ISSN 2385-2755
DiSSE Working papers
[online]

WORKING PAPERS SERIES
DIPARTIMENTO DI SCIENZE SOCIALI ED ECONOMICHE

Anticipating household vulnerability to food insecurity during large-scale crises

Marco Letta, Pierluigi Montalbano, Cristian Moralez Opazo, Federica Petruccelli

N. 24/2023

SAPIENZA - UNIVERSITY OF ROME
P.le Aldo Moro n.5 – 00185 Roma T(+39) 0649910563
CF80209930587 – P.IVA 02133771002

Anticipating household vulnerability to food insecurity during large-scale crises

Marco Letta

Sapienza University of Rome & Global Labour Organization (GLO)

Pierluigi Montalbano

Sapienza University of Rome & University of Sussex

Cristian Morales Opazo

Food and Agriculture Organization of the United Nations (FAO)

Federica Petruccelli*

Sapienza University of Rome

Abstract

Taking advantage of the World Bank's multi-topic surveys for Nigeria, collected pre and post the COVID-19 pandemic, we assess the ability of a risk sensitive vulnerability measure to anticipate households experiencing food insecurity conditions during the pandemic. The results are disappointing: we document that our vulnerability measure severely underperforms in predicting food-insecure households out-of-sample and that a simple data driven routine, although not able to close the gap, proves to do better even in a data-scarce environment. Sensitivity tests using only the pre-pandemic data show that the poor forecasting performance is not simply due to the discontinuity in the data-generating process (face-to-face vs. phone-based surveys) brought about by the mobility restrictions. This evidence has two important implications: i) there is a need to reconsider the effectiveness of the early warning models used by policymakers to identify vulnerability hotspots; ii) more methodological effort is required to address the limitations of current methods in predicting out-of-sample outcomes.

***JEL-Codes**:* C53, I10, Q12, O12

***Keywords**:* Vulnerability, Food Insecurity, Forecasting, Policy Targeting, COVID-19

* Corresponding author.

E-mail address: federica.petruccelli@uniroma1.it

1. Introduction

Considerable attention has been given to the detrimental effects of the pandemic crisis on the well-being and food security of impoverished individuals in developing countries (Amare et al., 2021; Béné, 2020; Béné et al., 2021; Bundervoet et al., 2022; Egger et al., 2021; Huss et al., 2021; Upton et al., 2021). However, less empirical focus has been placed on investigating whether it was possible to anticipate – so to target in advance – the socioeconomic groups most at risk, identified as vulnerable. In this work, we conduct a retrospective empirical analysis to shed light on this issue. In addressing this matter, we heed the recent call by Upton et al. (2022) to care about whether measurements at our disposal exhibit skills in predicting development outcomes out-of-sample and provide empirical validation of predictive models by testing their forecasting accuracy on publicly available data.

Drawing on household-level microdata taken from multi-purpose household surveys collected pre- and post-COVID-19 by the World Bank in Nigeria, we make two main amendments to the micro vulnerability literature. Firstly, we adapt the "risk-sensitive" longitudinal measure of vulnerability originally proposed by Calvo & Dercon (2005, 2013) by developing a household-level indicator of 'Vulnerability as the Threat of future Food Insecurity' (VTFI). Secondly, we apply a "train-test-compare" approach, commonly employed in the machine learning community, to our vulnerability analysis. This entails conducting a post-shock empirical validation to assess the out-of-sample forecasting performance of our model in predicting households that experienced food insecurity during the COVID-19 period— which is here considered as an exceptional large-scale "natural experiment." Additionally, we replicate this process using a simple machine learning model that employs an interpretable tree-based approach to predict food insecurity status beyond the observed data (Hastie et al., 2009).

We are particularly focused on precisely forecasting households experiencing food insecurity; however, the results of our analysis are disappointing. Our risk-sensitive vulnerability measure, VTFI, significantly underestimates the proportion of households facing food insecurity because of the COVID-19 pandemic, leading to a high number of false negatives. These findings remain consistent even when employing the train-test-compare approach exclusively with the pre-COVID-19 dataset. This check aims to address concerns related to potential disparities in the data generation process between the periods before and after the COVID-19. Conversely, the data-driven approach tends to overpredict the prevalence of food-insecure households. Notably, the latter approach outperforms traditional predictive tools, even when operating in data-scarce environments.

While acknowledging that the two methodologies are not directly comparable, our findings suggest that existing vulnerability measures offer limited and potentially misleading insights to policymakers. This undermines the cost-effectiveness and efficacy of resilience-building programs. Therefore, there is a pressing need to re-evaluate current targeting mechanisms, develop improved profiling methodologies, and conduct more empirical validation of predictive models for welfare outcomes. This issue is particularly acute in developing contexts where data scarcity is common (Aiken et al., 2022).

Before delving into the details of our analysis, it is important to clarify two preliminary points. Firstly, our focus in this study is not on examining the direct effects of the pandemic itself. Rather, we utilize the COVID-19 pandemic as a "stress test" to validate the empirical methodologies employed. Secondly, we acknowledge that *ex-ante* vulnerability predictions and *ex-post* observed outcomes are different metrics, and simple comparisons between these two statistics should be taken with a grain of salt. This is because the concept of vulnerability is an inherently forward-looking construct, and as such, it is neither directly observable nor linked to the actual manifestation of shocks (Imai et al., 2011; Magrini et al., 2018). Additionally, we recognize that the pandemic shocks differed from the routine shocks experienced by households in the pre-COVID-19 period (although the endogenous components of vulnerability should remain the same). Furthermore, we are also mindful of various methodological caveats and concerns in our work including data comparability, data scarcity, absence of *ex-ante* identification of the characteristics and probabilities associated with the pandemic "state of the world". We duly take these caveats into careful consideration both in our empirical analysis and robustness checks.

From a methodological viewpoint, our work contributes to the well-established theoretical and empirical literature on vulnerability to poverty and food insecurity and adverse shocks (Bogale, 2012; Calvo & Dercon, 2005, 2013; Chaudhuri et al., 2002; Gallardo, 2018; Hoddinott & Quisumbing, 2003; Magrini et al., 2018; Povel, 2015; Sileshi et al., 2019). It also aligns with the flourishing strand of works that leverage data-driven and cross-validation methods coupled with survey information in the service of poverty and food insecurity targeting, mapping, and monitoring (see, among others, Aiken et al., 2022; Aiken et al., 2023; Browne et al., 2021; Garbero & Letta, 2022; McBride et al., 2021; Zhou et al., 2021). Specifically, our methodological contribution is to make theory-based and data-driven literature meet advocating for a forecasting culture based on interpretability, domain knowledge, and economic intuition.

The rest of this paper is arranged as follows: the subsequent section reviews the relevant literature, Section 3 describes the context and data, Section 4 illustrates the empirical strategy and the predictive models, Section 5 presents the results of the empirical analysis, Section 6 shows the sensitivity checks, and Section 7 concludes and provides policy recommendations.

2. Literature review

In order to provide a comprehensive understanding of our contribution, it is essential to contextualize it within the relevant scientific literature. In this section, we will review two key strands of literature that inform our work on *ex-ante* profiling methodologies. The first strand is the traditional vulnerability literature, which is firmly grounded in standard microeconomic theory. The second strand pertains to the emerging research area that harnesses machine learning techniques for poverty and food insecurity mapping, monitoring, and targeting. Additionally, we will provide a summary of the literature concerning the impacts of the pandemic on the welfare of impoverished individuals in developing nations, particularly within the context of the food system.

2.1. The vulnerability literature

The recognition of significant movements in and out of poverty has shifted the focus towards household vulnerability as a basis for social protection strategies. However, assessing risk and vulnerability is challenging due to the varying definitions of vulnerability and the limited availability of data to measure it (Hoddinott and Quisumbing, 2003). Timely identification of vulnerable populations is crucial, as delays in providing assistance to the appropriate regions can result in inadequate support (Headey and Barrett, 2015). Upton et al. (2022) highlight that conventional measures of poverty and food security often work well in targeting chronically poor or food-insecure individuals but perform poorly in identifying temporary deprivation. Therefore, they are less effective in responding to major shocks when most needed. In contrast, vulnerability assessments serve as an *ex-ante* targeting criterion, allowing for the identification of individuals in greatest need before the occurrence of shocks. At the micro level, vulnerability to poverty is the most widely used notion of vulnerability. It refers to the likelihood that an individual or household will experience a level of welfare, typically proxied by consumption, below a fixed benchmark, such as the poverty line, in the near future (Hoddinott & Quisumbing, 2003). Although this notion of vulnerability is strictly linked to poverty, the two concepts are different. Vulnerability is a dynamic concept that looks toward the future and is unobservable at any point in time (Chaudhuri et al., 2002). It is said to be *ex-ante* and forward-looking because it is measured at time t, before the shock occurs (*ex-ante*), and refers to the probability of being poor at time $t+1$ (forward-looking). Conversely, poverty is an effective outcome that is observable at time t. It refers to the *ex-post* static status in which the individual lives at the exact moment it is observed and measured (Gallardo, 2018). Correlates of vulnerability may differ from those of poverty. This distinction plays a crucial role in designing policies and targeting only the poor could exclude a significant group of individuals who risk experiencing a welfare loss in case of downside shocks (Montalbano, 2011).

