

A Data-Driven Approach Improves Food Insecurity Crisis Prediction

Authors

E. C. Lentz^{1*}, H. Michelson², K. Baylis², Y. Zhou^{2†}

Affiliations

1 Lyndon B. Johnson School of Public Affairs, University of Texas at Austin, Austin TX 78713-8925.

2 Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana IL 61801.

* Corresponding authors: Erin Lentz, Lyndon B. Johnson School of Public Affairs, UT Austin PO BOX Y, Austin TX 78713-8925; +1-512-232-8353; erinclentz@utexas.edu; ORCID: 0000-0001-9934-9496. Hope Michelson, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 1301 W. Gregory Drive, Urbana IL 61801; +1-217-300-9669; hopecm@illinois.edu. Kathy Baylis, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, 1301 W. Gregory Drive, Urbana IL 61801; +1-217-244-6653; baylis@illinois.edu

A Data-Driven Approach Improves Food Insecurity Crisis Prediction

Abstract

Globally, over 800 million people are food insecure. Current methods for identifying food insecurity crises are not based on statistical models and fail to systematically incorporate readily available data on prices, weather, and demographics. As a result, policymakers cannot rapidly identify food insecure populations. These problems delay responses to mitigate hunger. We develop a replicable, near real-time model incorporating spatially and temporally granular market data, remotely-sensed rainfall and geographic data, and demographic characteristics. We train the model on 2010-2011 data from Malawi and forecast 2013 food security. Our model correctly identifies the food security status of 83 to 99% of the most food insecure village clusters in 2013, depending on the food security measure, while the prevailing approach correctly identifies between 0 and 10%. Our results show the power of modeling food insecurity to provide early warning and suggest model-driven approaches could dramatically improve food insecurity crisis response.

Keywords: food insecurity, crisis, prediction, early warning, Sub-Saharan Africa, famine

A Data-Driven Approach Improves Food Insecurity Crisis Prediction

Highlights

1. Faster humanitarian response during food crises saves lives and resources.
2. We present a new, transparent and replicable model-driven method to predict the onset of food crises across the world.
3. We apply the model to Malawi, a country with persistent chronic and acute food security problems.
4. Leveraging readily available data, our model substantially improves over the status quo global methods of prediction.
5. Our best forecasts predict the out-of-sample food security status for between 65 to 88 percent of village clusters.
6. Our analysis and results demonstrate the potential gains to real-time food security modeling.
7. We discuss the data and data integration necessary to take our methods to scale.

A Data-Driven Approach Improves Food Insecurity Crisis Prediction

I. Introduction

The global food insecure population is greater than 800 million and rising (FAO, 2017). Currently, international food crisis assessments are not based on statistical models. Instead, assessments use a convergence-of-evidence methodology, in which local stakeholders make infrequent projections based on available data. We envision and propose the following: a set of country-specific models combining readily available data to produce high-frequency, sub-national predictions of food security status. These predictions would complement the existing tools used to assess global food crises. Model predictions would inform status quo methods to support a more effective deployment of assessment resources by identifying particular regions in need of more careful scrutiny or by documenting locations where circumstances are deteriorating quickly. The research we present here is a first step towards achieving this larger objective of developing standardized methods that analysts can use to supplement the current approach, the Integrated Food Security Phase Classification System (IPC), to identify food crises. A lack of near-term, sub-national predictions delays effective response, and governments, nongovernmental organizations, and other donors face considerable challenges allocating scarce resources to mitigate hunger (Headey and Barrett, 2015). Thus, improving forecasts can quickly and directly improve the circumstances of affected populations.

Our model uses only readily accessible secondary data to predict sub-national food security, facilitating its application to other countries and contexts and avoiding primary data collection, a major limitation faced by other proposed estimation approaches (Headey and Barrett, 2015). Using Malawi as a test case, a country with frequent food insecurity challenges, we estimate and forecast three measures of food security as functions of remotely-sensed measures of weather shocks, market prices, and demographic data related to food access. We

propose our model as a transparent, replicable, and intuitive contribution to a redesigned global early warning system. Incorporating model-driven predictions into such a system would enhance and hasten humanitarian response and reduce the potential for political manipulation (de Waal, 2018).

The most widely accepted definition of food security includes a hierarchy of components: availability, access, and utilization; and one cross-cutting dynamic component, stability (Webb et al., 2006). Food must be available for people to access it; target populations must have the logistical wherewithal to access and consume it. Current efforts to anticipate food security crises have focused on using detailed, remotely sensed data to predict local crop production quantities, a measure of availability (e.g., Lobell et al. 2008). While our research builds on these production-focused methods, to our knowledge, ours is the first to predict subnational food insecurity using readily-available data to capture availability, access, and stability in an integrated model. Because efforts focused exclusively on crop production also by definition focus exclusively on availability (Niles et al., 2017; Shively, 2017; Hidrobo et al., 2018), they commonly disregard critical allocation dynamics: how harvested food moves through heterogeneous geographies, infrastructures, and market systems so that households can access it. As economist Amartya Sen famously noted (Sen, 1982. Page 1):

“Starvation is the characteristic of some people not *having* enough food to eat. It is not the characteristic of there *being* not enough food to eat. While the latter can be a cause of the former, it is but one of many *possible* causes.”

The global current standard for early warning to guide emergency aid allocation and response, the IPC, uses information related to both availability and access, but is characterized by its own set of limitations and challenges. IPC classifications are made quarterly or semi-annually, which are infrequent relative to the rate at which food security can worsen on the ground. Moreover, because the IPC uses a Delphic method relying on a convergence of evidence

approach rather than a formal model (IPC, 2012), it has faced criticism that it is too complex, requires an excess of detailed information that has uneven availability, produces assessments that are difficult to replicate and confirm, and shows vulnerability to political influence (The Economist, 2017; de Waal, 2018). Finally, the IPC fails to make use of a full scope of available data, particularly readily available secondary data, in a replicable and transparent manner (IPC, 2012).

In short, although the last decade has seen a dramatic increase in the available quantity and quality of data related to food availability and access, including high-resolution measures of soil quality, rainfall, and prices, no current food security early warning and monitoring system systematically incorporates these data into a predictive model.

We use linear and log-linear models based on 2010 data from Malawi to estimate three food security outcome measures, and then use those models to forecast food security outcomes three years later. Our best forecasts accurately predict the out-of-sample food security status for between 65 and 88% of village clusters, depending on the food security measure. Critically important for the effective targeting of resources, while our analysis indicates that the IPC fails to predict any of the most food insecure village clusters in 2013, our model correctly predicts 83-99% of the most food insecure village clusters in the sample. Because our model uses readily available monthly data, our approach generates food security predictions in near real time, providing assessments at least two months earlier than standard IPC quarterly evaluations. We are focused on model cost, replicability and scalability and we identify and evaluate a series of model trade-offs across (1) spatial granularity, (2) availability of a range of predictors, and (3) errors of inclusion and exclusion.

II. Materials and methods

We use readily available data to model the 2010 food security status of village clusters in Malawi, for three, commonly-used food security measures. We model each of these measures as functions of weather and price fluctuations, accounting for geographic variation and intra-annual seasonal trends. We then evaluate the model's accuracy using out-of-sample prediction of 2013 data.

The guiding principle of our modeling approach is to employ data likely to be widely available at high spatial granularity and, when possible, high frequency, allowing rapid, real-time assessment of sub-national food security. In response to concerns about the “black box” nature of the IPC (The Economist, 2017; de Waal, 2018), our method is reproducible, transparent, and parsimonious.

We study model performance across different combinations of spatial scales and data classes to explore the relative improvement in prediction associated with increasing spatial granularity and adding more (and potentially more costly to access) data to the model. Increasing granularity and incorporating additional variables are likely to improve prediction and targeting but both also increase the cost of data processing and access. Therefore, we ask, what is the right spatial level for prediction? And which data seem to contribute the most? We run model specifications at three increasingly granular spatial scales and progressively include three classes of variables as predictors. These data classes range along a continuum: from widely available, remotely sensed data collected daily to information about household assets and demographics derived from infrequently fielded surveys.

Governments, donors and national and international agencies face tradeoffs regarding modeling choices: valuable funds and time spent gathering additional information to refine

targeting could, instead, be spent on providing rapid assistance to more people (Basu, 1996). Our interest is to understand the amount of additional precision gained from estimations made at increasingly spatially granular scales and from estimations incorporating variables requiring more resources to access and process. In our results, we highlight the impacts of increasing granularity and adding predictors on accuracy, and assess tradeoffs.

We begin by describing our food security measures and the three classes of data that we use in the analysis. Further details on measuring food security, matching data across spatial scales, and data sources are available in the Supplementary Materials (SM), sections S1, S2, and S3, respectively. No institutional review board approval was required; we used only anonymized, secondary data. The collecting agencies received informed consent for personal data.

We predict three measures of food security used by international humanitarian organizations including the US Agency for International Development (USAID) and the United Nations World Food Programme (WFP): the reduced coping strategies index (rCSI), the household dietary diversity score (HDDS) and the food consumption score (FCS). We use the 2010-11 Living Standards Measurement Survey (LSMS) for Malawi as our source for our dependent variables because the survey's sampling strategy generates household-level food security outcomes with good temporal variation and spatial granularity. We take the average at the village-level for each measure. Further details on the LSMS sample are in the SM, along with Table S1, which summarizes data sources and how the data were processed.

The three dependent variable measures capture different dimensions of food insecurity; rCSI measures household coping strategies and is associated with food quantity, while FCS and HDDS reflect dietary quality (Maxwell et al., 2014; Vaitla et al., 2017). The HDDS is a count of the number of food categories that a household consumes in a week. The FCS weights this

count of food groups according to their nutrient density. Higher values of both the FCS and the HDDS indicate higher food security. The rCSI reflects the number of coping strategies a household uses from a set of five universal strategies to address possible food shortages by frequency and severity; a higher rCSI indicates lower food security. Most often, governments and international agencies apply cut-offs to categorize food security status rather than use the continuous measures of food security (Vaitla et al., 2017).

To predict the three measures of food security: rCSI, FCS, and HDDS, we employ what we term Class 1, 2, and 3 data. These are three nested classes of variables with increasing processing requirements and decreasing availability. We compare predictions of the food security measures using these three data classes to a naïve prediction of cluster (or village-level) average food security using the IPC food security assessment as the only right-hand side variable. We refer to the IPC assessment as our ***Class 0 data***.

The IPC is an appropriate benchmark for our model. It is used to assess food security quarterly or semi-annually in 26 countries; in addition to these countries, a committee of 17 West African countries uses IPC protocols with minor differences to assess food security. In Malawi, an IPC assessment exists for each of Malawi's 60 IPC livelihood zones. The IPC livelihood zones - combinations of livelihood zones and administrative districts - are the highest spatial level of aggregation in our analysis. The IPC assessment data and livelihood zone boundaries are available from the Famine Early Warning System Network (FEWS Net).

IPC-trained teams answer a series of questions and sift through available evidence including food prices, anthropometric measures, and mortality rates, to reach convergence on assigning a food insecurity classification to each IPC zone on a scale of one through five (IPC, 2012). The IPC aims for consistency in assigning food security classifications across the countries

where it is regularly used. A challenge is that availability and quality of the evidence used to forecast food insecurity varies across and within countries and over time. As a result, the IPC assessing teams inevitably rely on some interpretation of the evidence.

Class 1 data are readily available data including precipitation, market prices, soil quality and geographic variables combined with the Class 0 IPC food security assessments. Class 1 data are generally collected remotely and are widely available, making them especially useful for difficult-to-reach areas at risk of food insecurity. Agricultural product prices are often readily available and collected in person by government agents, although these can also be collected using cellular phone technologies. We generate measures of agriculturally-relevant precipitation generated from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset (Funk et al., 2015). We use the total amount of precipitation that fell during the October–April rainy season. For the same season, we define the length of the longest dry spell as the number of continuous days with no rain. To measure the beginning of the rainy season, we calculate the number of days after October 1st in which rainfall greater than 10 mm fell 3 days out of 5 (Guan et al, 2015). These three variables are taken from the prior agricultural season to predict food availability for the June/July maize harvest. We include the maximum amount of rainfall in the current month to control for possible flooding for flood-prone regions, which can affect transportation and local economic outcomes. We also include quarter fixed effects to capture seasonality.

We include a monthly average of local market price data of maize (the primary staple cereal), collected weekly from the Malawi Statistical Division (MSD) of the Ministry of Agriculture. In addition to the average price of maize in the market by month, we generate a measure of the number of weeks per month in which the maize price is missing in the nearest market reporting data as a measure of potential local availability and market thinness (Mallory and

Baylis, 2016). Along with these high-frequency weather and market variables, we use measures of elevation, distance to market, distance to road, amount of agricultural land in a 1km radius and an indicator for limited soil nutrient retention as a measure of soil quality. These data come from the 2010-11 LSMS survey but would all be available from secondary sources in other contexts.

Class 2 data are data likely to be available but also likely to require additional work to be accessed and processed, such as data on household roof type (a coarse predictor of poverty) and cell phone ownership. For the Class 2 data, we draw from the LSMS but note that these could come from other sources. We include the percent of households who own a cell phone and the roof type (metal versus thatch), both of which could be collected from sources other than recent surveys, such as cellular companies and remote sensing.

Class 3 data consist of infrequently gathered but publicly available household-level data including demographics and assets. These data are often available in a census or large household survey such as the Demographic Health Surveys or LSMS. Data for our Class 3 variables also come from the Malawi LSMS surveys: demographic data, including the gender and age of the household head, number of household members, and asset holdings.

We emphasize that we extract only a minimal set of variables from the LSMS: food security, basic demographics, and assets. Information on basic demographics and roofing could also be available from a census or rapid assessment tool. We stress this point in order to emphasize that the LSMS is not fundamental to our strategy. Surveys collecting representative, geo-referenced food security data could also be used to train and validate the initial model. For example, parsimonious, rapid assessments such as SMART surveys or monitoring tools using a community reporter model (e.g., Knippenberg, Jensen and Constatas,

2018) would suffice. Ideally, an initial model would be trained and tested using existing secondary data (such as the LSMS) and then the model's real-time predictions would be validated against periodic high-frequency, household food security data from rapid assessments. These findings could be further incorporated into the model to update and improve the accuracy of forecasts over time.

