 150, 200, 250, 300) and min. child weight (10, 20, 30, 40, 50, 60).

Table A7: A Comparision of Out-of Sample Predictive Accuracy (Compare with Table 9)
Robustness Test: We use the HAZ score instead of the stunted dummy as the target (predicted) variable

Method	Optimal Hyper-Parameter Values (via cross-validation)	RMSE
	Mozambique	
Linear Regression	None	147.9
Conditional Inference Decision Trees	Level of Significance = 0.90	149.6
Random Forest	Min. Node Size = 100 , Number of Var Tried = 7	147.4
Gradient Boosted Trees	eta = 0.05 , max depth = 3, num rounds = 200, min child wt = 30	148.3
	Malawi	
Linear Regression	None	137.0
Conditional Inference Decision Trees	Level of Significance = 0.95	138.4
Random Forest	Min. Node Size = 200 , Number of Var Tried = 9	136.6
Gradient Boosted Trees	eta = 0.05 , max depth =2, num rounds = 300, min child wt = 50	136.8
	Zimbabwe	
Linear Regression	None	139.4
Conditional Inference Decision Trees	Level of Significance = 0.9	140.1
Random Forest	Min. Node Size = 25 , Number of Var Tried = 5	139.4
Gradient Boosted Trees	eta = 0.3 , max depth =1, num rounds = 150, min child wt = 60	139.3
	Tanzania	
Linear Regression	None	127.9
Conditional Inference Decision Trees	Level of Significance = 0.95	130.4
Random Forest	Min. Node Size = 50 , Number of Var Tried = 5	127.9
Gradient Boosted Trees	eta = 0.4 , max depth = 1, num rounds = 200, min child wt = 50	128.1
	Zambia	
Linear Regression	None	154.8
Conditional Inference Decision Trees	Level of Significance = 0.90	155.7
Random Forest	Min. Node Size = 100 , Number of Var Tried = 6	154.6
Gradient Boosted Trees	eta = 0.1 , max depth = 2, num rounds = 300, min child wt = 10	154.4

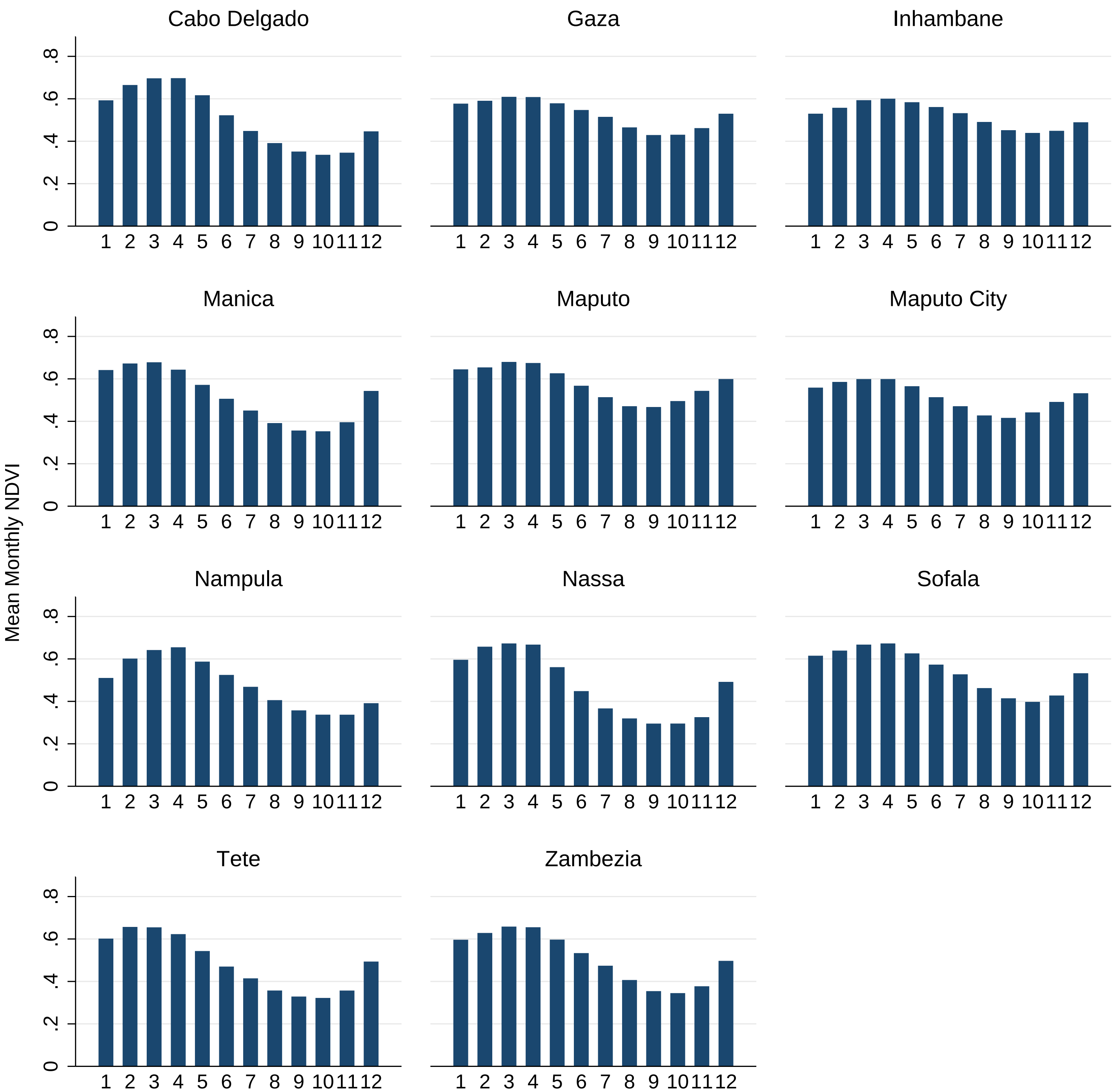
Notes: The models were trained on 70 percent of each set of surveys and tested on the remaining 30 percent. The hyper-parameters were optimized using 10-fold cross-validation (RMSE metric). The table reports the optimized (i.e. minimum) RMSE for each model. The hyperparameter candidate sets are as follows. Conditional Inference Decision Trees: Bonferroni level of significance (0.90, 0.95). Random Forests (R package Ranger): minimum node size (50, 100, 150, 200, 250, 300, 350), number of variables tried at each node (4, 5, 6, 7, 8, 9,10, 11, 12, 13, 14), and number of trees (1000). Gradient Boosted Trees (R package Xgboost): shrinkage (0.05, 0.1, 0.2, 0.3, 0.4), max depth (1, 2, 3, 4), number of rounds (50, 100, 150, 200, 250, 300) and min. child weight (10, 20, 30, 40, 50, 60).