One of the most robust "risk-sensitive" vulnerability measures, in the presence of panel data, is the one introduced by Calvo & Dercon (2005, 2013). They define vulnerability as a probability-weighted average of future indices of deprivation in different "states of the world". This measurement combines households' deprivation and shortfalls in welfare indicators with exposure to risks and satisfies the so-called "focus axiom", according to which the burden of future deprivation will not be compensated by any possible positive outcomes. Building upon the pioneering work of Christiaensen & Boisvert (2000), vulnerability measurements have recently been adapted to the field of food insecurity (Bogale, 2012; Das, 2021; Ibok et al., 2019; Mutabazi et al., 2015; Sileshi et al., 2019). However, a lack of shared understanding regarding fundamental terms, data requirements, and the interpretation of vulnerability to food insecurity has hindered progress in theory and methods. It is crucial to adopt a forward-looking approach that not only identifies currently food-insecure households but also those susceptible to shocks and risks such as natural disasters and extreme climate conditions. This approach enhances the effectiveness of safeguarding households from the detrimental impacts of such shocks. Therefore, similar to poverty, it is important to differentiate between food insecurity and vulnerability to food insecurity. Food insecurity indicators provide static measurements of food shortage for a household at a specific moment. On the other hand, vulnerability to food insecurity involves dynamic calculations that incorporate risks and shocks households might face, potentially affecting food consumption levels (Sileshi et al., 2019). In this context, vulnerability refers to the household's probability of falling below the food poverty line (Capaldo et al., 2010; Løvendal & Knowles, 2005; Sileshi et al., 2019).

In this study focusing on Nigeria, we propose an empirical application of the risk-sensitive measure introduced by Calvo & Dercon (2005), named vulnerability as the threat of food insecurity (VTFI), utilizing a hybrid threshold as proposed by Povel (2015).

2.2. Machine learning applications to poverty and food insecurity

In the last few years, in the wake of the increasing use of data science and artificial intelligence techniques in economics, as well as due to the recent availability of large amounts of information in the form of big and non-conventional data sources requiring computationally-intensive tools, a new body of research has focused on the leveraging of the machine- and deep-learning predictive algorithms for the mapping, targeting and monitoring of a variety of wellbeing outcomes.

Among the first to apply these tools in development economics, Blumenstock et al. (2015) use anonymized mobile phone data from Rwanda to show the potential of feature engineering and elastic net regularization techniques in predicting poverty and wealth status and generating high-resolution maps of poverty and wealth from call records. Jean et al. (2016) instead couple high-resolution daytime satellite imagery from five African countries with convolutional neural networks to predict local-level poverty. McBride & Nichols (2018) employ USAID poverty assessment tools and data to re-calibrate Proxy Means Test poverty-targeting tools towards the prioritization of out-of-sample performance, in place of the traditional in-sample focus, through cross-validation and stochastic ensemble methods. A recent work from Aiken et al. (2023) shows the importance of combining different data sources. They find that machine learning methods using mobile phone data can accurately identify ultra-poor households nearly as well as survey-based measures. By integrating mobile phone data with survey measurements, classifications achieve greater accuracy compared to those derived solely from a singular data source.

Machine learning routines have been also applied to predict food security and resilience outcomes. Hossain et al. (2019) employ machine learning models, such as random forest and extreme gradient boosting, as well as traditional methods on survey data from Bangladesh to predict household food insecurity measured through the caloric intake. Lentz et al. (2019) combine high-resolution market data, remote sensing information, and survey data to forecast household-level food security status in Malawi using Least Absolute Shrinkage and Selection Operator (LASSO) and logit models. Zhou et al. (2021), emphasizing the need for interpretability and transparency of predictive models for policy targeting, leverage gradient boosting, and random forest to predict household food insecurity in villages from three African countries. As for resilience, Knippenberg et al. (2019) apply LASSO and random forests to identify the best predictors of a resilience measure based on the Coping Strategy Index of Malawian households, while Garbero & Letta (2022) use a cross-country survey dataset and a battery of machine learning algorithms to predict household resilience to shocks. More recently, Haushofer et al. (2022) employ machine learning – specifically, generalized random forests – on data from an NGO Cash Transfer program implemented in Kenya to show that households that are social welfare maximizing to target, namely those delivering the largest treatment effects, are not those predicted to be most deprived.

In a recent overview, McBride et al. (2021) advocate the use of machine learning for poverty and food insecurity targeting, mapping, and monitoring, to fine-tune policy efforts on the ground and improve the design of effective early-warning mechanisms. At the same time, the authors warn about several important caveats regarding a wise policy-oriented use of these powerful tools. For instance, while the sudden availability of big data may seem to free researchers and agencies from the chronic microdata gaps in developing contexts, most of the approaches cited above still depend on traditional survey data for ground-truthing and validation of the models developed using remote sensing or other non-conventional information. McBride et al. (2021) also note that it is only possible to predict with accuracy states and processes that have been previously observed in data, and this presents a challenge for the development of early-warning systems and suggests that research should also focus on running simulations of extreme scenarios in the development of targeting models and prevention tools. Another crucial point is the distinction between contemporaneous and sequential prediction. While contemporaneous prediction is useful for poverty and malnutrition mapping, *sequential* prediction, which we call forecasting, is the analytical tool necessary for early-warning and targeting purposes (Browne et al., 2021; Tang et al., 2021). Finally, and differently from other domains, the importance of some degree of interpretability of model outputs should not be understated when it comes to the application of machine learning routines in the service of policy targeting. In this respect, the well-known trade-off between accuracy and transparency could lean more in favour of the latter at the expense of a loss in predictive accuracy (Browne et al., 2021; McBride et al., 2021).

2.3. The food security impacts of the COVID-19 crisis

Upton et al. (2021) highlight that the COVID-19 pandemic is only a seemingly unique event. Indeed, it has merely exposed the systemic and pre-existing vulnerabilities of rural populations situated within complex food systems in developing countries. As such, the pandemic-induced food security dynamics reflect the broader issues facing vulnerable rural populations confronting structural deprivation in the midst of myriad shocks and stressors (Béné et al., 2016). In sum, the pandemic acted as a threat amplifier and multiplier of shocks that are neither new nor rare to poor households in developing countries. Similarly, in a systematic review of the resilience capacity of local food systems, Béné (2020) emphasizes that the major direct impacts of the COVID-19 on the food system have been brought about by its effects on the income and purchasing power levels for all food system actors caused by non-pharmaceutical interventions (for example, mobility restrictions and lockdowns), and the subsequent negative effect these had on their access to food. Income losses and food insecurity are thus strictly interconnected, and their joint dynamics are key in explaining the pandemic’s direct and indirect effects on the food system. Upton et al. (2021) also highlight how rural households are exposed to food system shocks, particularly the shock produced by the COVID-19, not just as food producers but equally in their roles as food consumers or as workers within the broader agri-food value chain.

In a study on nine countries from Africa, Asia, and Latin America with phone-survey data on more than 30,000 respondents, Egger et al. (2021) document a generalized decline in income, employment, and food security (captured using self-reported missing or reduced meals) that began with the outset of the pandemic (March 2020). Using representative and high-frequency phone-survey data on 41,000 households from 31 countries (collected by the World Bank), Bundervoet et al. (2022) find that more than one-third of respondents lost their jobs and around two-thirds suffered income losses. Similar insights for geographic areas can be drawn from the results of Vu et al. (2022) for Vietnam: using a Bartik-type IV shift-share instrument approach; the authors find small changes in food insecurity risk at the national level during the pandemic, but substantial heterogeneity at a more granular district level, with a subsample of more vulnerable districts severely affected.

Overall, international agencies confirm that food security indicators have continued to deteriorate globally despite social protection interventions and extensive measures by national and international actors (FAO et al., 2021). The role of vulnerability in this context is central: vulnerable households determine the fragility of the entire food system as producers, processors, retailers, vendors, and consumers (FAO, 2021). Enhancing the resilience of supply chains and local agri-food systems, it is crucial to identify vulnerability hotspots across the food system and to understand how they respond to different shocks (Reddy et al., 2016). In conclusion, the literature suggests that a better understanding of vulnerability to food insecurity should be considered as policy priorities to increase resilience to future shocks (Béné et al., 2021; Bundervoet et al., 2022).