We conduct the estimation at three spatial scales for two periods: in 2010, our sample is drawn from (a) IPC zone ($n=55$, which excludes 5 urban zones), (b) traditional authority (TA) level ($n=281$), and (c) cluster of villages, or cluster level ($n=736$); in 2013, our sample is (a) IPC zone ($n = 46$); TA level ($n=153$) and village cluster level ($n=204$). Class 1, 2, and 3 data are aggregated to the appropriate spatial scale for each spatial scale of estimation. For example, the IPC-district level estimations use the share of households in the IPC zone that own a cellular phone, the zone's average total precipitation during the prior agricultural season, etc. Similarly, the cluster level estimations use the share of households in the cluster that own a cellular phone, or the cluster's average total precipitation during the prior agricultural season.

For the Class 1 data, we develop a protocol to combine monthly data for 72 geocoded markets with seasonal rainfall data to cluster-geocoded annual demographic and asset data. We match markets with gridded rainfall data from the most recent agricultural season. For Class 2 and 3 data, we combine this information with household-level cluster-geocoded information to capture demographic and assets data. See the SM (section S2) for details and Fig. S1 for a graphical representation of the matching protocol.

We use these data in a linear model to predict average cluster-level HDDS and rCSI, and a log-linear model to predict log FCS. We log FCS in the analysis to place more weight on the

lower end of the distribution. We use these regressions in our analysis to enable the replicability of our work in other locations and to evaluate how our model relates to existing food security research; in particular we are interested in understanding the potential magnitude and source of bias in the results when applied out-of-sample.

For each cluster-level spatial scale, we estimate:

$$FS_{it} = b_{0t} + b_1 Class1_{jt} + b_2 Class2_{jt} + b_3 Class3_{jt} + e_{it}$$

Where FS is the cluster average food security measure (rCSI, HDDS, logFCS) for cluster i and month t . $Class 1 - 3$ data, described above, are included additively across specifications and are averages at spatial level j , where j is IPC zone, TA, or cluster. At the IPC zone and traditional authority level, we compute the cluster-level food security as predicted by variables calculated at the IPC zone and TA level, respectively. The $Class 0$ estimations only include lagged IPC assessments, which are computed only at the IPC zone.

We also apply a machine learning technique, lasso (least absolute shrinkage and selection operator) regressions, to our data. We chose the lasso technique for a few reasons. First, the lasso yields easily interpretable coefficients, which is less true for more complex techniques. We believe that interpretable coefficients are critical for policymakers interested in understanding and responding to the drivers of food security. Second, lasso techniques are sparse prediction functions, meaning if a particular variable does not contribute to the prediction, its coefficient will reduce to zero. This helps us understand whether the variables we have included in our main linear and log-linear regressions add value to our predictions.

III. Results

We begin by presenting descriptive statistics. We then present our main results: predictions of food insecurity in the 2013 out-of-sample data. We evaluate our model's responsiveness to sample variation by splitting the data by region and apply a lasso regression to our data, comparing findings between our linear and log-linear regression results and the machine-learning model. We then consider errors of inclusion and exclusion associated with our main predictions. We validate our main 2010 model. We examine the role of prices in our out-of-sample predictions and compare results against the IPC.

Descriptive statistics

Summary statistics of our food security measures and explanatory variables are presented in Table 1. The 2010 measures are from Malawi's 2010-11 LSMS, which surveyed 12,271 households over 12 months. The 2010-11 sample is representative at a national level for each month and, for the crop year, for each district. The 2013 measures used for the out-of-sample prediction are from the 2013 Malawi LSMS, which surveyed 3,999 households. Approximately 60% of our sample reports an rCSI of zero, indicating that the majority of households did not use any coping strategies associated with food insecurity during the survey period.

[Table 1 about here]

Food security varies substantially over time and space. Mean rCSI peaks in April in the Malawi 2010-11 data, and FCS and HDDS scores are at their lowest levels in February and March, Malawi's lean season (Fig. 1a). Food security measures also vary systematically over space, with low levels of food security as measured by rCSI in the Shire valley, the southern tip of Malawi (Fig. 1b). Following the poverty prediction literature, we use the cluster-level average of household food security measures as outcomes in our model to reduce within-

cluster heterogeneity (Hyman et al., 2005; Jean et al., 2016). A cluster is similar to a village or group of small villages. In the SM (Fig. S2), we present the degree of variation of our 2010 household food security measures explained only by month and/or geographic variation.

[Figure 1 about here]

Out-of-sample predictions: spatial variation and data availability classes

To generate out-of-sample predictions, we first fit our model to our 2010 Class 0, 1, 2 and 3 data. We retain the coefficients from the fitted model and apply them to the 2013 data. Results show how much of the actual variation in food security status in 2013 correctly predicted by the model.

We first present the share of the variance of the three 2013 food security outcomes explained by our estimation models, as measured by R-squared (Fig. 2). We plot the R-squared of the predicted versus actual food security on the vertical axis for models at different levels of spatial aggregation, listed across the horizontal axis. For example, the “IPC zone” level on the x-axis uses data processed at the highest level of spatial aggregation (agro-ecological zones within districts). Moving to the right, the same data are processed at more granular levels of aggregation (the traditional authority or TA, and the cluster level, a cluster of villages). We use a different color to represent each food security measure and a different shape to represent the data classes employed in each iteration of the model.

[Figure 2 about here]

Our initial estimation uses only the IPC measure as a predictor (Class 0 – represented by circles in Fig. 2B). Because this metric is only compiled at the IPC zone level, higher levels of

spatial variation do not increase the amount of variation and it explains very little: from 0.00 for HDDS, 0.005 for logFCS and 0.016 for rCSI. In contrast to the “base” IPC only model, when we incorporate other relevant factors that explain food security and increase the degree of the spatial disaggregation, moving from the IPC zone to the TA and then village cluster, the share of variation explained increases. This is particularly true for the HDDS and logFCS, and particularly for specifications that include some measure of assets (Class 2 and Class 3 data). The specification using only Class 1 data (geographic and price variables) does substantially worse than those using Class 2 data (which add measures of three assets).

Our best predictions employ the full complement of available data (Class 1, 2, and 3) and are conducted at the most spatially disaggregated level, the cluster. Our predictions for the two dietary diversity measures HDDS and FCS, are relatively strong: explaining 64 and 65% of the variation respectively (measured by R-squared). The rCSI prediction fares worse, only predicting 12% of the variation. In comparison to the out-of-sample IPC-only (Class 0) results, our predictions’ explanatory power are substantial improvements: the variation in each food security measure explain by our model relative to the IPC: is 64% vs. 0.0% for the HDDS; 65% vs. 0.5% for the log FCS; and 12% vs 1.6% for the rCSI. Across geospatial levels, we identify substantial gains in prediction accuracy for the Malawi model when we move from Class 1 to Class 2 data. Our predictions continue to improve, albeit by smaller amounts, moving from Class 2 to Class 3 data.

While our model explains more variation in food security as we increase the spatial granularity of the predictor variables (moving to the right along the x-axis), the marginal benefits of granularity are largest when moving from IPC-level to TA-level. The additional explanatory power gained when moving from TA to cluster is smaller. Further, spatial granularity matters more when we include Class 2 or Class 3 data than Class 1 data. Thus, if the predictive model

relies solely on measures like rainfall, price and soil quality, a larger geographic aggregation may suffice. Both increasing spatial granularity and incorporating relevant variables that are less easily collected improves prediction and therefore targeting, but will increase the cost and complexity of data collection, processing, and access.

Comparing our out-of-sample results to our 2010 results, we find the amount of explained variation is similar for HDDS and logFCS; while the 2013 models explain 64% and 65% of the variation, the 2010 models explain 91% and 62% respectively (See Table 2). Our out-of-sample results for the rCSI explain less than the in-sample results; in 2010 we explained 34% compared to only 12% in 2013.

In general, our models explain less variation in rCSI, likely reflecting the non-linearity and lower variation in rCSI scores. In both years, the rCSI has a median value of 0, indicating the median respondent did not engage in any coping strategies. In 2010-11, almost two-thirds of the rCSI observations are 0, which limits the observed variation available to train our model. The model is more effective at explaining the variation in food security for the more normally distributed measures of dietary diversity such as HDDS and FCS, than for the truncated rCSI measure.

A second approach to out-of-sample predictions is training our data on some regions and estimating out-of-sample predictions for the remaining regions. This approach allows us to understand how well our model applies across geography, while using both years of data for training. Splitting the sample by geography could be valuable in situations where data is collected in one part of the country but not others. Using data from the southern and central region to predict the north, we find that, while informative, the model performs worse when applied to a different geographies than when used to forecast future years (See Table S2 in

the SM for out-of-sample regional results). This result might be explained by the fact that the north has a slightly different agro-ecology and market structure, and thus applying coefficients estimated from the south and central regions does not fit the northern data as well. We do not recommend this approach for extremely remote, conflict-ridden, or cut-off regions, as they may face substantially greater risk of food insecurity compared to regions where data collection is easier. In what follows, we focus on the out-of-sample predictions that vary by time, given our primary interest in improving early warning forecasting.

We also estimate the model for each of the food security measures using a lasso regression. We include all of the variables to allow the lasso to identify which have the most explanatory power. Table 3 shows that the lasso performs reasonably well, particularly for logFCS and HDDS; the lasso predicts 64% of the variation in logFCS (compared to 65% in our main model), 61% in HDDS (compared to 64%), and 6% in rCSI (compared to 11%). The lasso largely selects the variables that we include in our Class 3 models and signs are similar on the estimated coefficients (See Table S3 in the SM for the model coefficients). In fact, for two of the food security measures, the lasso retains all possible variables, confirming that these variables add explanatory power to our base model. The lasso's prediction accuracy is slightly worse for forecasting 2013 than for the base model, suggesting the lasso model may be overfitting. In other words, the lasso may be embedding more information from 2010 than is useful for predicting food security in later years. In contrast, the lasso does better than the standard ordinary least squares (OLS) regression when predicting over region, suggesting the more flexible functional form may perform better when being used to estimate across different climactic and market settings.

Out-of-sample predictions: errors of inclusion and exclusion

In what follows, we focus on the results of the specification employing the richest set of explanatory variables conducted at the most disaggregated spatial scale, estimating the model at the cluster level aggregation and incorporating Class 1, 2, and 3 data. These are the results presented furthest to the right on the x-axis in Fig. 2 and represented by the plus signs. Fig. 3 shows a scatterplot of predicted values against actual values for each of the three food security measures. The 45-degree, diagonal lines in each figure indicate where predictions equal measured values.

[Figure 3 about here]

Our initial results, described above, focus on the amount of variation we explain with our model for continuous outcomes (i.e., the R-squared). However, practitioners arguably care more about how well we classify clusters by food security category (i.e., above or below a cut-off). Such cut-offs are often used by humanitarian agencies to target aid to village clusters suffering from food insecurity as measured by rCSI, log FCS, or HDDS, and therefore predicting cut-offs rather than continuous values may be more useful for these agencies. To assess the accuracy of our predictions, we compute Type I errors of inclusion (i.e., targeting those who do not need assistance) and Type II errors of exclusion (i.e., not targeting those who are insecure) using commonly used cut-offs for food security status. Details on the selection of food security cut-offs are included in the SM.

We quantify errors of inclusion and exclusion by calculating the percent of correct predictions, i.e., those household clusters correctly identified as food secure or insecure. We then calculate the percent of the clusters the model predicts to be in a *better* food security state than they are (an exclusion or Type II error) and the percent we predict to be in a *worse* food security state than they are (an inclusion or Type I error). We compare these numbers against

random allocation and the allocation one would obtain if one assumed all clusters were in the largest category. See SM Fig. S4 for the results of analyses conducted at other spatial scales.

We present the cut-offs of food security measures in the scatterplots (Fig. 3), using vertical lines to represent the actual category cut-offs and horizontal lines for the predicted. We shade and label areas of the graph to identify where food security is over-predicted (areas I and II) and under-predicted (areas III and IV). Clusters in area I, where we predict food insecure clusters as food secure, are errors of exclusion. Clusters in area IV, where we predict food secure clusters as food insecure (area IV), are errors of inclusion. Areas II and III indicate areas where the continuous food security measure is over or under-predicted, but the predicted category is the same as the actual.

Given the priority to correctly categorize households while minimizing the number of food-insecure households that we miss, our model performs quite well. For the HDDS in Fig. 3A vertical dotted lines at 3 and 6 indicate food security category cutoffs of low and medium dietary diversity (proxies for low and medium food security, respectively). Most cluster averages lie between these two categories. While our model tends to under-predict food security in the middle-range of the distribution (we predict 79% have lower than actual HDDS values and 21% have higher values than their true level), we correctly classify the food security category of 88% of the sample, outperforming the simple assumption of all households being in the largest category (medium dietary diversity), which would correctly categorize 76%. Similarly, we observe more type I than type II errors (11% versus than less than 1%); thus, our model misses very few clusters that are truly food insecure.

For logFCS, shown in Fig. 3B, the model slightly under-predicts food security for the majority of the sample (we predict 82% have lower FCS values than their true level and 18%

have higher). The model performs especially well for households in the lowest food security category in the 2013 data (“borderline”), correctly classifying 83% of the most food insecure clusters. Overall, we correctly predict 76% of the cluster categories, under-estimating the food security category of 23% of the sample (area IV), and over-estimating in less than 1% (area I). This result is much better than random, which would correctly categorize only one third of households. On the other hand, due to the fact that the cut-off levels used by humanitarian agencies are designed to capture relatively rare events, our prediction performs slightly worse than if we simply assumed that each cluster was food secure, on average, which would correctly categorize 91% of the clusters.

For rCSI, shown in Fig. 3C, our best model obtains 65% accuracy, under-predicting the food security status of 61% of households, and over-predicting food security for 39%. As with the HDDS, our best model does better than the naïve prediction that all clusters are food secure on average, which would correctly classify 54%, and considerably better than random allocation, which would only accurately predict the food security category of one quarter of the clusters.

In the SM we include density plots for the predicted and actual values (Fig. S3) and figures presenting the prediction accuracy for the full sample and the tail-only samples (Fig. S4). Relative to predictions using the full 2010 sample, predictions using only the food insecure population decreases errors of exclusion while increasing errors of inclusion, as one might expect. A humanitarian agency aiming to reach most of the food insecure population may prefer fewer errors of exclusion relative to errors of inclusion.