Appendix Figure 1a
Harvest Cycles Across Districts: Malawi

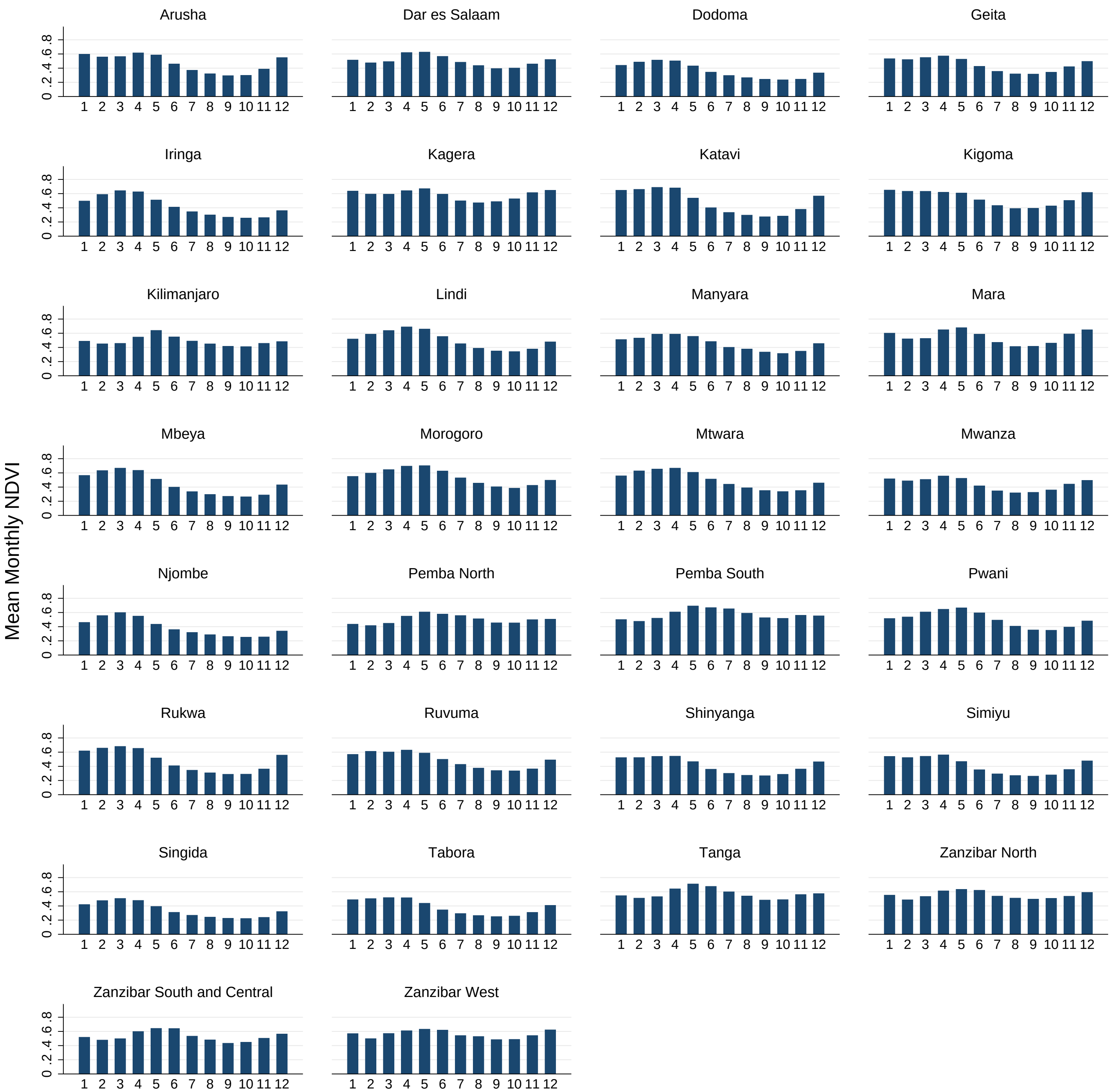