3. Context and data

Nigeria is the most populous African country; it was one of the first African countries affected by the pandemic and among the first to introduce non-pharmaceutical restrictions to control the spread of the virus. The pandemic and the collapse of international oil prices that followed posed a severe threat to Nigeria's already fragile economy (World Bank, 2020). Prior to the pandemic, Nigeria had been slowly recovering from the recession in 2016. However, the country still faced high poverty rates and significant exposure to local and global food price volatility, which further magnified the challenges posed by the COVID-19 pandemic (Benson et al., 2020). In addition to these economic pressures, weather shocks affecting agricultural production compounded the adverse effects on household consumption, particularly for vulnerable rural households with limited assets (Amare et al., 2018). A study by Bundervoet et al. (2022) reveal that 85% of Nigerian respondents stated that income from non-farm household enterprises and activities declined or entirely disappeared in the immediate aftermath of the pandemic. By matching the World Bank's LSMS-ISA pre-COVID-19 data with post-COVID-19 data (taken from High-Frequency Phone Surveys) for Nigeria, the study by Amare et al. (2021) demonstrates a staggering a 50% increase in household food insecurity compared to the pre-pandemic period. Moreover, 79% of households reported having suffered a reduction in total income since the start of the pandemic. Further, by carrying out an *ex-post* impact evaluation through a difference-in-differences identification strategy, the authors find that mobility restrictions had significant adverse effects on food security outcomes, with disproportionate impacts on poor households and those living in conflict-ridden areas.

Our research involves the use of combined datasets, comprising data collected before the pandemic gathered through standard face-to-face surveys – which are employed to predict vulnerable households in the vulnerability analysis – and phone-survey data collected after the pandemic – which are used to compare *ex-ante* predictions with post-COVID-19 observed outcomes in the targeting test. For the vulnerability analysis, wherein we predict households vulnerable to food insecurity and income loss, we rely on the Living Standard Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA)[1]. These data are nationally representative surveys collected by the World Bank in collaboration with national governments.

We use the panel dataset comprises four waves (2010-2011, 2012-2013, 2015-2016, and 2018-2019), which support the redesign and implementation of the General Household Survey (GHS). The GHS-Panel sample consists of 5,000 households, which is a subsample of the GHS core survey of 22,000 households. To ensure the continued integrity and representativeness of the sample, a partial refresh of the GHS-Panel sample was carried out during the fourth wave of the panel.

Regarding the sample size, a distinction must be made between VTFI and data-driven approaches. For the former, longitudinal data is required, so we only kept households re-interviewed in all the following waves, for which we have a complete set of information. This selection process led to a sample of 1,250 households per wave, with a full sample of 5,000 observations. However, for the data-driven approach, we can employ a larger sample size. In this case, machine learning models can be estimated without longitudinal information. The only requirement is that households in the post-COVID-19 data also appear in the last pre-COVID-19 survey, enabling us to match pre-pandemic input variables with post-pandemic outcomes for each household. As explained in more detail below, we train a machine learning model on the pre-COVID-19 data (the training set) and use it to predict post-COVID-19 outcomes (the testing set) and assess its forecasting performance. The total sample size is 3,715 household observations in the pre-COVID-19 data.

For the targeting test, we use data from the LSMS-Supported High-Frequency Phone Surveys on

[1] Data are available at: https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA#6

COVID-19 implemented in 2020.[2] The samples from this survey were selected from the sample of households interviewed in 2018/2019. The interviews consist of a 15-minute questionnaire covering topics such as employment status, household livelihoods, income losses, and coping strategies, access to basic needs, and knowledge of the COVID-19 and mitigation measures. We use the first round of the phone survey, carried out between April and May 2020, since it contains information on self-reported income losses. For the theory-based approach, the dataset is combined with the panel of the pre-COVID-19 data to allow a comparison between pre- and post-COVID-19 outcomes, leading to a final sample for the targeting test of 316 households. The sample size is relatively small because the vulnerability measure we employ requires longitudinal data. For this reason, we can implement the vulnerability-based targeting test only for households interviewed in all four waves before the pandemic and in the post-COVID-19 period. For the data-driven approach, we do not have these limitations for the reasons explained above. We can therefore use a larger sample size of 1,311 observations in the post-COVID-19 dataset.

Based on the availability of the data and the different methodologies used (theory based and data-driven), we explore different outcome variables. This clarification is important to underly that the two exercises are not directly comparable, but the machine learning practice is only applied to test these techniques using the available limited data. The theory-based approach requires the use of a continuous dependent variable. In this context, we begin by applying the logarithm to spatially and temporally adjusted total gross income, denominated in the Nigerian currency, the Naira. Subsequently, we establish its connection with the binary food insecurity outcome, as detailed below. Total gross income is given by the sum of farm and non-farm income. The farm income variable includes the gross value of the crop harvested and the value of livestock products and animals sold for that season. Non-farm income refers to self-employment job earnings from activities that differ from agricultural production. This variable is constructed by multiplying the monthly reported value of sales by the months of business activity in the previous twelve months. The total income variable includes the sum of the two categories of income described above, and all the income variables are expressed as logarithmic values. The outcome variable used for the post-COVID-19 analysis is the self-reported income losses since the pandemic's beginning. Households were asked whether the income from their major sources of livelihood has reduced, stayed the same, or increased.

Income losses have been identified as a significant driver of the COVID-19 pandemic's impact on food systems and food security in developing countries (Béné, 2020; Egger et al., 2021; Bundervoet et al., 2022). This observation holds true for the Nigerian context as well (Amare et al., 2021). Given

[2] Data are available at: https://microdata.worldbank.org/index.php/catalog/3712

our primary objective of addressing food insecurity through the income loss mechanism, we create an outcome variable assigned a value of 1 when a household encounters both income losses and food insecurity, and 0 otherwise. Food insecurity is measured through households' self-reported responses to the Food Insecurity Experience Scale (FIES) questions. The questionnaire contains the last three of the eight questions that are part of the standard FIES module,[3] capturing the more severe aspects of food insecurity. We classify a household as food insecure if it reports having experienced at least one of the food insecurity conditions asked about in the FIES module. This choice follows up the most recent food security literature that largely use subjective food security metrics à la FIES for identifying food security conditions in the aftermath of the COVID-19 (Bundervoet et al., 2022; Egger et al., 2021). In the methodology below, this variable will be used as the benchmark to identify the households vulnerable to food insecurity.

For the data-driven approach, as explained in more detail below, we train a machine learning model on the pre-COVID-19 data (the training set) and use it to predict post-COVID-19 outcomes (the testing set) and assess its forecasting performance. Unlike vulnerability models that are estimated using *continuous* variables for income or consumption, the machine learning routine we employ – a classification tree – only needs data on the same binary outcome variable for the training and testing sets to be employed. Moreover, the phone-survey post-COVID-19 data include only basic information and a much more limited set of essential variables compared to the pre-COVID-19 face-to-face surveys. For these reasons, we employ an outcome variable represented by a simple composite indicator of food insecurity using the available household-level FIES questions, namely a dummy variable taking value 1 if a household experienced *at least one* of the three food insecurity conditions and 0 otherwise. This is possible because we have data for this variable in both the pre- and post-COVID-19 surveys.

The main explanatory variables included in both analyses are basic household demographics, education characteristics, assets, and shocks. As for the latter, we use five different potential shocks, namely the three idiosyncratic shocks reported as the most severe by the households and two weather shocks. These are i) the death of a working member of the household, ii) the illness of an income-earning member of the household, iii) the rise in the price of major food items consumed, iv) drought, and v) flood. While it may appear that the shocks we have considered differ in their causes from the multiple shocks generated from the pandemic, they exhibit similar consequences and activate the same endogenous mechanisms of fragility for households. It is essential to approach the analysis of the pandemic within the broader context of ongoing risks and vulnerabilities faced by marginalized

[3] See Cafiero et al. (2018) and https://www.fao.org/in-action/voices-of-the-hungry/fies/en/.

populations. In this respect, it has been already stressed that the pandemic has exacerbated pre-existing vulnerabilities to different risk factors that rural populations in poor countries regularly face (Upton et al., 2021). It has magnified the effects of existing vulnerabilities and exposed communities to a range of intersecting challenges, including economic downturns, disruptions in supply chains, reduced access to resources, and increased food insecurity.