Validation of 2010 model

The 2013 data allow us to examine the accuracy of our predictive model, given the coefficients on the predictors generated from the 2010 data. Here we validate our model estimated using the 2010 data by asking whether our predictive variables affect food security as expected. We find that across all three food security measures, the signs and significance of included variables are consistent with theory (Barrett, 2010) and other empirical analyses (Jones et al., 2014; Shively, 2017). Cluster-level coefficient estimates are presented in Table 2 (Table S4 presents these estimates for the clusters in the tails). Holding other variables constant, greater precipitation, earlier start of the rainy season and shorter dry spells during the last agricultural season (for HDDS) and better soil quality improves current cluster average food security (likely through increased production, and thus greater food availability); measures of improved food access such as lower food prices, proximity to road and markets, cell-phone ownership and assets are all associated with greater food security.

[Table 2 about here]

In-sample, our most spatially granular estimations with Class 1, 2 and 3 data predict 62% of the variation in HDDS, 61% of the variation in logFCS and 34% of rCSI. The explanatory power is comparable to that obtained in cross-sectional studies that include fixed effects as well as highly detailed information on production and farm size (Jones et al., 2014). Yet, we note that models used in such studies are data intensive and not well-suited for prediction.

The role of price data

While the data available for early warning has expanded, key data remain missing, incomplete or difficult to access. In particular, high-frequency (e.g., monthly) price data collected from tertiary (rural, remote) markets are rare and there is little research to date on whether primary and secondary market prices are adequately integrated with more remote, rural market prices.

In Table 3, we show results using our high-frequency, high geo-spatial resolution prices, compared to (1) having no price data and (2) using the Global Information and Early Warning System (GIEWS) price data for Malawi, which includes monthly maize prices for six cities. In each model, we retain quarterly fixed effects, which help capture seasonal effects, such as pre-harvest hunger.

[Table 3 about here]

When relying on the GIEWS data, we find similar results for logFCS and HDDS, but the rCSI predictions are worse, reflecting that the rCSI is more sensitive to prices. This finding is even more pronounced when using only the quarterly fixed effects and no price data; we lose further accuracy in our prediction of rCSI, but we retain good predictive power for logFCS and HDDS (See Table S2 in the SM for results using an alternative specification with month fixed effects). These results suggest that food prices are particularly valuable for understanding rCSI in our Malawi model. Additional research would help us to understand whether focusing early warning data collection efforts on prices from a sample of tertiary markets could yield substantial improvements in early warning, and whether these improvements matter for CSI, for other food security measures, or whether relying on price data from primary markets is adequate.

Comparison to the IPC, the current approach to food insecurity early warning

A final and primary contribution of this work is validation of the IPC. For the four quarterly assessments that overlap with Malawian household survey data in 2010-11, we find that the IPC assessment (i.e., Class 0 model) is correlated with household-level measures of food insecurity. See Table 4. Among the three food security measures, the IPC is most closely related to measures of rCSI. The rCSI is believed to capture inadequate quantities of food

consumed, which is consistent with acute food insecurity, the focus of the IPC (de Waal, 2018). Overall, the IPC assessments anticipate the LSMS household food insecurity measures. That said, a significant constraint of the IPC is that while it may work relatively well on average, it does not capture the distribution of food security outcomes, and therefore may miss the food insecure clusters and households it is meant to identify.

[Table 4 about here]

We compare the sensitivity of predictions from our model against those of the IPC. Our sensitivity measure is the percentage of food insecure clusters correctly predicted to be food insecure (also known as a true positive rate) for our 2013 out-of-sample data. High sensitivity indicates greater accuracy in predicting food insecurity hotspots. In Table 5, we present this sensitivity measure for both our model and for a model using only IPC classification values. For FCS, the IPC value only model fails to identify any of the “borderline” food insecure clusters, compared to our model which identifies more than three quarters, resulting in a sensitivity of 83% (no clusters had average FCS scores that were in the poor category). Our model captures four times more of the clusters classified as food insecure based on the rCSI (86% versus 10%). Because the measure is relatively more discrete, and the majority of households are in the medium dietary diversity category, the HDDS results are similar for the two models (100% versus 99%).

[Table 5 about here]

IV. Discussion

Our approach makes several contributions, improving upon the current best practice in early warning predictions and food insecurity analysis. The model decreases the lag associated with

current early warning, helping to identify food insecure populations at least two months earlier than the IPC; faster responses save lives (Nikulkov et al., 2016; Gelli et al., 2017). It also targets food insecure communities at a more spatially granular level, better allocating scarce resources. By developing a model that relies on widely available, pre-existing, spatially disaggregated, and temporally frequent data, we ensure that our model can be applied to most countries with sizable food insecure populations and, depending on the method of price data collection, generate predictions in near-real time. We prioritize parsimony and modeling transparency with the objective of reproducibility, including to contexts where other data may be difficult to obtain, such as conflict zones, and the model could be expanded to incorporate data gathered quickly via remote-sensing or cell-phone technology. Thus, the method could be especially useful in areas where on-the-ground information is limited or entirely unavailable. By utilizing only easily accessible, secondary data to construct a transparent, replicable and intuitive early warning system, we not only can enhance humanitarian response but also address current concerns of politicization of early warning resulting, in part, from perceptions that current approaches are black boxes (The Economist, 2017).

Leveraging readily available data improves early warning. Our results demonstrate that a relatively small set of readily available, secondary data can effectively predict food security, and improve upon current best practice in early warning, the IPC. Our use of readily available data means that agencies, donors, and governments aiming to better predict food security could adopt our approach without additional on-the-ground data collection efforts. Note that while nearly all governments regularly collect the market prices of basic foodstuffs as data inputs into central bank inflation calculations in addition to other applications, these data are not always as widely collected and made available as they are in Malawi. Our research demonstrates the importance of such data for humanitarian applications.

By utilizing relatively simple regression techniques, we have shown that our model is consistent both across food security measures and at different spatial scales. Further, the variables are associated with food security outcomes as expected based on existing theory and research. More complex machine-learning based models may improve predictions, and remain an important avenue for future research. In related work, we apply additional machine-learning methods to this same prediction problem (Authors, 2018). In prediction, variable selection involves a trade-off between bias and variance. We could take an approach where we add a large set of variables, higher order effects and their interactions, which might reduce the prediction variance and increase the prediction bias. Yet, to adopt models without validating that their results are consistent with expectations runs the real risk of biased predictions, potentially harming the very people our models are intended to serve. For these reasons, the simple parametric model proposed here serves as a robust and interpretable alternative.

Frequency matters. This model can produce high-frequency predictions of sub-national food insecurity. People fall in and out of food security over the course of a year, but currently, standard data collection techniques largely miss such dynamics. Many food security surveys are outdated (for prediction) by the time the data are available. The IPC is assessed quarterly at most, sometimes semi-annually. The data we use in our model are available in near-real time, and the dynamic variables are higher frequency than the IPC; precipitation data are available daily and price data are released weekly or monthly. Even this one-month lag in our model provides information earlier and more frequently than currently available approaches.

Researchers and analysts desiring a model that more closely approximates real-time could use other techniques to collect prices, such as cell-phone based data collection, particularly in

regions where government price data collection does not exist. While cellular data is a promising way to increase data frequency, real methodological concerns remain, especially for the collection of household-level data. Potential sample biases in cellular phone survey data are not yet well understood (Blumenstock et al., 2015).

Spatial granularity matters. Our model predicts local food security status at a significantly more granular spatial level than what is currently available. Moreover, spatial granularity in the data can further improve prediction. In particular, when spatially detailed asset and demographic information are available, conducting the analysis at a more spatially granular scale substantially improves predicting power. If such data are not available and the predictive model is limited to spatially correlated Class 1 data measures like precipitation and market data, conducting the analysis at a larger geographic aggregation may suffice.

Targeting tradeoffs. Our model highlights the tradeoffs faced by food security analysts and policymakers balancing errors of inclusion and exclusion. Our results show that if the objective is to identify those clusters most at risk of food insecurity, analysts would do well to choose a measure that is normally distributed and has categories defined such that most of the measure is not in a single category, particularly if that category is defined as being food secure. As a second best, weighting observations by their food insecurity status to place more emphasis on the lower tail of the distribution can help. The cost is that this approach increases the number of potentially truly food secure households predicted to be food insecure (errors of inclusion). The severity of this drawback depends on the goal of and funding available for programming.

Policy implications and future work. Our approach is designed to be replicable in other locations by relying on readily available data and utilizing consistent protocols to incorporate

data at different spatial levels within a site. Across sites, each type of data is similarly scaled, enabling cross-site comparisons (Sachs et al., 2010). Our approach can be applied to locations that lack recent census or household survey information, such as conflict areas or those in the midst of another crisis, such as a natural disaster.

A limitation of our data is that Malawi did not experience any major shock during this time period. For this reason, we believe that our current model may best be used as a complement to the more Delphic approach used by the IPC that incorporates difficult-to-predict shocks such as conflict or catastrophic events. Predictions of these events remain an open and unresolved area of research (Bazzi et al., 2018). In applications of this model to places with long-term conflict, it may be sensible to include an indicator variable in the model for such regions of instability. Methodologically, such a spatial fixed effects approach to incorporating conflict presupposes that conflict, and its effect on food security, is time invariant. We also note that prices and market assessments provide some information about infrastructure disruptions and shortages likely to occur during conflict, natural disasters, mass migration, or other difficult-to-predict crises. Thus, our model could be expanded to incorporate emergent efforts to predict these events, such as tracking population movements through cell-phone tower usage, collecting data via cell-phones, or scraping reporting in newspapers (Blumenstock et al., 2015).

A second limitation is the degree to which Malawi can serve as a generalizable test case for extending this work into other countries. Two points are relevant with respect to the external validity of our work. First, we argue that the method we propose to integrate and analyze readily available data can be meaningfully applied in other countries. First, we do not intend for the coefficients from our Malawi model to be applied to other countries; the estimated coefficients are unlikely to be applicable in other settings, with different weather patterns and

market infrastructure. Malawi is a small country with chronic food security problems punctuated by crises and may provide limited insight into other areas facing different challenges. Instead, we intend the *approach* to be applicable in other contexts. A second concern is that Malawi may be distinct in its available relative abundance of secondary data. In particular, Malawi is distinguished by good quality price data collected reliably and weekly with high spatial granularity (72 markets in a relatively small area) and multiple rounds of LSMS data with more community observations per unit area than found in other LSMS surveys in the region. Application in other countries will require solving country-specific challenges related to identifying appropriate food security outcome data and price data to run and validate the model (e.g., using SMS or other innovative technologies). In Table S5 in the SM, we summarize data sources that could support extending this work elsewhere.

For these reasons, future work validating our models in other contexts is warranted. Our models are more effective at predicting HDDS and logFCS relative to rCSI in Malawi during 2010-2011 and 2013. This may reflect the experience of food insecurity specific to Malawi or it may reflect the underlying distribution of each food security measure (Maxwell et al., 2014). Extending our model to other locations and during periods of more severe food insecurity present exciting possibilities to better understand which outcome measure might be best suited to prediction. Our model improves on the status quo in food security early warning. Building a better, transparent, data-driven early warning system that is replicable and intuitive can save lives and resources, and encourage policymakers to pay more attention to early warnings of hunger and famine.

References

- Barrett, C. B., 2010. Measuring food insecurity. *Science* 327(5967): 825-828 (2010).
- Basu, K., 1996. Relief programs: When it may be better to give food rather than cash. *World Dev.* 24(1): 91-96.
- Bazzi, S., R. A. Blair, C. Blattman, O. Dube, M. Gudgeon, R. Peck, 2018. The Promise and Pitfalls of Conflict Prediction: Evidence from Columbia and Indonesia. January. *Working Paper*.
- Blumenstock, J., G. Cadamuro, R. On, 2015. Predicting poverty and wealth from mobile phone metadata. *Science*. 350. 1073-1076.
- Boyle, S. A., C. M. Kennedy, J. Torres, K. Colman, P. E. Perez-Estigarribia, N. U. de la Sancha, 2014. High-resolution satellite imagery is an important yet underutilized resource in conservation biology. *PLOS One*. 9(1).
- Brown, M., K. Grace, G. Shively, K. Johnson, M. Carroll, 2014. Using satellite remote sensing and household survey data to assess human health and nutrition response to environmental change. *Popul Environ* 36: 48-72.
- De Waal, A., 2018. *Mass Starvation: The history and future of famine*. Wiley Press: Cambridge: UK.
- The Economist, 2017. The Economist explains: How is famine declared? The UN relies on technocratic calculations but politics still play a part. *The Economist*. (April 4).
- Famine Early Warning System Network, 2018. Data from “Famine Early Warning System Network Data Center.” <http://fews.net/data> Accessed January 3, 2018.
- Food and Agricultural Organization, 2017. *The State of Food Security and Nutrition in the World: Building resilience for peace and food security*. Accessed January 2, 2018. <http://www.fao.org/3/a-I7695e.pdf>

- Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, L., A. Hoell, J. Michaelson, 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* 2, 150066. doi:10.1038/sdata.2015.66 2015
- GADM, 2018. Data from “GADM Maps and Data.” Version 3.2. <https://gadm.org/data.html> Accessed January 3, 2018.
- Gelli, A., N.L. Aberman, A. Margolies, M. Santacroce, B. Baulch, E. Chirwa, 2017. Lean-season food transfers affect children’s diets and household food insecurity: evidence from a quasi-experiment in Malawi. *J Nutr* 147(5): 869-878.
- Guan, K., B. M. Sultan, M. Biasutti, D. B. Lobell, 2015. What aspects of future rainfall changes matter for crop yields in West Africa. *Geophys Res Lett* 42(19).
- Headey, D., C.B. Barrett. 2015. Opinion: Measuring development resilience in the world’s poorest countries. *Proceedings of the National Academy of Sciences*. 112(37): 11423-11425.
- Hidrobo, M., J. Hoddinott, N. Kumar, E. Olivier, 2018. Social protection, food security and asset formation. *World Dev* 101:88-103.
- Hyman, G., C. Larrea, A. Farrow, 2005. Methods, results, and policy implications of poverty and food security mapping assessments. *Food Policy* (30): 453-460.
- IPC, 2012. *Technical Manual Version 2.0: Evidence and Standards for Better Food Security Decisions* (Food and Agricultural Organization: Rome).
- Jean, N., M. Burke, M. Xie, W. M. Davis, D. B. Lobell, S. Ermon, 2016. Combining satellite imagery and machine learning to predict poverty. *Science* (19 Aug): 790-794.
- Johnson, K., A. Jacob, M. Brown, 2013. Forest cover associated with improved child health and nutrition: Evidence from the Malawi Demographic and Health Survey and satellite data. *Global Health Science and Practice*. 1(2): 237-248.