Appendix Figure 1b

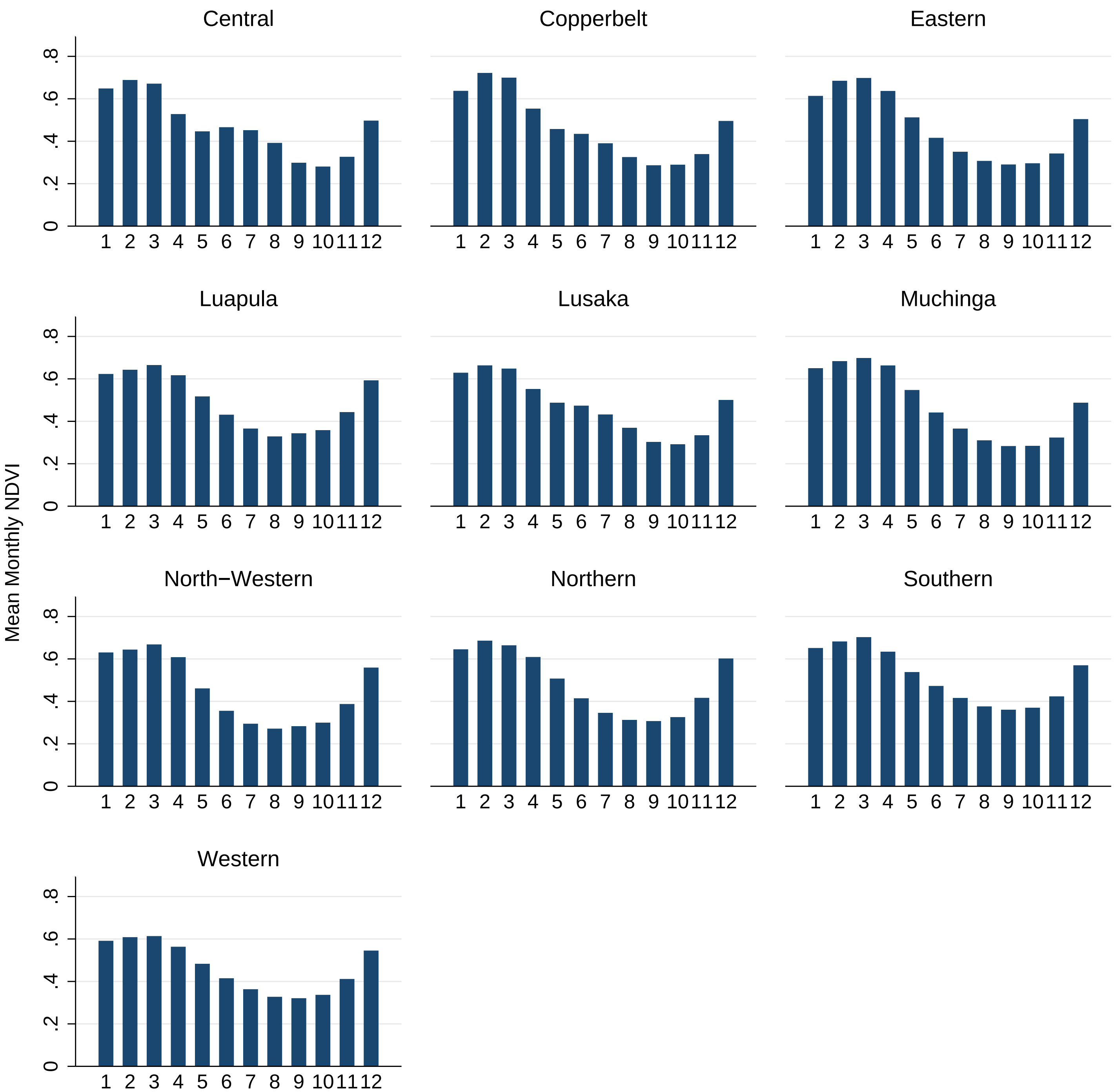
Harvest Cycles Across Provinces: Mozambique