To this end, in our empirical analysis we adopted a combination of proxies for shocks. These include self-reported episodes directly experienced by households and two objective measures of weather shocks which are computed using historical rainfall data. In fact, the dataset also includes a set of georeferenced variables, which are used to construct negative and positive values of rainfall anomalies from which the dummies for drought and flood are constructed.[4] Table A.1 in Appendix A reports the description of the entire set of variables used for both analyses, while Table A.2 displays their descriptive statistics.

We use *lagged* input variables with the only exception being the shock variables. In other words, for each wave, the outcome variable is at time *t*, whereas all the input variables, except contemporaneous shock dummies, are at time *t-1* and thus are taken from the previous wave. In this way, we can use the entire set of household characteristics available from the last pre-COVID-19 LSMS-ISA survey (2018-2019) as inputs employed, jointly with shock dummies, to predict the post-COVID-19 food insecurity outcome variable.

As is common practice for poverty and resilience assessments (Chaudhuri et al., 2002; Calvo & Dercon, 2013; Povel, 2015; Ligon & Schechter, 2003; Barrett & Constas, 2014; Constas et al., 2014; Cissé & Barrett, 2018), our analyses make use of thresholds to discriminate vulnerable and non-vulnerable households. Inspired by Povel (2015), we here use a hybrid threshold using both FIES indicators and the initial level of welfare of each individual. This allows us to predict as vulnerable those individuals who feel food insecure but also experienced an objective loss of welfare.

---

[4] Following Dell et al. (2014), we construct the measure of rainfall anomalies as a deviation of the level of rainfall in the previous twelve months from the historical average as follows:

Rainfall anomalies: $\left(\frac{R_{it}-R_i}{R_{it}^{SD}}\right)$

where R_{it} indicates the last twelve months' rainfall at the location of household *i* for year *t*. R_i is the historical average rainfall at the location of household *i,* calculated based on the local government area (lga). R_{it}^{SD} is the standard deviation of rainfall at the location of household *i.* Second, we identify as droughts rainfall anomalies under the 25th percentile of the total distribution and as heavy rains/floods values of anomalies over the 75th percentile of the distribution. Third, dummy variables for droughts and floods are created; the first takes value 1 if the household experienced drought anomalies, and the second takes value 1 if the household experienced heavy rains/flood rainfall anomalies.

.

4. Empirical strategy

To provide a comprehensive retrospective evaluation of the accuracy of theory-based and data-driven methodologies in predicting households affected by the COVID-19 shock, we apply the following methodological steps: i) we first develop forecasts of vulnerable households using data from the pre-COVID-19 face-to-face surveys; ii) we then generate confusion matrices to carry out an out-of-sample validation of the forecasts using the post-COVID-19 data from the phone surveys.

4.1. Vulnerability as the threat of food insecurity (VTFI)

To assess vulnerable households, we adopt a risk sensitive measure of "vulnerability as the threat of food insecurity" (VTFI), inspired by Calvo & Dercon (2005; 2013) with the extension proposed by Povel (2015). This approach allows us to estimate household-specific vulnerability measurements that consider the possible "states of the world" that households could face, coupled with their respective probabilities. Using information about the occurrence of the shocks and estimating the related loss of income, we can predict the deprivation indexes associated with all the different states of the world considered, which are represented by the different combinations of shocks that a household might potentially face, listed in Section 3. This provides us with an *ex-ante* measurement of household vulnerability. Precisely, we measure vulnerability as follows:

$$VTFI_i = \sum_{j=1}^{N_i} \left(p_{ij} \times x_{ij}^{\alpha}\right) \tag{1}$$

where $N_i = \sum_{k=0}^{K_i} \frac{K_i!}{(K_i-k)!k!}$ represents the number of possible states of the world. In our case, we consider five shocks so $N_i = \sum_{k=0}^{5} \frac{5!}{(5-k)!k!} = 32$ states of the world. p_{ij} denotes the probability that the state of the world j will occur and ranges between zero and one. It is measured as:

$$p_{ij} = \prod_{q=1}^{Q_{ij}} p_{ijq} \times \prod_{l=1}^{L_{i,i\neq q}} \left(1 - p_{ijl}\right) \tag{2}$$

where $\prod_{q=1}^{Q_{ij}} p_{ijq}$ yields the probability that Q_i risks will occur in state of the world j while $\prod_{l=1}^{L_{i,i\neq q}} \left(1 - p_{ijl}\right)$ represents the probability that L_i other shocks will not occur in the same situation. x_{ij}^{α} indicates the deprivation index and is measured as $x_{ij} = \sum_{q=1}^{Q_{ij}} \frac{s_{ijq}}{y_i}$ where s_{ijq} represents the severity of shock q in state of the world j, namely the loss of income (log) in that state of the world, and y_i is the threshold and represents the household income (log). To target households vulnerable to food insecurity, inspired by the hybrid method proposed by Povel (2015), we combine the income loss with a food insecurity line constructed using FIES data.[5] To estimate income losses associated to the experience shocks, assuming that shocks are independent and identically distributed (IID), we calculate the elasticities of income to each shock separately, as follows:

$$y_{it} = \alpha_i + \beta_1 \, shock_{it} + \beta_2 X'_{it-1} + \varepsilon_{it} \tag{3}$$

y_{it} refers to the household i log of total income at time *t*; α_i captures household fixed effects that absorb many potential time-invariant confounders, and $shock_{it}$ indicates the shocks that enter the regression separately. X'_{it-1} is a set of household characteristics, including age, sex and education of the head of household, household size and a set of welfare indicators (Tropical Livestock Units and an index for the non-agricultural assets held by the household) at time *t-1* that enable an *ex-ante* computation of the outcome variable since explanatory variables at time *t-1* cannot be affected by the shock at time *t*. The coefficient β_j captures the percentage change in the log income given by each shock. To obtain a household-specific value for the loss experienced, we multiply the estimated coefficients for each shock by each household income values:

$$loss_{it} = \beta_{ji} \, * y_{it} \tag{4}$$

To control for household-specific characteristics, we re-estimate the predicted loss (*ploss*) using the following regression:

$$ploss_{it} = \alpha_i + \beta_j \, X'_{it-1} + \varepsilon_{it} \tag{5}$$

α_i captures household fixed effects. X'_{it-1} includes the same variables as Equation (3). To satisfy the focus axiom, we divide each predicted household-specific *ploss* s_{ijq} by household total income y_i and replace negative losses (gains) with zeros. Finally, we sum all the different losses to get an overall index of deprivation x_{ij} that household i faces in the different states of the world *j* as follows:

$$x_{ij} = \sum_{q=1}^{Q_{ij}} \frac{s_{ijq}}{y_i} \tag{6}$$

As we assume that households are risk-averse, we set the parameter α=2 so that we square the value x_{ij}.[6] Once we have state- and household-specific deprivation indexes, as in Calvo & Dercon (2005,

[5] A viable alternative to FIES could be the use of income or consumption data. Unfortunately, in the post-COVID-19 phone survey data from Nigeria these data were not collected.
[6] The parameter $\boldsymbol{\alpha}$ regulates the strength of "risk sensitivity" of our vulnerability measure, where 1 means risk neutrality (Povel, 2015).

2013) and Povel (2015), we calculate the household-associated independent probabilities of experiencing shocks in the state of the world j by using a logit model as follows:

$$Pr(shock_{it}) = F(X'_{it-1}) \quad (7)$$

$shock_{it}$ at time *t* is predicted using explanatory variables from time *t-1* and X'_{it-1} includes age, sex and education of the head of the household and household size, and a set of welfare variables as in Equation (3). The household- and state-specific probabilities are then computed as in Equation (2) and multiplied by the household- and state-specific index of deprivation previously calculated in Equation (6).

The VTFI is then calculated using Equation (1). Households that are vulnerable to income loss are identified as those above the median value of the vulnerability distribution which is the standard cut-off universally employed for classification tasks (Hastie et al., 2009)[7]. Furthermore, since the COVID-19 pandemic represents a rare event in the form of a major exogenous shock, we also separately consider the worst-case scenario, which refers only to the state of the world in which all the shocks occur simultaneously, and we assign this a probability of occurrence equal to one, for reasons that will be explained in detail in the next section. Because we are interested in food insecurity through the channel of income losses, following Povel (2015), we use a hybrid threshold to discriminate vulnerable from non-vulnerable households that includes a food insecurity cut-off represented by the FIES[8] in the analysis of income loss. Hence, the final sample of vulnerable households comprises those that are vulnerable to income losses and that also reported having experienced food insecurity conditions before the COVID-19 pandemic. In the targeting test, we will then compare these food-insecure households predicted as vulnerable by our VTFI measure with households experiencing income losses and food insecurity during the pandemic (post-COVID-19 dataset), in line with the cross-country evidence provided by Bundervoet et al. (2022) and Egger et al. (2021) in relation to the close link between pandemic-induced income losses and heightened food insecurity at the household level. As clarified above, we cannot assume that all the pre-COVID-19 vulnerable households should be necessarily affected by income losses and food insecurity in the post-COVID-19 as households may lose out because of unlucky realisations, or *vice versa* can be better off for upcoming reasons that were not under control *ex-ante*. Although this hampers a strict correspondence between the two statistics, it is undeniable that a certain degree of positive correlation should be in place across the two measures and that certain thresholds in terms of both sensitivity and

---

[7] Additional outcomes, obtained by employing the 25th and 75th percentiles as cut-off points, are reported in Section 6.
[8] More details can be found in Section 3.

accuracy should be met. This assessment is key for making choices in terms of appropriate targeting.