- Jones, A., A. Shrinivas, R. Bezner-Kerr, 2014. Farm production diversity is associated with greater household dietary diversity in Malawi: findings from nationally representative data. *Food Policy* 46:1-12.
- Knippenberg, E., N. Jensen, M. Constanas, 2019. “Quantifying household resilience with high frequency data: Temporal dynamics and methodological options.” *World Development*. 121:1-15.
- Lentz, E. C., C. B. Barrett, 2013. The economics and nutritional impacts of food assistance policies and programs. *Food Policy* 42:151-163.
- Lobell, D., M. B. Burke, C. Tebaldi, M. Mastrandrea, W. P. Falcon, R. L. Naylor, 2008. Prioritizing Climate Change Adaptation Needs for Food Security in 2030. *Science*. 319(5863): 607-610.
- Mallory, M., K. Baylis, 2016. The food corporation of India and the public distribution system: Impacts on market integration in wheat, rice, and pearl millet. *Agribusiness* 30(2): 225-250.
- Maxwell, D., B. Vaitla, J. Coates, 2014. How do indicators of household food insecurity measure up? An empirical comparison from Ethiopia. *Food Policy* 47:107-116.
- Nikulkov, A., C. B. Barrett, A. G. Mude, L. M. Wein, 2016. Assessing the impact of U. S. food assistance delivery policies on child mortality in northern Kenya. *PLoS ONE* 11(12) (December): e0168432. doi:10.1371/journal.pone.0168432
- Niles, M., M. Brown, 2017. A multi-country assessment of factors related to smallholder food security in varying rainfall conditions. *Scientific Reports*. 7(16277): 1-11.
- ODAV (Vulnerability Analysis and Mapping Branch) 2008. World Food Programme (WFP) Food Consumption Analysis. WFP: Rome.
- Sachs, J., R. Remans, S. Smukler, S. J. Andelman, K. G. Cassman, D. Castle, R. DeFries, G. Denning, J. Fanzo, L. E. Jackson, R. Leemans, J. Lehmann, J. C. Milder, S. Naeem, G. Nziguheba, C. A. Palm, P. L. Pingali, J. P. Reganold, D. D. Richter, S. J. Scherr, J. Sircely, C.

- Sullivan, T. P. Tomich, P. A. Sanchez, 2010. Monitoring the world's agriculture. *Nature* 466: 558–560.
- Sen, A., 1982. *Poverty and Famines: An Essay on Entitlement and Deprivation*. Oxford University Press.
- Shively, G., 2017. Infrastructure mitigates the sensitivity of child growth to local agriculture and rainfall in Nepal and Uganda. *Proc Natl Acad Sci USA* 114(5): 903-908 (2017).
- Swindale, A., P. Bilinsky, 2006. Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide: Version 2.
https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf
Accessed January 3, 2018.
- Vaitla, B. J. Coates, L. Glaeser, C. Hillbruner, P. Biswal, D. Maxwell, 2017. The measurement of household food security: Correlation and latent variable analysis of alternative indicators in a large multi-country dataset. *Food Policy* 68: 193-205.
- Vagen, T. G., L.A. Winowiecki, J.E. Tondoh, L.T. Desta, T. Gumbricht, 2016. Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma* 263: 216-225.
- Webb, P. J. Coates, E. A. Frongillo, B. L. Rogers, A. Swindale, P. Bilinsky, 2006. Measuring household food insecurity: why it's so important and yet so difficult to do. *J Nutr* 136: 1404S–1408S.
- World Bank, 2018. “Living Standards Measurement Surveys: Malawi 2010-11” and “Living Standards Measurement Surveys: Malawi 2010-13.”
<http://microdata.worldbank.org/index.php/catalog/lsms/about> Accessed January 3, 2018.

Data and materials availability:

LSMS data: <http://microdata.worldbank.org/index.php/catalog/lsms>.

Weather data: CHIRPS <http://chg.geog.ucsb.edu/data/chirps/>;

Elevation: NASA SRTM 90m <ftp://xftp.jrc.it/pub/srtmV4/arc.asci/> ;

Soil: FAO Harmonized World Soil Database <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/> ;

Agricultural land cover: GlobCover v 2.3 http://due.esrin.esa.int/page_globcover.php.

Global Information and Early Warning System (GIEWS): Food Price Monitoring Tool
<http://www.fao.org/giews/food-prices/tool/public/#/dataset/domestic>

Additional data related to this paper (i.e., IPC and price data) may be requested from the authors.

FIGURES AND TABLES

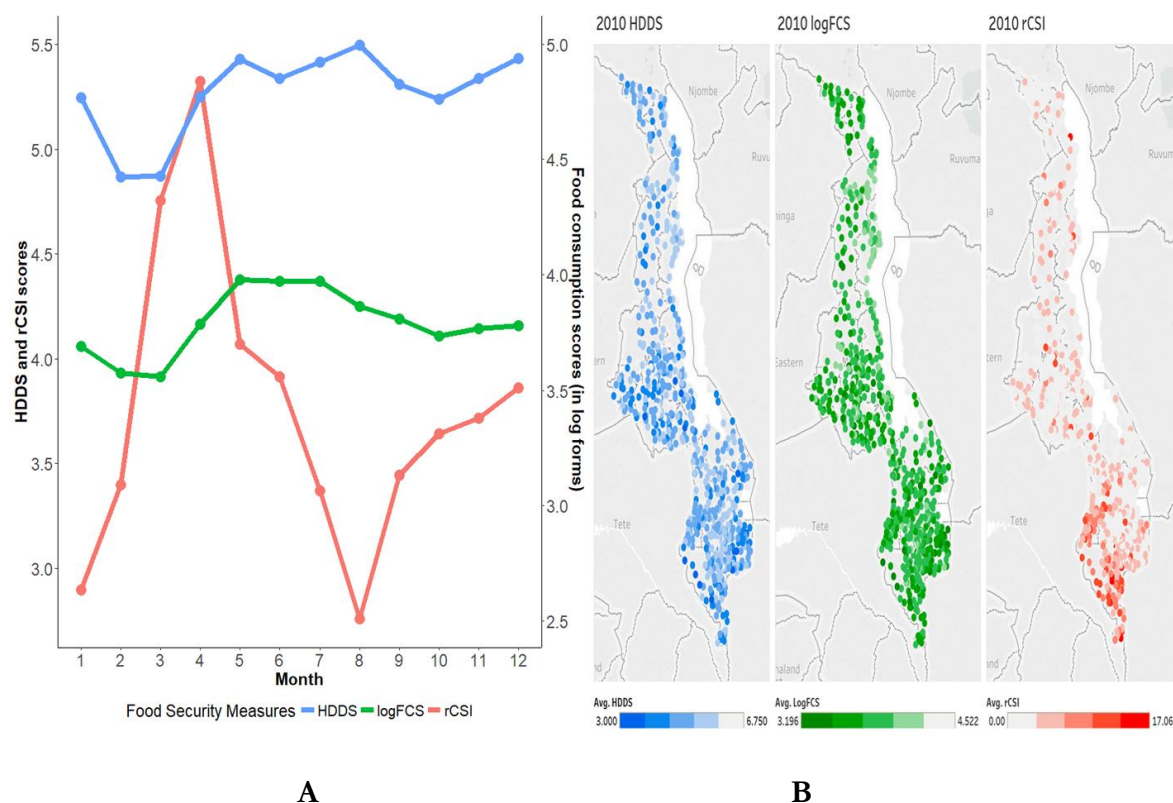
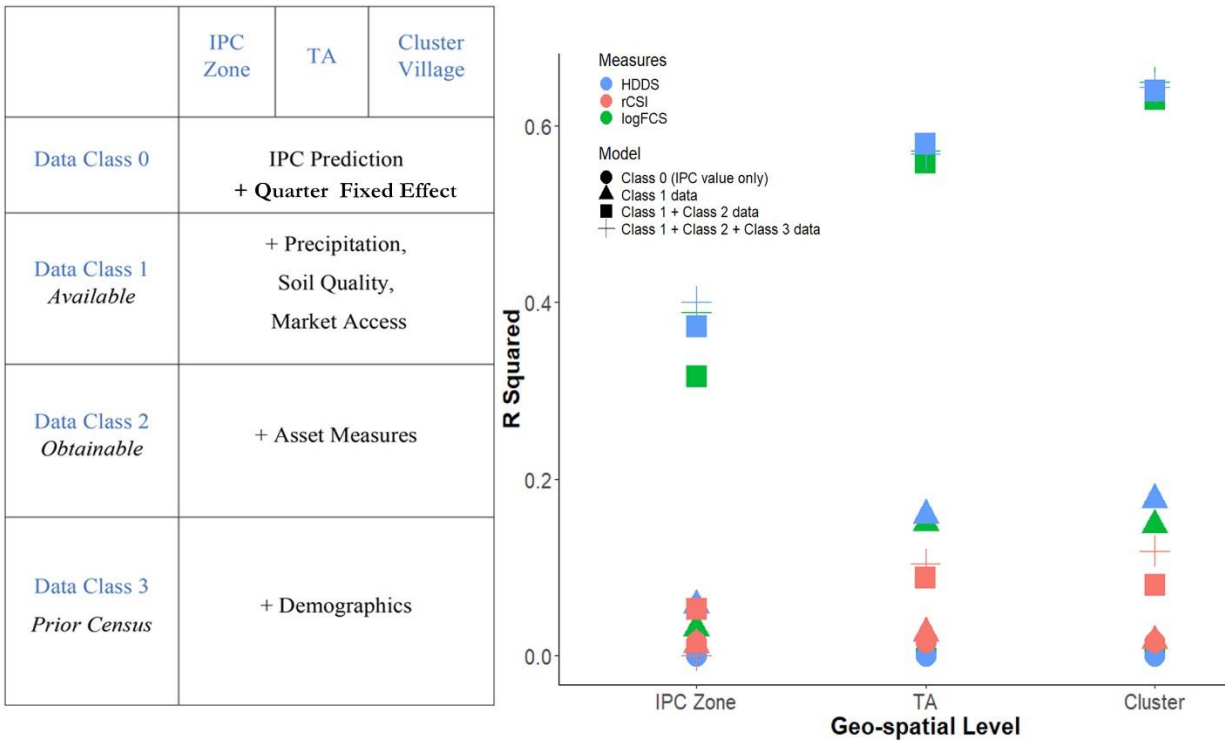


Fig. 1. Food security measures vary across space and time. (A) Average food security measures by month for 2010 vary by season. January = 1. Read rCSI (reduced coping strategies index) and HDDS (Household Dietary Diversity Score) against the left axis and FCS (Food Consumption Score) against the right axis. A high rCSI indicates the use of more coping strategies, thus low food security, whereas a high HDDS or FCS indicates high food security. **(B)** Darker shades of cluster-level average food security measures in Malawi 2010-11 LSMS data indicate lower regional food security. Specifically, for the household dietary diversity score (HDDS, left panel) and log food consumption score (logFCS, central panel), darker shades indicate lower dietary quality. For the reduced coping strategies index (rCSI, right panel), a darker shade indicates that on average households in the cluster are employing more coping strategies.



A **B**

Fig. 2. The share of variation in out-of-sample cluster-level food security predicted by our models improves with greater spatial granularity and richer data. (A) We predict food security outcomes using three levels of spatial granularity and 4 classes of models. Class 0 data include the IPC early warning value only. Class 1 data contains: past IPC values, precipitation, market prices, market access measures, and soil quality. Class 2 data contains: share of households owing cellular phone and share of dwellings with metal versus thatch roof. Class 3 data contains: household demographics and assets. **(B)** The explanatory power of the models, measured as R-squared, increases with the Classes of data and the spatial granularity of data used. Our best model explains 64% of the variation in cluster-averages of household dietary diversity (HDDS) and 65% of logged Food Consumption Scores (logFCS). However, at the most spatially disaggregated level, the cluster level, additional household and demographic variables add little additional information.

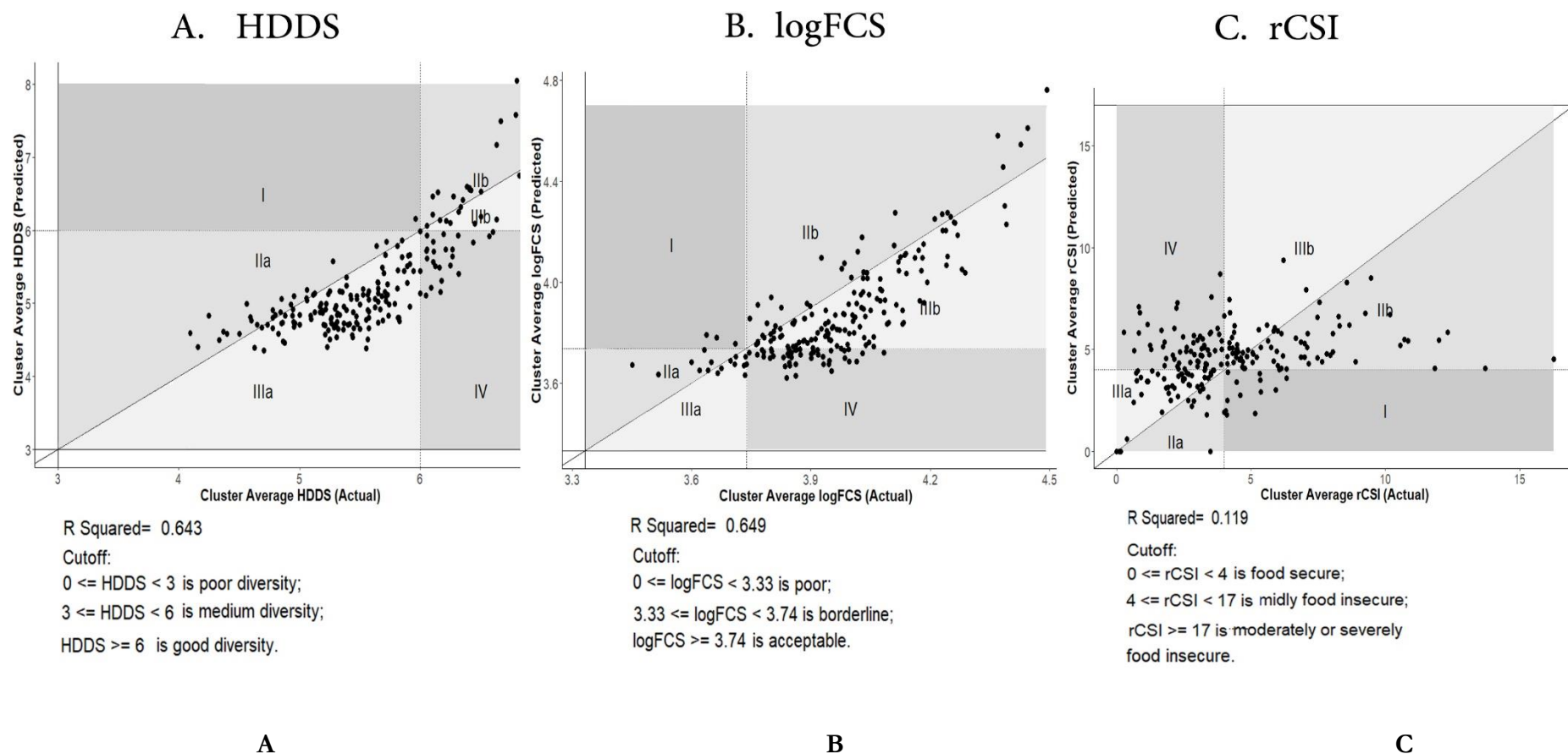


Fig. 3. Actual versus predicted out-of-sample cluster food security estimates, with errors of exclusion and inclusion, where errors of exclusion may be of more concern to relief agencies. Across all three panels, Area I shows areas in which the model under predicts food insecurity status (errors of exclusion) and Area IV shows areas in which the model over predicts food insecurity (errors of inclusion). Areas IIa,

I Ib, IIIa and IIIb are more nuanced. Areas IIa and IIIa show: clusters where the model correctly predicted that households were food secure but under or over predicted the continuous measure, respectively. Area I Ib and area IIIb show: clusters where categorical status was correctly predicted to be food insecure but continuous measures were under and over predicted, respectively. **(A)** Our model correctly classifies 88% of the household dietary diversity scores. **(B)** Our model correctly predicts 76% of the log food consumption score categories. **(C)** Our model correctly predicts 65% of reduced coping strategies index categories.