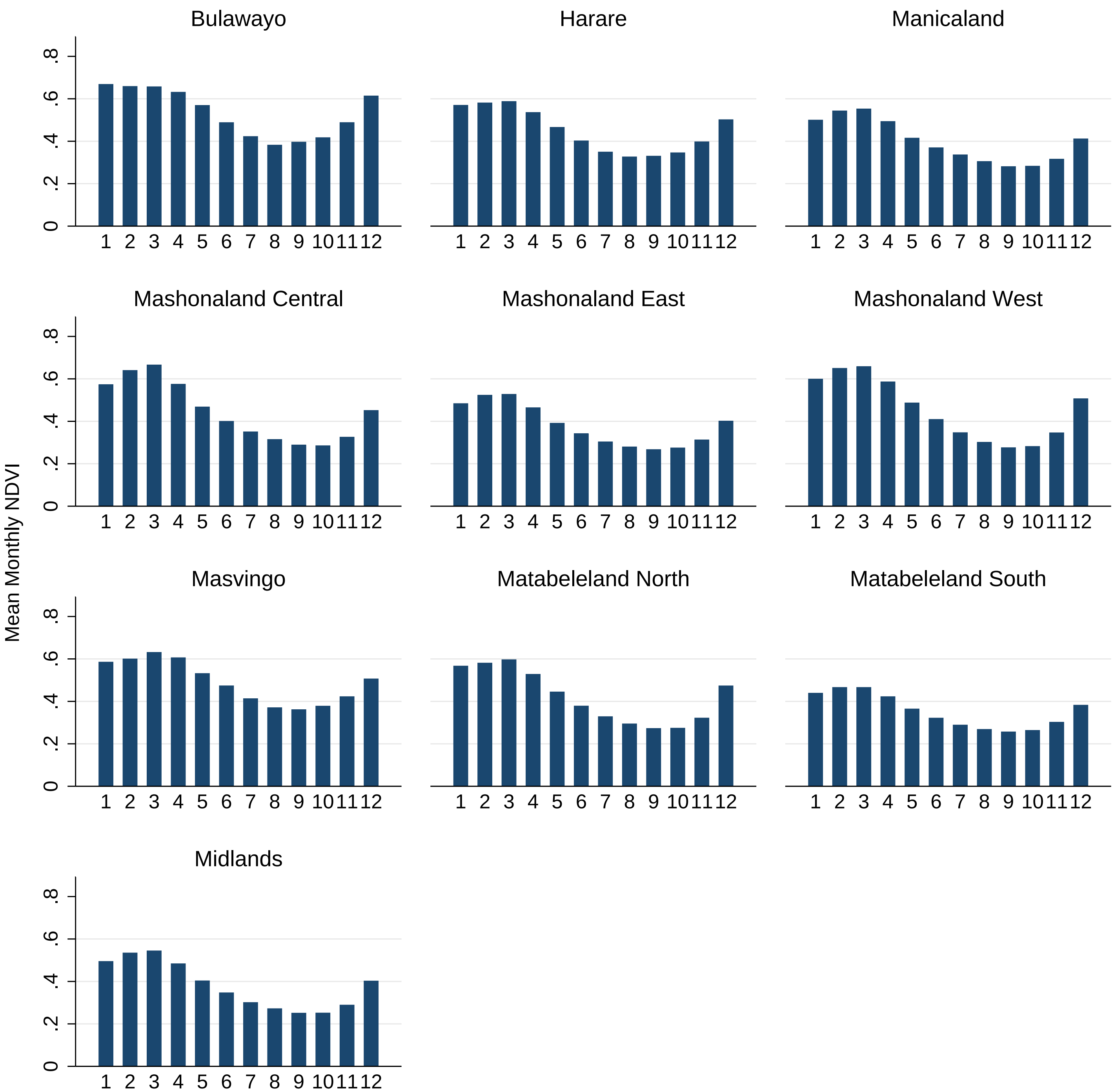
Appendix Figure 1c
Harvest Cycles Across Regions: Tanzania



Appendix Figure 1d
Harvest Cycles Across Provinces: Zambia



Appendix Figure 1e
Harvest Cycles Across Provinces: Zimbabwe



Appendix Figure 2
HDI Comparisons in the Broader African Context