4.2. Classification tree

As already mentioned above, we also explore the performance of a data driven approach to test whether this could represent a potential path to follow for future investigations in targeting analyses. To this end, we employ a simple and straightforward machine learning predictive routine named classification trees (Hastie et al., 2009). Unlike the previous vulnerability model, which is strongly rooted in standard economic theory, classification trees are purely data-driven predictive techniques aimed at maximizing the out-of-sample predictive performance of a given outcome of interest.

At the core of the machine learning mindset lies the "firewall" principle: none of the data involved in generating the predictive model should be used to evaluate its predictive performance (Mullainathan & Spiess, 2017). In this spirit, we use pre-COVID-19 data to tune our classification tree and select the best model and then use it to predict the food insecurity outcome for the unseen observations in the post-COVID-19 dataset.

Classification trees are particularly suited for applications in which the decision rule needs to be transparent and must be communicated (Lantz, 2019), which is in line with recent calls for interpretability in the use of machine learning for policy targeting (Browne et al., 2021; McBride et al., 2021). Indeed, the output of a decision tree is intuitive and can also be easily understood by someone without strong statistical skills, such as decision-makers, which makes this technique appealing for policy targeting purposes. From a technical point of view, classification trees are based on a process called recursive binary splitting: the algorithm divides the data into progressively smaller subsets to identify recurring patterns that can be used for predicting a specific binary output. The model chooses the variables and split-probability that maximizes the so-called *information gain* which indicates the signals (information) we get from each variable. The information gain is calculated using the *entropy* which is a measure of impurity that controls where the classification tree decides to split the data. The aim is to find variables and split points that will produce the purest subsets, thus having the lowest entropy value possible.[9] Trees are considered highly flexible methods because non-linearities and interactions are automatically captured by the sequence of splits in the tree. A drawback of classification trees is that they tend to be prone to overfitting on the training data:

[9] The model looks for feature values that split the data in partitions that contain primarily a single class. A segment of data that contains a single class is said to be pure. One criterion to measure purity is the entropy measure. As an indication of how mixed class values are in a sample, the entropy is measured from 0 to 1, where 0 indicates maximum homogeneity, and 1 indicates maximum disorder. The algorithm calculates homogeneity changes resulting from a split based on entropy. This calculation is called information gain. When a feature has a high information gain, it is more likely to create homogeneous groups after a split. When the information gain is zero, splitting this feature will not reduce entropy (Lantz, 2019).

a high number of branches and leaves in a tree is likely to overfit the data, leading to a final model which performs very well in-sample but poorly out-of-sample. The solution to this issue is to reduce the size and complexity of the tree by "pruning" the tree. Pruning means setting a penalization cost for flexibility, which is referred to as a "complexity parameter" (*cp*). To select the optimal value of *cp,* which maximizes the out-of-sample accuracy of our model, we run 10-fold cross-validation on the full set of pre-COVID-19 data, compare the ten-resulting cross-validation errors, and apply the complexity parameter associated with the lowest cross-validation error. This *cp* is then selected for the model used to predict unseen observations belonging to the held-out post-COVID-19 data.

Besides transparency, another critical reason why we select classification trees from many available machine learning techniques is that we have substantial missing data in our samples. Just like traditional econometric approaches, most machine learning routines do not handle missing data unless they are imputed. In our case, given the numerous cases where values for several variables are missing for many observations, this would imply either accepting a drastic loss in sample size (which is already limited) or opting for massive imputation of missing values, which would very likely result in data distortion. Unlike other methods, classification trees automatically handle missing data using surrogate splits in the case of missing observations (Lantz, 2019).

Before performing our classification exercise, we must tackle the challenge of data imbalance of the primary outcome of interest, food insecurity. In fact, in the pre-COVID-19 sample, food insecurity (as measured using the composite FIES indicator described above) is, from a statistical point of view, a "rare" event, meaning that there are many more food-secure households than food-insecure ones. When facing data imbalance issues like this, machine learning routines simply tend to predict the over-represented class (Y = 0 in our case). This happens because the algorithms aim to maximize overall accuracy and provide the lowest total error rate, irrespective of which class the errors come from. Clearly, a model that predicts all households as food secure would be useless for our purpose. To prevent this from happening, we use a popular data rebalancing technique, the Synthetic Minority Oversampling Technique (SMOTE) routine by Chawla et al. (2002). Using SMOTE, we can rebalance the frequency of two classes of our outcome variable, food insecurity, in the pre-COVID-19 sample. Specifically, SMOTE oversamples the under-represented cases and undersamples the majority class, leading to a smaller but rebalanced dataset. Clearly, we run SMOTE only on the pre-COVID-19 dataset, leaving the post-COVID-19 data untouched.

As we want to assess how accurate, given available pre-shock information, machine learning is in identifying vulnerable households before a food security crisis occurs, we need to simulate the occurrence of the shock in the covariate space of the post-COVID-19 dataset. We do so by adopting a “trick” that leads the algorithm to “believe” that a massive sequence of shocks has taken place in the post-COVID-19 data: we take all the shock dummies available in the last pre-COVID-19 wave (food price spikes, droughts, floods, illnesses, and deaths of household members) and artificially switch them to 1 for all the observations in the sample. Through this technique, the algorithm, which has previously learned from the pre-COVID-19 data that shocks tend to be associated with food insecurity status, “recognizes” the abrupt change in the patterns of key features and consequently also takes its decisions on the classification of observations based on this change. It should be noted that a need for the simulation of unobserved events and large-scale shocks has recently been emphasized in the specialized literature (McBride et al., 2021).

Conceptually, by doing this, we are mimicking the occurrence of the COVID-19 generalized disruption by simulating, for the entire sample, the simultaneous occurrence of all the shocks and stressors that most households reported facing at various points in time across all waves. This simulation relies on a reasoning analogous to that behind the worst-case scenario in the VTFI model and draws heavily on the argument advanced by Upton et al. (2021), already reported above that the COVID-19 shock is less a new shock or unique event than a dramatic, recent manifestation of familiar shocks, stressors, and uncertainties that burden poor populations in developing contexts.

4.3. Targeting evaluation

As Upton et al. (2022) explain, a key performance test of any measurement, especially one estimated in a population-representative sample, is represented by its out-of-sample performance, namely when applied to observations not from the original estimation sample. Therefore, we compare the predictions generated by our developed models with the actual outcome data from the COVID-19 phone surveys. This comparison is carried out using confusion matrices (Lantz, 2019; Hastie et al., 2009), a straightforward tool widely employed to evaluate predictive performance. In the case of classification problems with a binary outcome variable, like the one in our study, it consists of a simple two-way table. For a comprehensive explanation, please refer to Table B.1 in Appendix B.

5. Empirical results

This section contains the outcomes of the vulnerability-based targeting analysis for both the VTFI and the classification tree discussed earlier.

5.1. VTFI model

Table 1 compares the predicted VTFI for the scenario in which we consider all the different states of the world pre-COVID-19 (Equation 1) with the real status of households that experienced both income losses and food insecurity in the post-COVID-19 period. The results show that our model successfully predicts 88% of the households that did not experience an income loss coupled with food insecurity after the pandemic. This indicates a high "specificity" of the model, which refers to its ability to accurately identify negative cases[10] (Y = 0). However, the model performs poorly in predicting vulnerable households, with a "sensitivity" value of only 16%. Sensitivity measures the model's ability to correctly classify positive cases (Y = 1), in this case, households that are food insecure. In our pursuit of evaluating the effectiveness of models in identifying households vulnerable to food insecurity during significant shocks, our primary focus clearly lies on the sensitivity performance. The observed low sensitivity value indicates limitations in the model's ability to accurately identify and forecast households that are at risk of food insecurity.