Table 1. Summary statistics of 2010 and 2013 Malawi food security measures and predictors for both household and cluster levels. Ranges of possible values are in parentheses in column 1. Higher rCSI values indicates lower food security; lower HDDS and FCS values indicate lower food security. The 2010-11 LSMS sample is representative at a national level for each month and, for the crop year, for each district. The 2013 food security measures are used for the out-of-sample predictions. We present the cluster average of the household-level variables, which includes food security measures and asset measures.

	Year 2010 (12270 households)					Year 2013 (3999 households)				
Variable	Mean	Median	Std. Dev.	Min	Max	Mean	Median	Std. Dev.	Min	Max
Household logFCS (0-4.72)	3.82	3.83	0.37	1.1	4.72	3.91	3.91	0.36	2.2	4.72
Household rCSI (0-42)	3.68	0	6.51	0	42	4.23	0	6.95	0	42
Household HDDS (0-12)	5.18	5	1.27	1	7	5.56	6	1.15	1	7
	Year 2010 (768 clusters)					Year 2013 (204 clusters)				
Cluster mean logFCS	3.87	3.86	0.21	3.23	4.53	3.96	3.95	0.18	3.45	4.49
Cluster mean rCSI	3.68	2.88	3.05	0.00	17.25	4.26	3.70	2.66	0.00	16.28
Cluster mean HDDS	5.18	5.19	0.70	3.00	6.75	5.55	5.55	0.57	4.10	6.86
Total rainfall (meters)	1.00	0.99	0.18	0.57	1.58	0.93	0.90	0.17	0.59	1.58
First day of rain	45.02	41.00	10.25	3.00	72.00	53.36	51.00	13.26	30.00	72.00

Max days without rain	23.76	21.00	8.14	10.00	52.00	24.30	24.00	6.75	12.00	43.00
Rainfall in flood prone regions (meters)	2.04	0.00	9.66	0.00	75.97	1.11	0.00	7.34	0.00	69.15
Number of cellphones owned	0.60	0.44	0.60	0.00	4.13	0.94	0.70	0.71	0.05	4.40
Maize price (log form)	3.34	3.41	0.39	2.40	5.19	4.46	4.50	0.26	2.77	5.05
Market thinness	0.46	0.49	0.29	0.00	1.00	0.27	0.25	0.16	0.00	1.00
Percent of non-natural roof	0.36	0.49	0.25	0.00	1.00	0.46	0.40	0.27	0.00	1.00
Household size	4.56	4.56	0.68	2.31	7.06	5.00	4.89	0.72	3.58	7.53
Household age	42.15	42.06	4.65	30.81	56.38	42.56	42.37	4.06	33.63	56.19
Household gender (1 for male, 2 for female)	1.24	1.25	0.12	1.00	1.69	1.23	1.22	0.10	1.00	1.50
Asset Index	0.00	-0.02	0.35	-0.83	1.05	-0.02	-0.30	0.54	-0.30	2.94
Distance to road (km)	8.37	4.36	10.15	0.07	56.19	7.68	4.18	8.40	0.06	44.68
Distance to Admarc market (km)	8.06	6.56	5.77	0.38	37.32	7.81	6.24	5.04	1.20	32.89
Percentage of agricultural land	0.36	0.49	0.25	0.00	1.00	0.34	0.41	0.22	0.00	0.75
Nutrition retention constrained	0.28	0.00	0.44	0.00	1.00	0.26	0.05	0.38	0.00	1.00
Elevation (km)	0.87	0.90	0.35	0.04	1.73	0.94	1.02	0.29	0.12	1.55
IPC Value (1 month lag)	1.17	1.00	0.40	1.00	3.00	1.08	1.00	0.22	1.00	3.00

Table 2: Regression results for each food security measure using 2010 LSMS data for Malawi

confirms that food security measures are associated with common drivers. The results are estimated at the cluster-level and include predictors from all Class 1 + Class 2 + Class 3. Standard errors are presented in parentheses and asterisks indicate level of statistical significance of coefficients where three asterisks indicate 1%; two indicate 5% and one indicates 10%.

	logFCS	HDDS	rCSI
IPC Value (Previous month)	-0.023 (0.015)	-0.081* (0.047)	2.223*** (0.274)
Total rainfall during rainy season	0.104*** (0.032)	0.101 (0.103)	-1.302** (0.598)
First day of rain	-0.001 (0.001)	-0.002 (0.002)	0.017* (0.010)
Maximum days without rain	-0.0005 (0.001)	-0.005** (0.002)	-0.047*** (0.013)
Maize price	-0.001** (0.0004)	-0.002* (0.001)	0.025*** (0.007)
Maize market thinness	-0.010 (0.017)	-0.085 (0.055)	-0.272 (0.317)
Percent of agricultural land	0.011 (0.020)	0.034 (0.066)	-0.117 (0.382)
Elevation	-0.060*** (0.018)	-0.048 (0.060)	-1.216*** (0.346)
Soil nutrient retention constrained	-0.014 (0.012)	-0.079** (0.040)	0.509** (0.229)
Distance to nearest road	-0.002*** (0.001)	-0.010*** (0.002)	-0.018* (0.010)
Distance to nearest Admarc market	0.003*** (0.001)	0.005* (0.003)	-0.007 (0.017)
Roof type	0.111*** (0.029)	0.572*** (0.093)	0.158 (0.537)
Number of cellphones owned	0.143*** (0.014)	0.332*** (0.046)	-1.084*** (0.266)
Household size	-0.015* (0.008)	-0.041 (0.025)	0.631*** (0.146)
Household age	-0.003** (0.001)	-0.015*** (0.004)	-0.016 (0.021)
Household gender	-0.040 (0.043)	0.064 (0.138)	0.445 (0.799)
Asset index	0.123*** (0.017)	0.462*** (0.055)	-0.704** (0.317)

Quarter 1	-0.059*** (0.015)	-0.193*** (0.048)	-0.169 (0.276)
Quarter 2	-0.012 (0.016)	-0.048 (0.053)	-0.166 (0.304)
Quarter 3	0.007 (0.013)	-0.057 (0.043)	-1.121*** (0.248)
Observations	760	760	760
R ²	0.605	0.623	0.337

Table 3: The comparison of model with different specifications. The results are out-of-sample predictions in 2013. Models are estimated at the cluster-level, using only the IPC value and include predictors from all Class 1 + Class 2 + Class 3, respectively. The percentage of food insecure clusters that are predicted to be in the insecure category is also known as the true positive rate or “sensitivity” of prediction. The categorical accuracy is the percentage of clusters that are correctly predicted to be the same as the actual food security category.

	R squared			Sensitivity of Food Insecure Category			Categorical accuracy		
	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI
Main model	0.649	0.643	0.119	0.833	0.994	0.86	0.759	0.882	0.650
Main model without cluster price variables	0.638	0.661	0.119	0.222	0.968	0.258	0.916	0.882	0.591
Main Model with region fixed effect and without cluster price variables	0.557	0.554	0.073	0.278	0.955	0.290	0.916	0.862	0.616
Main Model with GIEWS price and without cluster price variables	0.634	0.661	0.036	0.722	0.994	0.978	0.793	0.892	0.468
LASSO	0.640	0.611	0.061	0.833	1	1	0.665	0.857	0.463

Table 4: In-sample IPC value regression results indicate that the IPC is significantly associated with food insecurity but the explanatory power (R-squared) is quite low. Standard errors are presented in parentheses and asterisks indicate level of statistical significance of coefficients where three asterisks indicate 1%; two indicate 5% and one indicates 10%.

	(1)	(2)	(3)
	IPC Value	IPC Value	IPC Value
rCSI	0.055***		

	(0.004)		
logFCS		-0.282*** (0.069)	
HDDS			-0.087*** (0.021)
Constant	0.963*** (0.021)	2.255*** (0.267)	1.616*** (0.109)
Observations	760	760	760
R ²	0.175	0.022	0.023

Table 5: The percentage of food insecure clusters correctly predicted to be food insecure. The results are out-of-sample predictions at the food insecure category in 2013. They are estimated at the cluster-level, first using only the IPC value and then including predictors from all data Classes (“Full Model”). The percentage of food insecure clusters that are predicted to be in the insecure category is also known as the true positive rate or “sensitivity” of prediction.

Model	HDDS	logFCS	rCSI
IPC only	1.000	0.000	0.097
Full model	0.994	0.833	0.860

SUPPLEMENTARY MATERIALS

The supplementary material provided includes details on: (S1) four challenges associated with food security measurement for early warning; (S2) protocols to facilitate the application of our approach to other contexts; (S3) additional information on food security measures and cutoffs, and on the construction of Class 1 – 3 variables; and (S4) additional results. The additional results include: density plots of actual values versus out-of-sample predictions for 2013 for all spatial levels; results using month fixed effects; coefficient estimates from the tail-only subsample for 2010 at the cluster-level including all variables; and accuracy of categorical predictions for all data classes and spatial levels using both the full sample and the food insecure subsample.

S1. Food security measurement challenges

Four challenges have impeded the development of effective predictions for food security early warning systems. Below, we describe the limitations and how our model aims to address each one.

First, availability estimates, such as crop models, miss crucial determinants of food security (Sen, 1982). Increased availability of high-frequency, high-resolution precipitation data, ground cover data, and soil moisture data has launched a series of efforts to better predict crop production for food security in low income countries (Lobell et al., 2008). These efforts include crop-monitoring systems, such as Geoglam, intended to alert policymakers and others when a crop failure may be imminent. A limitation of these models is that they exclusively model crop production – leaving out available information on market prices or local average demographic characteristics that affect where and how production shortfalls will translate into household outcomes (Niles et al., 2017; Shively, 2017; Hidrobo et al., 2018). Other existing

availability measures, such as food balance sheets or the availability data from Global Information and Early Warning System, are often too spatially and temporally aggregated to be of use for high frequency, sub-national forecasting. We incorporate high frequency market data and prices – measures of access and stability that are relevant to food security.

A second limitation of existing approaches to measure food security is that individual and household-level measures of food security are generally from one-off, cross-sectional household surveys that are quickly out of date and reveal little about the dynamics, stability and future of food security (Nikulkov et al., 2016). Seasonal hunger, for example, is more common than suggested by food security measures derived from annual surveys. Figure 1 in the main text shows the sensitivity of three different food security measures to seasonal variation in Malawi in 2010, discussed further below. Further, even when available, household questions about food security are generally retrospective, revealing household circumstances weeks or months prior to the assessment, but do not capture prospective need (Lentz and Barrett, 2013). For these reasons, household surveys offer limited insights into future food insecurity, highlighting the need for predictive approaches.

Third, our model fills an important geographical gap: most food security measures are captured at individual and household level, or at the national level - leaving a critical gap at the meso-level, such as in villages, sub-districts or districts, where crises happen and interventions generally take place. These sub-national levels are particularly relevant to identifying food and nutrition insecurity given that food insecurity can be a highly localized phenomenon, with some zones within a country experiencing acute – and regular – insecurity while others do not (Sen, 1982; Lentz and Barrett 2013; Boyle et al., 2014; Brown et al., 2014). Further, the limited attention to the meso-level hinders sub-national targeting of responses.

Finally, our research takes a data-driven approach to the status quo approach for assessing sub-national food security, the IPC. The IPC is used for decision-making around food security and famine declarations, by local governments, USAID, European Commission, and UN agencies including the WFP, UNICEF, and FAO. IPC-trained teams answer a series of questions and sift through available evidence including food prices, anthropometric measures, and mortality rates, to assess food insecurity within and across countries (IPC, 2012). An IPC analysis requires a working consensus of analysts from key stakeholders, such as government ministry members, NGO organization members, early warning systems analysts, and others; an IPC compatible analysis does not require consensus from key stakeholders and may be completed by a single organization; other requirements (e.g., consistent use of evidence) remain the same across the two analyses. Because an objective of the IPC is to forecast food insecurity levels to provide early warning to governments and humanitarian agencies, its classification process relies on the best data available at the time of the assessment; the availability and quality of data can vary across time and space. The IPC system requires some interpretation on the part of the assessing team, and the IPC has faced the criticism that it is too complex, not easily replicable, not consistent across countries, and too reliant on the availability of detailed information. Taken together, concerns have been raised that its methods make it vulnerable to political influence (The Economist, 2017; de Waal, 2018).

2. Methods: Protocol for matching data

One challenge associated with using both availability and access measures is how to combine information from various sources collected at a range of geographic scales. In our model, we wish to link price data from a geocoded market location with gridded weather data and other cluster-geocoded household information. Fig. S1 shows the location of the IPC zones and 72 markets in Malawi.