Table 1. Vulnerability targeting performance – VTFI model

		Real status		
		Income losses & FI = 0	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	91	178	269
	Vulnerability = 1	13	34	47
	Total	104	212	316
	Correctly Predicted	**88% Specificity**	**16% Sensitivity**	**39% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity calculated considering all the states of the world using the train set (before the COVID-19). This takes value 1 if the household is classified as vulnerable and 0 otherwise. The real status represents the observed income losses and food insecurity outcome in the test set (after the COVID-19). Income losses & FI is a combined binary variable which takes value 1 if the household reported both to have lost income and experienced food insecurity conditions during the pandemic.

The predictive model's inability to predict vulnerable households results in an overall accuracy rate of only 39%. This is primarily because the majority of households (212 out of 316) in the phone-survey data reported experiencing both income losses and food insecurity. It appears that the vulnerability-based targeting approach is structurally unable to accurately predict that a significant proportion of households would be susceptible to income losses and food insecurity in the event of a major shock such as the COVID-19 crisis.

5.2. Classification tree

Figure 1 below shows the classification tree constructed using pre-COVID-19 information from the LSMS-ISA face-to-face surveys. The tree is composed of those variables that have been automatically selected by the algorithm as the ones more correlated with the outcome variable, the FIES food

[10] The negative cases refer to the non-food-insecure households and thus the food-secure ones.

security indicator. These variables include a dummy referring to the female gender of the household head, a dummy capturing flood shocks, a variable indicating the value of Tropical Livestock Units for small ruminants, and three asset measurements, namely the agricultural and non-agricultural wealth indices, and the overall wealth index (combining both productive and non-productive assets).

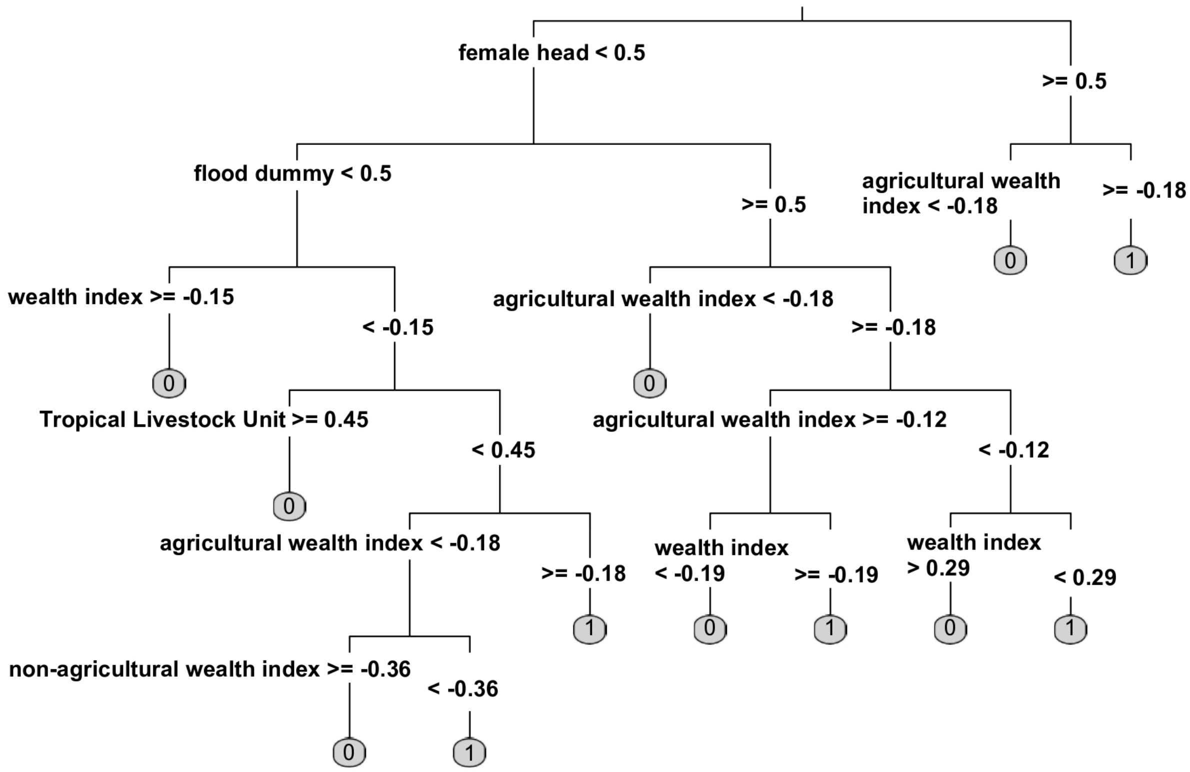

Figure 1. Classification tree of food-insecure households

These variables interact with each other and, depending on whether the value of each variable is below or above the reported thresholds, the tree classifies each observation in the pre-COVID-19 training set as food insecure (Y=1) or not (Y=0). It is interesting to note that although the relationships depicted in the tree only signal correlation and not causation, the variables selected by the algorithm are consistent with the findings on the drivers and determinants of poverty traps highlighted by the well-established poverty trap literature (Barrett & Carter, 2013; Carter & Barrett, 2006; Carter et al., 2008). This means that, despite being quite simple, the tree is able to detect critical recurring patterns embedded in the structure of the data.

How accurate is this tree in forecasting households that have experienced food insecurity conditions during the COVID-19 pandemic? The answer can be found in the confusion matrix shown in Table 2 below: 80% of food-insecure households during the COVID-19 crisis have been correctly predicted by the algorithm, thanks to the "multiple-shock" simulation that we artificially introduced into the post-COVID-19 shock data. The tree is, however, much less accurate in identifying non-food-insecure households, with a specificity of only 29%. For this reason, the overall forecasting accuracy of the model is good, 69%, but not high.

Table 2. Tree-based targeting performance

		Real status		
		FIES = 0	**FIES = 1**	**Total**
Predicted status	**FIES = 0**	82	202	284
	FIES = 1	205	822	1,027
	Total	287	1,024	1,311
	Correctly Predicted	**29% Specificity**	**80% Sensitivity**	**69% Accuracy**

Notes: the predicted status refers to the predicted out-of-sample food-insecure households using the train set (before the COVID-19). The real status represents the observed food-insecure households of the test set (after the COVID-19). FIES is a binary outcome that takes value 1 if the household is predicted (predicted status) or reported (real status) to be food insecure and 0 otherwise.

Despite operating in a relatively data-poor environment, the simple machine learning model depicted in Figure 1 can anticipate more than three out of four households experiencing food insecurity due to the COVID-19 shock. This gives an idea of the significant potential of machine learning techniques in identifying food insecurity, and vulnerability hotspots, in line with the findings of other recent contributions (Aiken et al., 2022; Garbero & Letta, 2022; McBride et al., 2021; Zhou et al., 2021). However, the targeting performance of the classification tree is still imperfect and, overall, unsatisfactory, as the tree tends to overpredict food insecurity and maximize sensitivity at the expense of specificity and overall accuracy.

6. Sensitivity checks

In this section of the paper, we present a series of sensitivity checks to further examine the robustness of our findings. Sensitivity checks are crucial to assess the stability and reliability of our main results, ensuring that they are not overly dependent on specific modelling assumptions or data variations. By conducting these additional analyses, we aim to address potential concerns and strengthen the overall validity of our research.

6.1. Different vulnerability cut-offs

It is important to highlight that the model's underperformance showed in Table 1 is not exclusively linked to the choice of the median as the cut-off for distinguishing between vulnerable and non-vulnerable households. This is evident, as reported in Table 3, where we present supplementary results utilizing alternative cut-off points, specifically the 25th and 75th percentiles of the vulnerability distribution. Remarkably, the outcomes remain consistent, suggesting that the model's suboptimal performance is not reliant on a specific threshold choice.

Table 3. Vulnerability targeting performance – VTFI model – different cut-offs

		Real status		
25th percentile cut-off		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	83	165	284
	Vulnerability = 1	21	47	68
	Total	104	212	316
	Correctly Predicted	**80% Specificity**	**22% Sensitivity**	**41% Accuracy**
75th percentile cut-off		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	95	197	292
	Vulnerability = 1	9	15	24
	Total	104	212	316
	Correctly Predicted	**91% Specificity**	**7% Sensitivity**	**35% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity calculated considering all the states of the world using the train set (before the COVID-19). This takes value 1 if the household is classified as vulnerable, defined as those above the 25th and 75th percentile of the VTFI distribution, and 0 otherwise. The real status represents the observed income losses and food insecurity outcome in the test set (after the COVID-19). Income losses & FI is a combined binary variable which takes value 1 if the household reported both to have lost income and experienced food insecurity conditions during the pandemic.

6.2. Worst-case scenario

The disappointing outcome from the VTFI model results could be attributed to the fact that the COVID-19 is indeed an extreme scenario which could be mimicked by the simultaneous occurrence of the many idiosyncratic and covariate shocks that plague households' welfare in developing contexts (Upton et al., 2021). However, the estimated vulnerability model assigns a very low probability to such a pandemic "state of the world" since it is considered a "fat-tail risk", namely a devastating but extremely rare event.

To address this issue, a sensitivity analysis can be conducted on the pre-COVID-19 data by replacing the vulnerability measurement with one calculated under the assumption that all shocks occur simultaneously. This approach trains the model to consider a rare event as certain (probability assigned = 1) and may improve its ability to identify vulnerable households in the face of a "systemic risk". Table 4 displays the outcome for this worst-case scenario.

Although this model specification seems to perform slightly better in terms of sensitivity and accuracy, the results are still disappointing, meaning that the targeting failures are not due to an inherent inability of standard vulnerability models to spot and adequately consider extreme scenarios and catastrophic disasters.

Table 4. Vulnerability targeting performance (worst-case scenario) – VTFI model

		Real status		
		Income losses & FI = 0	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	78	152	230
	Vulnerability = 1	26	60	86
	Total	104	212	316
	Correctly Predicted	**75% Specificity**	**28% Sensitivity**	**44% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity calculated considering only the worst-case scenario (the household experiences all the shocks simultaneously) using the train set (before the COVID-19). This takes value 1 if the household is classified as vulnerable and 0 otherwise. The real status represents the observed income losses and food insecurity outcome in the test set (after the COVID-19). Income losses & FI is a combined binary variable which takes value 1 if the household reported both to have lost income and experienced food insecurity conditions during the pandemic.

6.3. Data comparability issue

The discrepancy in data collection between the pre- and post-COVID-19 surveys poses an additional area of concern. It is important to highlight that, despite differences in data collection methods (face-to-face versus phone-based), both survey types have undergone rigorous validation, demonstrating a high degree of statistical equivalence. The phone-based data have been carefully adjusted to address statistical biases resulting from the distinct collection modes. Notably, even though the pandemic has concluded, these phone-based surveys have now officially become part of the World Bank's LSMS-ISA collection. This inclusion aims to enhance the frequency and timeliness of panel surveys in client countries. Currently, they are integrated into ongoing data collection efforts in several countries, including Nigeria, and are regularly conducted alongside face-to-face surveys (Gourlay et al., 2021).

However, to gauge the magnitude of this concern, we also employ the same analysis using only the pre-COVID-19 data for both the prediction and the test sets: the former is composed of the first three waves of the panel, and the latter is tested using the fourth wave. This test thus uses a fully harmonized pre-pandemic dataset of standard field surveys.

Our main conclusions are not overturned by this check, and the results, summarized in Table 5, still report a disappointing performance in terms of predictive accuracy, suggesting that the poor performance is not linked to disparities in the data collection methods employed before and after the onset of COVID-19.

Table 5. Vulnerability targeting performance (only using pre-COVID-19 dataset) – VTFI model

		Real status		
		Income losses & FI = 0	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	833	260	1,093
	Vulnerability = 1	63	42	105
	Total	896	302	1,198
	Correctly Predicted	**93% Specificity**	**14% Sensitivity**	**73% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity calculated using the train set (the first three waves pre-COVID-19). This takes value 1 if the household is classified as vulnerable and 0 otherwise. The real status represents the observed income losses and food insecurity outcome reported in the test set (the fourth wave pre-COVID-19). Income losses & FI is a combined binary variable which takes value 1 if the household reported both to have lost income and experienced food insecurity conditions in the wave before the COVID-19.

6.4. Subsample analysis by treatment (lockdown) status

In addressing concerns about the targeting approach, we acknowledge that one objection may be related to the inclusion of the full sample size of households in the post-COVID-19 phone surveys, without distinguishing between those exposed to non-pharmaceutical interventions (lockdowns, mobility restrictions, etc.) and those merely experiencing the spread of the contagion and the spillovers of the global pandemic. The COVID-19 pandemic has been global and determined a black-out of within-country and international supply chains, triggering a deep recession that left no communities unaffected, especially in developing contexts. Indeed, as reported in the data section, even the questions of the Phone Surveys themselves make explicit reference to the arrival of the COVID-19 pandemic in asking interviewed households whether they had experienced income losses or food insecurity, regardless of their exposure to lockdown measures.

We argue that all households in the post-COVID sample should be considered "treated" in the sense that they have been exposed to the COVID-19 crisis; therefore, it is essential to include all households in the targeting analysis. Nonetheless, to address this concern, we also conducted a subsample analysis focusing on households exposed to lockdowns (treated) and those not subjected to mobility restrictions (untreated), using the separate samples of the "lockdown" variable as employed by Amare et al. (2021).

These results are reported below in Table 6 for households exposed to lockdowns (*treated group*) and in Table 7 for those not exposed to mobility restrictions (*untreated group*), for the analyses covering both all the states of the world and the worst-case scenario. Overall, the estimates are not significantly different and confirm the unsatisfactory performance of the targeting model in identifying the hardest-hit households.

Table 6. Vulnerability targeting performance – VTFI model – *Treated group*

		Real status		
All the states of the world		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	20	40	60
	Vulnerability = 1	2	7	9
	Total	22	47	69
	Correctly Predicted	**91% Specificity**	**15% Sensitivity**	**39% Accuracy**
Worst-case scenario		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	20	36	56
	Vulnerability = 1	2	11	13
	Total	22	47	69
	Correctly Predicted	**91% Specificity**	**23% Sensitivity**	**45% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity both for all the states of the world and the worst-case scenario (the household experiences all the shocks simultaneously) using the train set (before the COVID-19). This takes value 1 if the household is classified as vulnerable and 0 otherwise. The real status represents the observed income losses and food insecurity outcome in the test set (after the COVID-19) and in the subgroups of household exposed to lockdowns (treated). Income losses & FI is a combined binary variable which takes value 1 if the household exposed to lockdowns during the pandemic reported both to have lost income and experienced food insecurity conditions.

Table 7. Vulnerability targeting performance – VTFI model – *Untreated group*

		Real status		
All the states of the world		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	71	138	209
	Vulnerability = 1	11	27	38
	Total	82	165	247
	Correctly Predicted	**87% Specificity**	**16% Sensitivity**	**40% Accuracy**
Worst-case scenario		**Income losses & FI = 0**	**Income losses & FI = 1**	**Total**
Predicted status	**Vulnerability = 0**	58	116	174
	Vulnerability = 1	24	49	73
	Total	82	165	247
	Correctly Predicted	**71% Specificity**	**30% Sensitivity**	**43% Accuracy**

Notes: the predicted status refers to the combined vulnerability to income losses and food insecurity both for all the states of the world and the worst-case scenario (the household experiences all the shocks simultaneously) using the train set (before the COVID-19). This takes value 1 if the household is classified as vulnerable and 0 otherwise. The real status represents the observed income losses and food insecurity outcome in the test set (after the COVID-19) and in the subgroups of household not exposed to lockdowns (untreated). Income losses & FI is a combined binary variable which takes value 1 if the household not exposed to lockdowns during the pandemic reported both to have lost income and experienced food insecurity conditions.

7. Discussion and conclusion

In developing emergency interventions to address large-scale shocks, governments face the challenge of rebalancing exposure to myriad shocks in real time (Upton et al., 2021). On the other hand, knowing in advance which groups and communities are most at risk in case of a future shock would enable *ex-ante* preventive policies and resilience-building efforts aimed at minimizing exposure before the shock occurs. This is even more urgent in the current historical period, where new economic and political tensions are emerging, increasing the overall risk exposure of the most vulnerable households and their liquidity and behavioural constraints. For instance, the recent war in Ukraine caused a rise in food prices and fossil fuels, compromising access to food and food production for low-income countries and vulnerable groups (IMF, 2022). Similarly, due to its unpredictable nature, the COVID-19 pandemic caught most governments unprepared and revealed a shortage of both preventive and absorptive policies.

In our study we conducted a COVID-19 stress test not to assess the well-documented welfare impacts of the pandemic *ex-post*, but rather to retrospectively evaluate the *ex-ante* forecasting ability of a well-established vulnerability measure in predicting food insecurity and informing policy targeting. Acknowledging the inherent challenges of making an exact comparison between *ex-ante* and *ex-post* statistics, we estimated the out-of-sample performances of the theory-based and data-driven approaches to the prediction of vulnerability to food insecurity. The results revealed disappointing performances, with the vulnerability model failing to accurately anticipate households experiencing food insecurity during the COVID-19 pandemic. Further examinations indicate that the poor performance observed is not attributable to discrepancies in the data collection method between pre- and post-COVID-19. In fact, similar results persist when conducting the analysis exclusively with the pre-COVID-19 dataset, thereby utilizing a fully harmonized dataset.