[Figure S1 about here]

We assign market prices to each IPC zone or TA using the following approach. First, we assume that people visit the market closest to them by the straightest path or Euclidean distance. Thiessen polygon boundaries are drawn based on the midpoint between market locations. All households within a Thiessen polygon boundary are closest to the market within that polygon boundary. This approach creates a marketshed for households, based on market proximity. We then overlay the IPC zones onto the Thiessen polygons and clip each Thiessen polygon by the IPC zone that they fall into. See Fig. S2. From a spatial perspective, each IPC zone may include some, one, or none of our 72 sampled maize markets. We assume that people within an IPC zone experience the market prices closest to them, regardless of whether that market is located within the IPC zone boundaries. An alternative approach would be to assign people to markets based on road networks (e.g., it may be faster to go to a market that is 5 km away on a paved road than a market that is 3km away on a dirt path). This would be a valuable approach, but Malawi's road network information does not allow us to make this refinement. The final market price for each IPC zone is the average of markets prices weighted by population from the 2011 Landscan files within the intersection of the IPC zone and marketshed. This weighting ensures that the resulting IPC-level price represents the price accessible to the average household in the zone. We repeat this process for the TA polygons, whereas for the clusters, we simply use the price of the marketshed into which each cluster falls.

[Figure S2 about here]

For precipitation data, we extract values from raster data on precipitation and take an average of the pixel values for each IPC zone and TA polygon. At the cluster level, we take the average of the pixel values within a 2 km buffer around the coordinates of each cluster. We use a 2 km buffer because LSMS cluster geocodes are deviated by up to 2 km to preserve anonymity of the household. The precipitation raster dataset comes from Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) (Funk et al., 2015). The 2012 Malawi livelihood zone shapefile is from FEWS Net (FEWS Net, 2018) and the 2015 Malawi administrative boundaries are from the GADM database (GADM, 2018).

Techniques for joining data at different spatial levels are new (Boyle et al., 2014; Brown et al., 2014; Johnson et al., 2013), and a benefit of this protocol is that it can be readily applied to other contexts. Livelihood zone maps and administrative maps exist for almost all the other countries and the CHIRPS rainfall is a quasi-global dataset. Price data, wherever they are collected, could be transformed to prices at the IPC, TA or cluster level with this protocol. Google Maps can be one source for market geo-coordinates.

3. Data

Measures of Food Security

Many different measures of food security exist (Barrett, 2010). Food security monitoring or analysis focusing on access and utilization often rely on household or individual level measures of food security, such as the reduced coping strategies index (rCSI) (a measure of five household coping strategies associated with food intake that are weighted by frequency and severity) and the household dietary diversity score (HDDS) and food consumption score (FCS) (measures of food groups consumed). Major donors, such as USAID, and United

Nations organizations, such as the WFP, use subsets of these measures for program monitoring. Each measure captures a slightly different aspect of food insecurity; rCSI is thought to better capture quantity of food consumed, while FCS and HDDS better capture the nutritional value of food consumed (Maxwell et al., 2014; Vaitla et al., 2017).

The dependent variables we use in our analysis (FCS, HDDS, and rCSI) are measured at the household level and come from the 2010-11 and 2013 Malawi Living Standards Measurement Survey (LSMS) (in Malawi also called the Integrated Household Survey). Households are surveyed over 12 months, and the LSMS sample is designed to be representative for the nation each month, and each district over the year (World Bank, 2018). One challenge is different clusters are sampled each month, implying that cluster fixed effects will be perfectly correlated with months during the year, and month fixed effects will be correlated with clusters; thus, standard granular fixed effects approaches from one year of data may overfit.

Higher values of HDDS and FCS indicate greater food security. The HDDS is a count of the number of food categories that each household consumes over a week in our sample with possible values ranging from zero to twelve; sample responses ranged from three to eight. Thus, the true HDDS distribution is clustered at integers. The FCS is similar to the HDDS but weights different food groups according to their nutrient density. Higher rCSI scores, conversely, indicate greater food insecurity. The rCSI sums the number of coping strategies a household has taken to mitigate food insecurity, from a possible five strategies, weighted by frequency of use in the past week and severity of the coping strategy; food secure households not engaging in any food-related coping strategies have scores of 0. The theoretical maximum score is 56.

There are no universal cut-offs for HDDS or rCSI (Maxwell et al., 2014); we apply cutoffs for

HDDS from neighboring Mozambique to Malawi (Vaitla et al., 2017; Swindale and Bilinsky, 2006) and cutoffs for the rCSI based on provisional findings (Maxwell et al., 2014). For rCSI, a score of 0-4 indicates food secure, 5-10 indicates moderate food insecurity; and 11 and above, high food insecurity (Vaitla et al., 2017). For FCS, cut-offs of 28 and 42 are more appropriate for populations that frequently consume sugar and oil (ODAV, 2008). We log FCS to place more weight on the lower end of the distribution. Households are considered to have poor food consumption status with logFCS below 1.447; borderline, between 1.447 and 1.623; and acceptable, with logFCS above 1.623.

We explore sources of variation in our food security measures. When we decompose the variation of the three food security measures from 2010 using simple fixed-effects only models, we find that month of the data collection explains 1.3% of the variation for rCSI, 1.9% of the variation for HDDS and 2.1% of the variation for the FCS. For our year of data, the spatial level explains more of the variation for all measures than does month of collection. For the rCSI, the IPC zone, TA, and cluster explain 8.3%, 13.8%, and 21.7% of the variation in household food security scores, respectively. The degree of variation in the measures explained by time and geography are presented in Fig. S2. For HDDS, IPC zones explain 10.3% of variation, TAs 20.1% of variation and clusters 30% of the variation. For logFCS, IPC zones explain 12.8% of the variation, TAs 23.9% and clusters 32.8%. Thus, within a given year, spatial variation explains a substantial fraction of the variation in household-level food insecurity. Note that because of the sampling that all households in a cluster are sampled within the same month in a given survey wave, the cluster fixed effects also pick up the effects of the sampling month. Nonetheless, we also find substantial variation among households in the same month, living in the same village.

Class 1 data

Knowing that food security varies over space and time does not explain what specific characteristics underlie this variation. To explain the variation in food insecurity, we draw on readily and widely available Class 1 data on rainfall, prices, and IPC food security assessments. Below, we describe data sources and processing. A summary table, S1, describes the various sources of data, and how they were matched.

Rainfall data is from the CHIRPS data set, which collects daily data on rainfall at a high spatial resolution (Funk et al., 2015). We use the CHIRPS data to generate several agronomically-relevant measures of precipitation. First, we measure the total rainfall during the unimodal rainy season (from October to April) for the prior year, on the assumption that current food security is largely affected by last year's harvest. Second, we develop a measure of the beginning of the rains for the prior agricultural season, to capture late-onset rain, which affect agricultural yield (Guan et al., 2015). Third, we develop a measure for length of dry spells during the past rainy season. As discussed, we integrated the CHIRPS data with the IPC data using protocols on population weights and geographic matching.

Price data for maize for 72 markets distributed across Malawi between 2000 and 2017 were collected on a weekly basis by the Malawi Statistical Division of the Ministry of Agriculture. The price data include weekly maize prices from a variety of markets, including more remote, rural markets and urban markets. The Ministry of Agriculture purposively sampled these markets to be representative of markets across Malawi (Ministry of Agriculture personal communication). See Fig. S1 for locations (some monitored urban markets are in close proximity and appear on the map to share a location). We use the price data to calculate monthly price averages by market and to assess the monthly thickness of the market (measured as the share of missing prices per month).

Several weekly price observations are missing, which itself may contain information about maize availability. We use the following procedure to interpolate price data. For the missing weekly prices in a given month, we take an average of any available weekly prices in that month. If all weekly prices are missing in that month for this particular market, we use a spatial interpolation technique under the assumption that there is spatial price integration between the markets. Specifically, we replace the missing price at a market with the price in its 'k'th nearest neighbor market. In other words, $k = 1$ if the price in its nearest market is available. If not, we replace the missing price using the second nearest market. We iterate this process until data for each month is interpolated. For the few months in our data where none the markets have price data, we use temporal interpolation to replace the missing price with interpolated values from other time periods in the same market.

The market thinness measure for each market is the proportion of weeks in a given month that price data for maize are missing for a particular market. The measure is 0 if we have price for all weeks in that market and it is 1 if the price information is missing for all weeks in that month. We employed the same approach as our price data to compute the average market thinness measure at the IPC, TA and cluster level.

IPC food security assessment data FEWS Net provided quarterly IPC assessment levels for Malawi from July of 2009 to April 2016. The roughly quarterly IPC assessments in Malawi focus on identifying current acute food insecurity, which is often “on top of” chronic food insecurity. Nearly 78% of Malawi’s IPC assessments between 2009 and 2016 were classified as phase 1, or “minimal.” About one-fifth (19%) were assessed at phase 2, “stressed,” and only 2.5% were phase 3, “crisis.” The periods of greatest phase 3 food insecurity occurred in 2015

(7% of assessments for that year), and 13 IPC zones were evaluated as in “crisis” at the end of 2015.

The IPC has never been validated against household-level food security data. An initial undertaking was to test the predictive power of the IPC on household level reported food security access. We find that the IPC is a strong predictor of household food security status; moreover, the IPC is well predicted by our simple weather and price data. These results are presented in the main text of the document.

Geographic data We include easily measured geographic variables such as elevation and distance to road and market. We also include a variable that captures the percent of agricultural land in the area, which should be readily available from satellite imagery or pre-processed data. Last, we include a measure of soil quality, controlling for nutrient retention ability of the soil. While not all countries have soil maps, work by the World Agroforestry Centre (ICRAF) to map soils throughout Africa will make these data more readily available (Vagen et al., 2016).

Class 2 data

We also include what we term **Class 2** data. **Class 2** data is a mix of time-varying and time-invariant variables and may be available from government or private sources but likely to require some additional work or expense to be accessed and processed. These are either (1) potentially available as remote sensing products such as characterizations of the roof material of homes within a given geographic region (a proxy for socioeconomic status), or (2) available from cell phone companies or newer national censuses or representative surveys such as cell phone ownership. We derive these variables from Malawi’s 2010-11 LSMS (World Bank,

2018) but in the absence of a relatively recent household survey, these variables could be accessed from other sources.

Class 3 data

Class 3 data are household-level and likely to be available from a country census, a LSMS or Demographic and Health Surveys. These data could also be collected through rapid techniques, such as cell-phone based surveys.

We use **Class 3** data on household-level demographics and assets from the LSMS. These data would add the most precision to targeting as they include richer household-level access measures. Demographic characteristics include (a) gender of head of household (b) age of household head (c) household size and (d) household dependency ratio. Household assets are compiled into an asset index computed using principal components analysis and including the following assets: refrigerator, television radio, bicycle, motorcycle and car.

4. Additional Results

Distribution of predicted values

The density plots for predicted and actual values are for our three measures of food security shown in Fig. S3. The predicted values shown in these plots are all from models using the full set of predictors (Class 1 + Class 2 + Class 3); data vary in terms of spatial granularity of the variables. Results for Class 1 and Class 1 + Class 2 alone models are available upon request. The cluster level averages of food security measures obtained from the household surveys are shown in red (labeled as “Actual”) as a reference point to our model predications. Predicted values using variables at the IPC zone, TA and cluster level are shown in blue, purple, yellow

respectively. As is shown in the figures, as the level of spatial granularity increases, the distribution of the predicated values moves towards the actual distribution both in the center and the spread of the distributions. At the higher aggregation level, the predicted values centers around the mean of the predication and that the mean can be relatively off from the actual mean. The prediction model can capture more heterogeneity among different clusters as the level of spatial granularity increases. All of these model predictions slightly underpredict the actual food security status across the three food security measures. Prediction values using only the IPC value are shown in teal, mostly clustered around the mean of the distribution. This prediction would leave out the clusters at the two extremes of the distribution, which may include those clusters of most concern from a humanitarian perspective.

[Figure S3 about here]

Values at the household level (as opposed to cluster averages) are shown in green, with a much larger spread and a longer tail. This difference suggests that even though our prediction model does a good job at fitting the values at the cluster level, it is hard to capture the household-level variation that lies within the clusters with the data that would be readily available via satellite imagery or past censuses.

Month fixed effects

In Table S2, we include out-of-sample predictions using month fixed effects. Compared to specifications in Table 3, which includes quarterly fixed effects, the models underperform. The LSMS sample, designed to be representative for the nation each month, and each district over the year, is not representative for each district each month (World Bank, 2018). The out-of-sample predictions with month fixed effects therefore overfit the model. Month fixed

effects could be useful for predictions based off of additional years of training data, or data representative at more granular spatial-temporal scales.

[Table S2 about here]

2010-11 tail regression results

We predict food security using only data from more food insecure households, first, to observe if the factors that affect varying degrees of food security differ over the distribution of food security and, second, to observe how placing more emphasis on food insecure households affects our predictions. The “tail” is defined by excluding the households that fall in the most food secure category of each food secure measure: rCSI greater than 4, logFCS smaller than 1.623 and HDDS smaller than 6. The cluster average tail food security measures are based on the cluster average of the remaining households at the tail (Table S4). The signs and significance of the coefficients on the variables included in the tails-only model closely match estimates from the full distribution (full results available upon request from the authors). Importantly, when estimating just the tails of the distribution, the length of dry spells largely negatively affects food security.

[Table S4 about here]

Accuracy of categorical predictions for full sample and subsample of food insecure households (“tail subsample”)

Many relief agencies use food security categories instead of the full distribution of the food security measure for targeting assistance. Thus, practitioners and policymakers may care more about accurately predicting the category of food security than the continuous measure.

Further, agencies might be more worried about predictions that minimize errors of omission:

mis-categorizing food insecure households in need of assistance as food secure. In Figure S4, we present the accuracy, defined as the percent of the observations correctly classified, the percent of Type I errors (errors of inclusion), and the percent of Type II errors (errors of exclusion), for all of our model predictions on the 2013 data. On the x-axis, from left to right, models grow more spatially granular, moving from IPZ zone to TA to cluster level. For each spatial level, we include accuracy, Type I and Type II results for Class 1 data, and then add in Class 2 and Class 3.

As can be seen in Figure S4, increasing spatial granularity generally improves accuracy for all measures for the full sample. It does not always reduce the percent of Type II errors, as can be observed by considering both logFCS and HDDS, which have very low Type II errors at the IPC zone level of analysis. Adding in Class 2 and Class 3 household asset information does not always improve accuracy, particularly at more coarse spatial scales of analysis. However, adding in asset data does, on average, reduce the percent of Type II errors (or errors of exclusion), which might be of particular importance to practitioners.

[Figure S4 about here]

When we use the model trained only on food insecure households (i.e. the ‘tail only’ models), we observe that the accuracy drops precipitously for most models. However, almost all errors are now Type I; Type II errors virtually disappear. While this approach of only using food insecure households to train the model is extreme, the results suggest that weighting households by their level of food insecurity might help reduce errors of exclusion. This approach could be most useful for emergency response agencies that are willing to trade some accuracy to identify more of the food insecure population for programming.

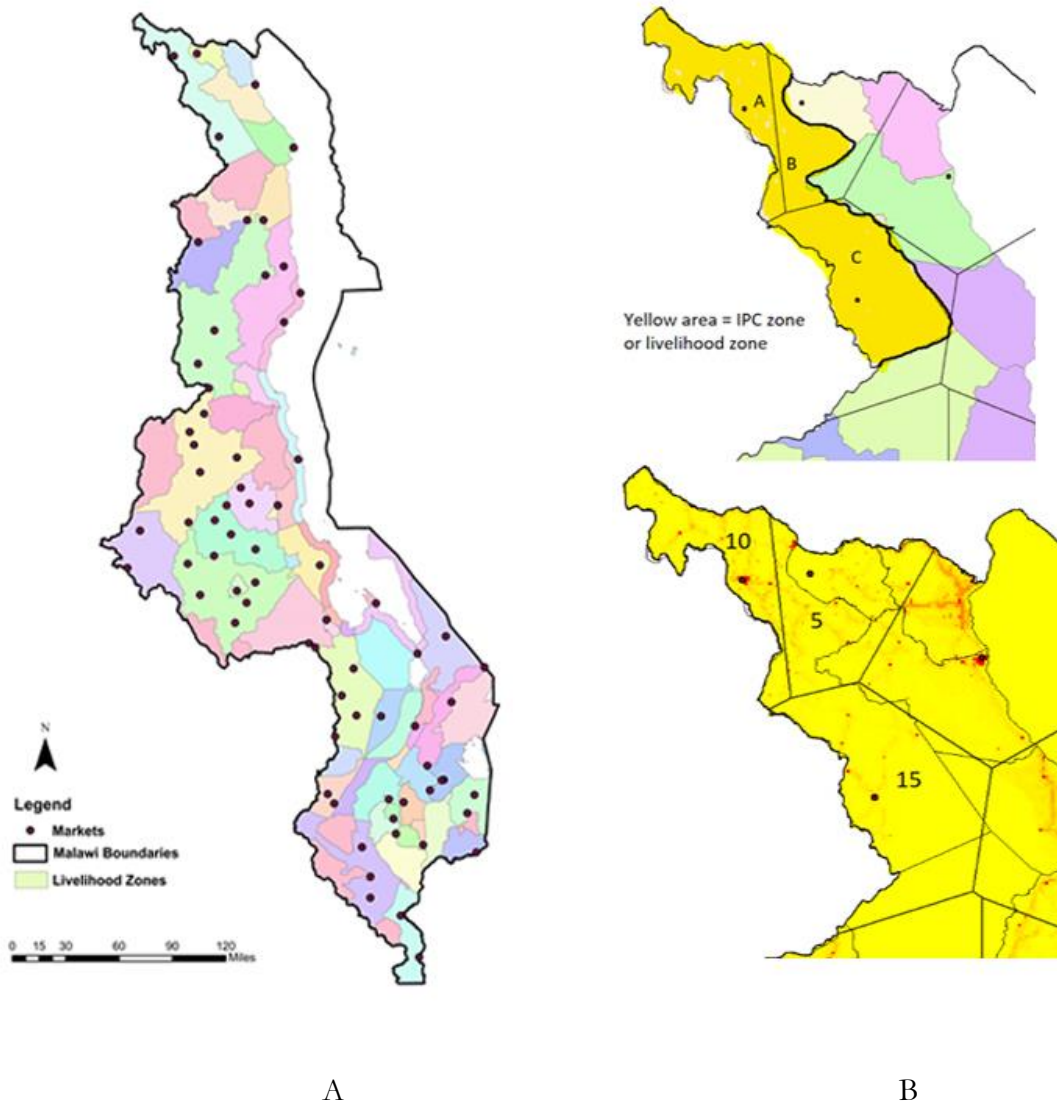


Fig. S1: Demonstration of application of our protocol for matching data at different spatial scales. (A) The map shows the delineation of IPC zones, with 72 markets indicated as dots. Some markets are close to one another (e.g., urban markets in Blantyre or Lilongwe) and therefore do not appear as distinct dots. (B) To match residents in an IPC zone to a market, Thiessen polygons are drawn around each market. We weight the Thiessen polygons that fall within a livelihood zone by population where the dot within each polygon is the closest market. For example, suppose Zone A included a population of 10 people; Zone B included a population of 5 people; and Zone C included a population of 15 people. We compute the total population for the livelihood zone as the following: Zone A (10) + Zone B (5) + Zone C (15) = 30 people.

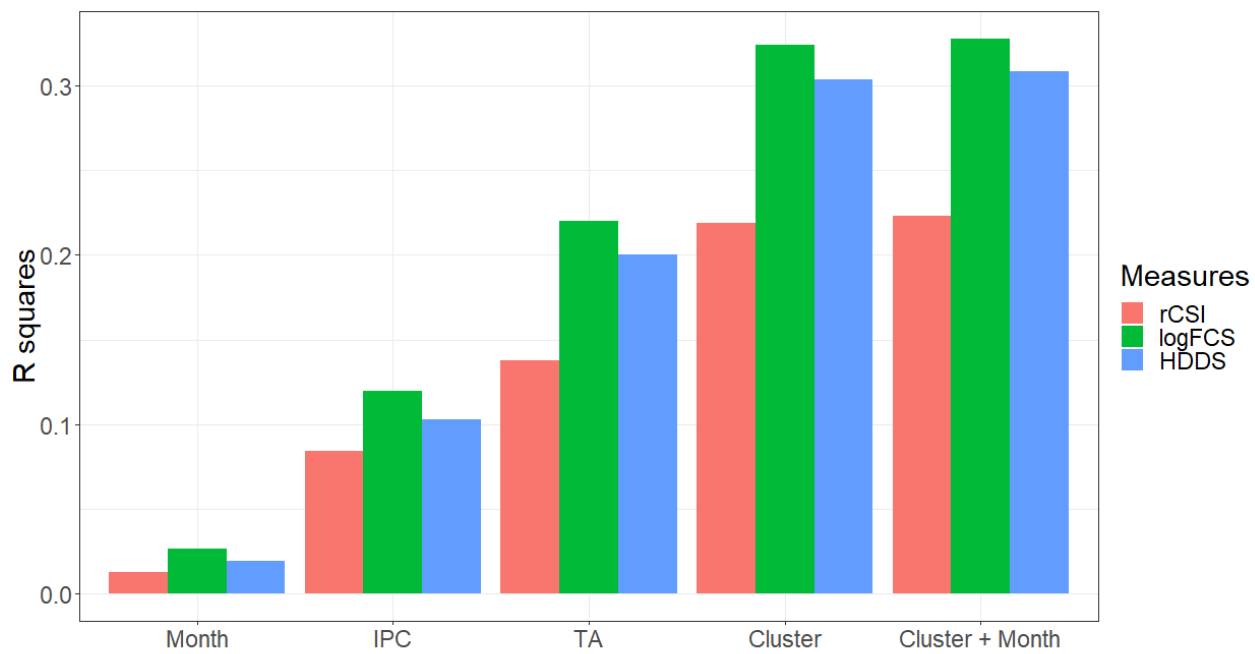


Fig. S2: A series of fixed effect models show that spatial level and temporal frequency influence food security status. Bars present the share of variation of each of the 2010 HH Food Security measures explained by month and or geographic identifier.

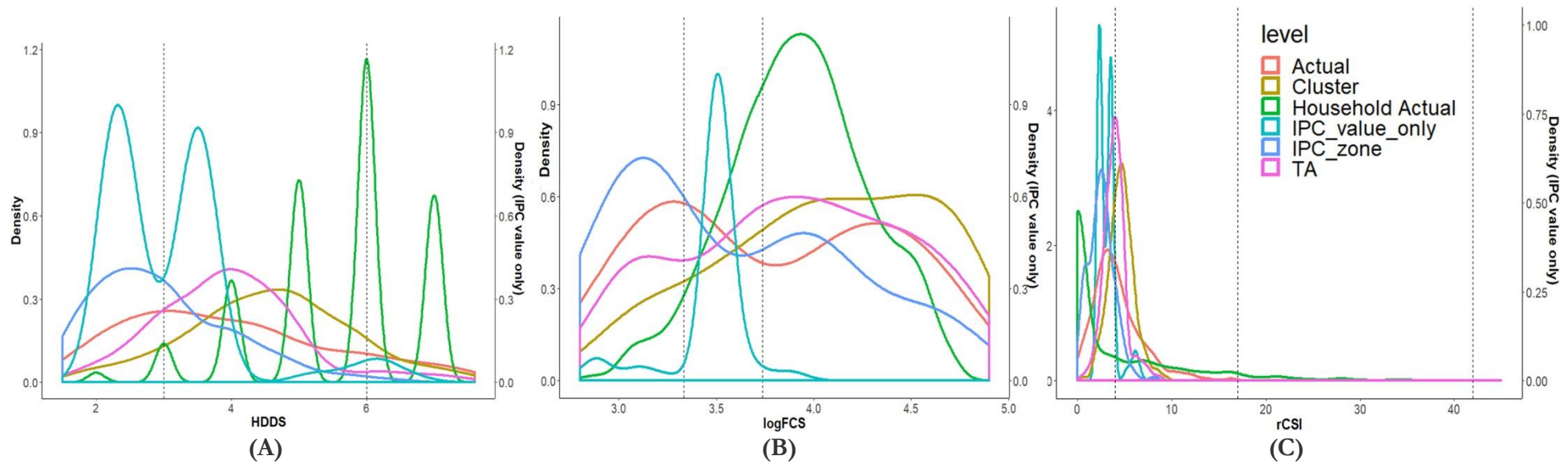


Fig S3: As the level of spatial granularity increases, explanatory power increases and the distribution of the predicated values moves towards the actual distribution both in the center and the spread of the distributions. Panels show, from left to right, Density plots of actual and predicted value of HDDS at different spatial levels (A) Density plots of actual and predicted value of logFCS at different spatial levels (B) Density plots of actual and predicted value of rCSI at different spatial levels (C). The cluster level food security measure averages obtained from the household surveys are shown in red (labeled as “Actual”) as a comparison for our model predictions. Predicted values using variables at the IPC zone, TA and cluster level are shown in blue, purple, and yellow respectively. Prediction values using only the IPC value are shown in teal and are mostly clustered around the mean of the distribution. This prediction leaves out the clusters at the two extremes of the distribution.

Panel A: Full Sample

Panel B: Tail Subsample

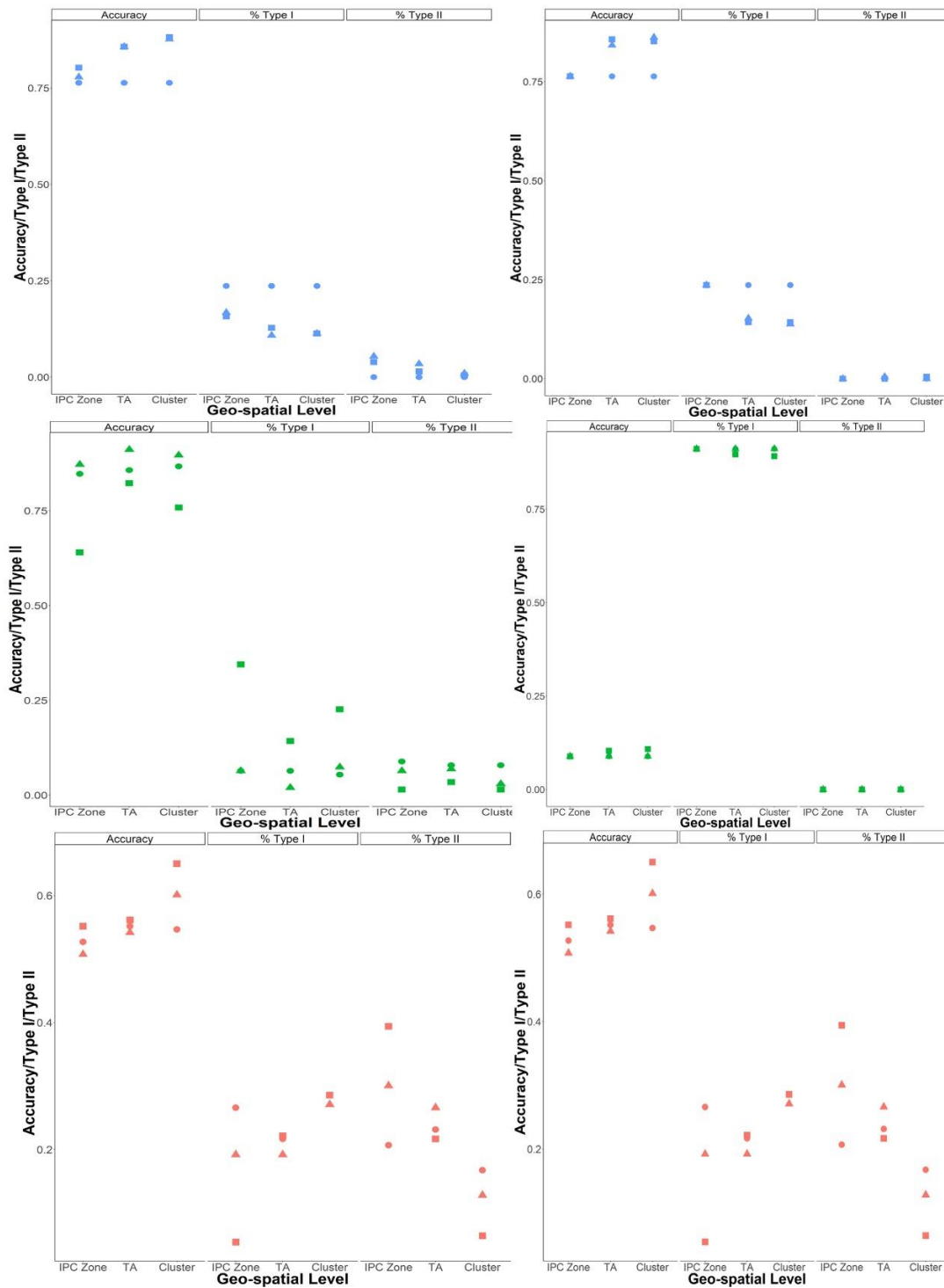


Fig S4: We predict 2013 food security using only 2010 data, limiting both sample(s) to the most food insecure households and confirm that factors that affect food security do not

substantially differ for the subset of the most insecure households. Row 1's blue values represent HDDS. Row 2's green values represent logFCS. Row three's red values represent rCSI. Circles indicate Class 1 data; triangles indicate Class 1 + Class 2; squares indicate Class 1 + Class 2 + Class 3. The "tail" is defined by excluding the households that fall in the most food secure category of each food secure measures: rCSI greater than 4, logFCS smaller than 1.623 and HDDS smaller than 6. The cluster average tail food security measures are based on the cluster average of the remaining households at the tail. The signs and significance of the coefficients on the variables included in the tails-only models closely match estimates from the full distribution. When estimating just the tails of the distribution, the length of dry spells largely negatively affects food security but other measures are consistent between the full distribution and the tails-only distribution. Results of categorical prediction for models including both full-samples and tail subsamples indicate the full sample's accuracy is more sensitive to spatial scale and data class than tail subsample. The accuracy in categorical predictions is similar between model 1 to model 3 for the same food security measure because the predicted values tend to fall into the same category. The same holds for the percent of type I and type II errors.

Table S1: Description of the sources and usage of data used in the Malawi model

Data	Units/interpretation	Collected by	Level of Spatial granularity	Temporal frequency	How the data are used
Dependent Variables					
rCSI	The rCSI measure combines the frequency of use of and severity of five coping strategies. A score of 0-4 indicates food secure; 5-10 indicates moderate food insecurity; 11+ indicates high food insecurity.	2010-2011 and 2013 Malawi LSMS (Household Survey)	Household	Once per household per survey	The LSMS samples different households each month. The resulting sample is representative for the nation, each month and for each district, for the year. We run analyses using both the household and cluster-levels food security measures. Cluster-level measures are the average across households.
HDDS	HDSD possible values range from zero to twelve, with higher values indicating higher food security.				
FCS	FCS cutoffs are 28 and 42 for communities that consume sugar and oil, with higher FCS scores indicating greater security. We log FCS to weight the more food insecure tail of the distribution. A logFCS indicates poor food consumption with a score <1.447; borderline with 1.447 and 1.623; acceptable is > 1.623				
Class 0 Data					
IPC values	1 = minimal ... 5 = famine	2010 - 2017 quarterly IPC values and shape file for IPC zones from FEWs	60 IPC zones (intersections of agro-ecological zones and districts)	Quarterly	FEWS shape files designate IPC-livelihood zones. We use the IPC data to predict the LSMS food security measures and as a benchmark to understand how incorporating additional variables improves our predictions.
Class 1 Data					
Rainfall	Millimeters	CHIRPS	0.05 x 0.05 degree spatial resolution	Daily	LSMS cluster geocodes are deviated by up to 2km to preserve anonymity. We compute average precipitation for each pixel, averaged for each spatial unit. We compute three agronomic-related rainfall variables (1) total precipitation during the prior unimodal rainy season (October to April) (2) we measure the beginning of the rains for the prior agricultural season

					(3) length of dry spells during the past rainy season.
Prices	Maize prices in Malawi kwacha per kg	Malawi Statistical Division of the Ministry of Agriculture	72 markets, including remote and rural markets	Weekly between 2011 and 2017	We draw Thiessen polygons, which assign market prices to each household. We then compute a weighted average marketshed price to each spatial unit of analysis. We compute two price variables (1) monthly price averages by market (see SA on treatment of missing prices) and (2) thinness of the market, measured as the share of missing prices per month.
Soil Quality	1 = limited soil nutrient retention; 0 = ow	World Agroforestry Centre soil maps	500 meter resolution	Once observation/measure per pixel	Data averaged for each relevant unit for spatial analysis from the Moderate Resolution Imaging Spectrometer (MODIS) platform
Elevation	Meters	2010-2011 and 2013 Malawi LSMS (Village Survey)	Village	Once per household per survey	LSMS cluster geocodes are deviated by up to 2km to preserve anonymity. A small amount of random error is introduced by these measures.
Distance to market	Kilometers				
Distance to Road	Kilometers				
Agricultural land	Proportion of agricultural land in 1 km distance				
Class 2 Data					
Roof Type	1 = metal roof; 0 = ow	2010-2011 and 2013 Malawi LSMS (Household Survey)	Household	Once per household per survey	Data are averaged across HH to relevant unit for spatial analysis.
Cell phone ownership	1 = own a cellphone; 0 = ow				
Class 3 Data					
Gender of household head (HH)	1 = male; 0 = female	2010-2011 and 2013 Malawi LSMS (Household Survey)	Household	Once per household per survey	Data are averaged across HH to relevant unit for spatial analysis.
Age of HH	Age				
Number of household members	Count				
Dependency ratio	Number of children (<15) and elderly (>65) divided by total number of HH members				
Asset holdings	Asset index including refrigerator, television, radio, bicycle, motorcycle, and car, computed using principal components analysis				

Table S2: The comparison of model with different specifications. The split by year results are out-of-sample predictions in 2013. Models are estimated at the cluster-level, using the IPC value and include predictors from all Class 1 + Class 2 + Class 3, respectively. The percentage of food insecure clusters that are predicted to be in the insecure category is also known as the true positive rate or “sensitivity” of prediction. The categorical accuracy is the percentage of clusters that are correctly predicted to be the same as the actual food security category.

	R squared			Sensitivity of food Insecure Category			Categorical accuracy		
	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI
Split by year									
Main model	0.649	0.643	0.119	0.833	0.994	0.86	0.897	0.882	0.635
Main model without MSD maize price with MSD maize price in previous month	0.634	0.663	0.089	0.722	1.000	0.892	0.877	0.882	0.611
Main Model without MSD maize price with quarter by region fixed effect	0.558	0.550	0.073	0.444	0.968	0.548	0.887	0.862	0.571
Main Model without MSD maize price with month fixed effect with region fixed effect	0.613	0.603	0.117	0.278	0.968	0.430	0.906	0.877	0.611
Main model with month fixed effect	0.625	0.655	0.126	0.222	0.955	0.624	0.916	0.887	0.596
Split by region									
Main model	0.574	0.660	0.115	0.313	0.772	0.417	0.943	0.769	0.736
Main model without MSD maize price	0.576	0.676	0.063	0.313	0.748	0.222	0.934	0.764	0.693
LASSO Results	0.579	0.671	0.117	0.375	0.862	0.417	0.934	0.788	0.721

Table S3: LASSO regression coefficients. The results are estimated at the cluster-level and include predictors from all Class 1 + Class 2 + Class 3. The split by year results are out-of-sample predictions in 2013. The split by region results are out-of-sample prediction in the North region. Models are estimated at the cluster-level, using the IPC value and include predictors from all Class 1 + Class 2 + Class 3, respectively.

	Split by year			Split by region		
	logFCS	HDDS	rCSI	logFCS	HDDS	rCSI
Best Alpha	1.000	0.000	0.000	0.100	0.100	0.000
Best Lambda	0.001	0.075	0.305	0.019	0.038	0.305
IPC Value (Previous month)	-0.021	-0.089	1.181	-0.015	-0.072	1.125
Total rainfall during rainy season	0.111	0.182	-1.379	0.055	0.048	-0.727
First day of rain	-0.001	-0.002	0.008	0.000	-0.001	0.013
Maximum rainfall in flood prone region	0.000	0.003	-0.018	0.000	0.003	-0.025
Maximum days without rain	.	-0.003	-0.045	.	-0.001	-0.061
Growing degree days	0.001	0.001	-0.053	0.001	0.003	-0.051
Mean temperature	0.018	0.036	0.279	0.007	0.018	0.287
Heat degree days	0.000	-0.002	0.052	-0.001	-0.002	0.050
Maize price	-0.001	-0.002	0.012	.	0.000	0.005
Maize market thinness	-0.014	-0.094	-0.551	-0.015	-0.124	-0.659
Maize price (Previous month)	0.000	0.000	0.020	0.000	0.001	0.005
Maize market thinness (Previous month)	0.006	0.101	-0.051	.	0.050	-0.290
GIEWs maize price	-0.002	-0.002	0.085	0.000	0.002	0.025
Percent of agricultural land	0.014	0.038	-0.137	.	0.038	0.294
Elevation	-0.010	0.054	0.145	-0.018	.	-0.366
Soil nutrient retention constrained	-0.008	-0.069	0.345	-0.004	-0.074	0.530
Distance to nearest road	-0.002	-0.011	-0.006	-0.002	-0.009	-0.005
Distance to nearest Admarc market	0.003	0.004	0.009	0.004	0.007	0.017
Roof type	0.126	0.534	-0.053	0.160	0.693	0.029
Number of cellphones owned	0.140	0.287	-0.696	0.113	0.250	-0.972
Household size	-0.009	-0.026	0.531	0.000	-0.002	0.632
Household age	-0.002	-0.013	-0.026	-0.002	-0.015	-0.022
Household gender	-0.064	-0.039	1.476	-0.084	-0.032	2.459
Asset index	0.112	0.421	-0.360	0.083	0.320	-0.491
Quarter 1	-0.062	-0.185	0.327	-0.073	-0.242	0.732
Quarter 2	-0.010	-0.037	-0.126	.	.	-0.078
Quarter 3	0.001	-0.062	-0.848	0.014	.	-0.622
Central Region	0.017	0.005	0.587	.	0.023	-0.158
South Region	-0.012	-0.154	0.841	.	-0.021	0.157

Table S4: The cluster average tail food security measures are based on the cluster average of the remaining households at the tail. The results are estimated at the cluster-level and include predictors from all Class 1 + Class 2 + Class 3. Standard errors are presented in parentheses and asterisks indicate level of statistical significance of coefficients where three asterisks indicate 1%; two indicate 5% and one indicates 10%.

	logFCS tail	HDDS tail	rCSI tail
IPC Value (Previous month)	-0.012 (0.020)	-0.078 (0.051)	0.553 (0.381)
Total rainfall during rainy season	0.049 (0.051)	0.086 (0.113)	-2.419** (1.104)
First day of rain	-0.001 (0.001)	-0.002 (0.002)	0.045** (0.019)
Maximum days without rain	-0.0002 (0.001)	-0.005** (0.002)	-0.051** (0.024)
Maize price	-0.001 (0.0005)	-0.002 (0.001)	0.016 (0.010)
Maize Market Thinness	0.004 (0.023)	-0.088 (0.059)	-0.909* (0.539)
Percent of agricultural land	0.022 (0.031)	0.040 (0.074)	-0.101 (0.685)
Elevation	-0.020 (0.027)	-0.048 (0.065)	-1.591*** (0.594)
Soil nutrient retention constrained	-0.006 (0.017)	-0.076* (0.043)	0.689* (0.396)
Distance to nearest road	-0.001 (0.001)	-0.010*** (0.002)	-0.013 (0.016)
Distance to nearest admarc market	0.001 (0.001)	0.007** (0.003)	0.029 (0.033)
Roof type	0.036 (0.055)	0.441*** (0.110)	2.413** (0.989)
Number of cellphones owned	-0.017 (0.046)	0.310*** (0.074)	-2.457*** (0.698)
Household size	-0.003 (0.012)	-0.032 (0.028)	0.987*** (0.231)
Household age	-0.0001 (0.002)	-0.011*** (0.004)	-0.062* (0.037)
Household gender	-0.072 (0.065)	0.115 (0.153)	-1.249 (1.398)
Asset index	0.070*** (0.026)	0.476*** (0.062)	-0.895* (0.509)
Quarter 1	-0.013 (0.021)	-0.215*** (0.053)	0.933** (0.423)

Quarter 2	-0.023 (0.024)	-0.083 (0.058)	0.754 (0.498)
Quarter 3	0.013 (0.024)	-0.056 (0.048)	0.060 (0.458)
Observations	208	648	259
R ²	0.109	0.430	0.309

Table S5: General data requirements for Malawi and other countries

Data	Spatial scale required	Possible sources	Assessment of availability	Malawi
Dependent Variables				
Food security measure with geo-codes or spatially granular identifiers.	Household or village-level	from existing data including LSMS, or other existing surveys representative of the country or (sub-national) region of interest; rapid assessments also possible	Variable by country. In some countries, undertaking a rapid, representative food security assessment may be necessary	2010-2011 and 2013 Malawi LSMS (Household Survey)
Class 0 Data				
IPC values		Available for 38 countries, quarterly or semi-annual assessments	Available for all IPC countries	2010 - 2017 quarterly IPC values and shape file for IPC zones from FEWs
Class 1 Data				
Rainfall		CHIRPS	Available for all countries	CHIRPS
Prices		Available from many national statistical offices	Variable coverage of secondary and tertiary markets.	Malawi Statistical Division of the Ministry of Agriculture
Soil Quality		World Agroforestry Centre soil maps	Available for all countries	World Agroforestry Centre soil maps
Elevation	Village-level	Can generate with village coordinates		2010-2011 and 2013 Malawi LSMS (Village Survey)
Distance to market	Village-level	Can generate with village coordinates		
Distance to Road	Village-level	Can generate with village coordinates		
Agricultural land	Village-level	Can generate with village coordinates		
Class 2 Data				
Roof Type	Village-level	Can generate from satellite images using village coordinates	Available for all countries	

Cell phone ownership	Village-level	Census or LSMS or information from cell phone companies	Variable by country. Could be gathered with food security data	2010-2011 and 2013 Malawi LSMS (Household Survey)
<i>Class 3 Data</i>				
Gender of household head (HH)		Census or LSMS	Census available for many countries at spatially disaggregated scale	2010-2011 and 2013 Malawi LSMS (Household Survey)
Age of HH		Census or LSMS	Variable by country. Could be gathered with food security data	
Number of household members		Census or LSMS	Variable by country. Could be gathered with food security data	
Dependency ratio		Census or LSMS	Variable by country. Could be gathered with food security data	
Asset holdings		Census or LSMS	Variable by country. Could be gathered with food security data	