On the other hand, we also tested whether a simple machine learning approach (classification tree), which is specifically suited for out-of-sample predictions, exhibits some skills in predicting a development outcome even when operating in the data-poor environment typical of developing contexts. While still suboptimal, the results show that the classification tree still fare much better than the theory-based model, suggesting that machine learning could represent a potential ally for the development of good targeting models.

Readers may be concerned that models' forecasting failures are not due to their intrinsic shortcomings in out-of-sample performances but to the 'black swan' nature of the COVID-19 crisis. Thus, we stress again that the development literature agrees that, while unique in its causes, the pandemic shock is akin to a multi-stressor shock whose consequences have merely amplified the pre-existing vulnerabilities to different risk factors to which rural populations living in poor countries are routinely exposed (Upton et al., 2021). The lack of a satisfactory forecasting model hampers cost-effectiveness of the predictive tool and raises concerns, particularly considering that vulnerable households play a crucial role in determining the fragility of the entire food system chain, involving producers, processors, retailers, vendors, and consumers (FAO, 2021). From a methodological perspective, the results are in line with the evidence recently provided by Upton et al. (2022) on the disappointing out-of-sample performance of popular resilience metrics, and reinforce the view that predictive models of poverty, vulnerability, resilience, and other development outcomes should always be subject to rigorous empirical validation of their forecasting ability on held-out, unseen data.

These findings are subject to a number of caveats: i) the presence of non-trivial data issues (for example, the focus on income losses rather than consumption behaviour due to the lack of consumption data in the post-COVID-19 surveys, the use of subjective measures of food insecurity, discrepancies in comparing different datasets) and the fact that all the methodologies were employed in data-poor environments with small sample sizes, which is especially penalizing for data driven routines; ii) the decision to proxy shock probabilities and corresponding severities using relatively limited panel data; iii) the lack of a rigorous identification of all the potential mechanisms at play in the vulnerability framework (for example, the hypothesis of non-transferability across states of the world of insurance mechanisms, whereby households can smooth away variations in outcomes over states of the world); and iv) the assumption underlying the simulations, namely that the COVID-19 shock can be proxied by the simultaneous occurrence of all the food insecurity and income shocks households routinely face in rural developing contexts.

Despite these caveats, we highlight a lack of appropriate forecasting and preventive tools. While mapping and targeting subnational vulnerability hotspots and food insecurity pockets is an adamant policy priority, standard models failed to address this need in the case of the COVID-19 shock. Therefore, efforts should be dedicated to refining targeting mechanisms and developing improved profiling methodologies to inform preventive interventions. This would enable policymakers to take action before shocks occur and implement early-warning systems to identify vulnerability hotspots in anticipation of future food system crises. In turn, this would enhance the cost-effectiveness and efficacy of resilience-building programs aimed at strengthening local agri-food systems' capacity to absorb shocks.

From a data-oriented perspective, our work underscores the fact that predictive models, regardless of their theoretical soundness or computational power, are limited by the availability of data in data-scarce environments. Lastly, improving the interoperability of traditional survey data with non-conventional data sources, such as big data, crowdsourced data, and citizen-generated information, appears to be a promising route (Aiken et al., 2023).

References

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. E. (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature*, *603*(7903), 864-870.

Aiken, E. L., Bedoya, G., Blumenstock, J. E., & Coville, A. (2023). Program targeting with machine learning and mobile phone data: Evidence from an anti-poverty intervention in Afghanistan. *Journal of Development Economics*, *161*, 103016.

Amare, M., Abay, K. A., Tiberti, L., & Chamberlin, J. (2021). COVID-19 and food security: Panel data evidence from Nigeria. *Food Policy*, *101*, 102099.

Amare, M., Jensen, N. D., Shiferaw, B., & Cissé, J. D. (2018). Rainfall shocks and agricultural productivity: Implication for rural household consumption. *Agricultural systems*, *166*, 79-89.

Barrett, C. B., & Carter, M. R. (2013). The economics of poverty traps and persistent poverty: Empirical and policy implications. *The Journal of Development Studies*, *49*(7), 976-990.

Barrett, C. B., & Constas, M. A. (2014). Toward a theory of resilience for international development applications. *Proceedings of the National Academy of Sciences*, *111*(40), 14625-14630.

Béné, C. (2020). Resilience of local food systems and links to food security–A review of some important concepts in the context of COVID-19 and other shocks. *Food Security*, 1-18.

Béné, C., Al-Hassan, R. M., Amarasinghe, O., Fong, P., Ocran, J., Onumah, E., ... & Mills, D. J. (2016). Is resilience socially constructed? Empirical evidence from Fiji, Ghana, Sri Lanka, and vietnam. *Global Environmental Change*, *38*, 153-170.

Béné, C., Bakker, D., Rodriguez, M. C., Even, B., Melo, J., & Sonneveld, A. (2021). *Impacts of COVID-19 on people's food security: Foundations for a more resilient food system*. Intl Food Policy Res Inst.

Benson, T., Amare, M., & Ogunniyi, A. (2020). *The relative commercial orientation of smallholder farmers in Nigeria: Household and crop value-chain analyses* (Vol. 66). Intl Food Policy Res Inst.

Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, *350*(6264), 1073-1076.

Bogale, A. (2012). Vulnerability of smallholder rural households to food insecurity in Eastern

Ethiopia. *Food Security*, *4*(4), 581-591.

Browne, C., Matteson, D. S., McBride, L., Hu, L., Liu, Y., Sun, Y., ... & Barrett, C. B. (2021). Multivariate random forest prediction of poverty and malnutrition prevalence. *PloS one*, *16*(9), e0255519.

Bundervoet, T., Dávalos, M. E., & Garcia, N. (2022). The Short-Term Impacts of COVID-19 on Households in Developing Countries. *World Development*, 153, 105844.

Cafiero, C., Viviani, S., & Nord, M. (2018). Food security measurement in a global context: The food insecurity experience scale. *Measurement*, *116*, 146-152.

Calvo, C., & Dercon, S. (2005). Measuring individual vulnerability. Department of economics discussion paper series, 229. Oxford: Oxford University.

Calvo, C., & Dercon, S. (2013). Vulnerability to individual and aggregate poverty. *Social Choice and Welfare*, *41*(4), 721-740.

Capaldo J, Karfakis P, Knowles M, Smulders M (2010) A Model of Vulnerability to Food Insecurity. ESA Working Paper Series No.10-03. Rome: Food and Agriculture Organisation of the United Nations

Carter, M. R., & Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies*, *42*(2), 178-199.

Carter, M. R., Little, P. D., Mogues, T., & Negatu, W. (2008). Poverty traps and natural disasters in Ethiopia and Honduras. In *Social Protection for the Poor and Poorest* (pp. 85-118). Palgrave Macmillan, London.

Chaudhuri, S., Jalan, J., & Suryahadi, A. (2002). Assessing household vulnerability to poverty from cross-sectional data: A methodology and estimates from Indonesia. *Department of Economics Discussion Paper Series 0102-52, Columbia University.*

Christiaensen, L. J., & Boisvert, R. N. (2000). On measuring household food vulnerability: Case evidence from Northern Mali.

Chawla, N. V., & Bowyer, K. W., & Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321-357.

Cissé, J. D., & Barrett, C. B. (2018). Estimating development resilience: A conditional moments-based approach. *Journal of Development Economics*, *135*, 272-284.

Constas, M., Frankenberger, T., & Hoddinott, J. (2014). Resilience measurement principles: Toward an agenda for measurement design. *Food Security Information Network, Resilience Measurement Technical Working Group, Technical Series*, *1*.

Das, M. (2021). Vulnerability to food insecurity: A decomposition exercise for rural india using the expected utility approach. *Social Indicators Research*, *156*(1), 167-199.

Dell, M., Jones, B. F., & Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature*, *52*(3), 740-98.

Egger, D., Miguel, E., Warren, S. S., Shenoy, A., Collins, E., Karlan, D., ... & Vernot, C. (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science advances*, *7*(6), eabe0997.

FAO, IFAD, UNICEF, WFP and WHO. (2021). The State of Food Security and Nutrition in the World 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all. Rome, FAO. Available at: https://doi.org/10.4060/cb4474en

FAO. 2021. *The State of Food and Agriculture 2021. Making agrifood systems more resilient to shocks and stresses*. Rome, FAO. Available at: https://doi.org/10.4060/cb4476en

Gallardo, M. (2018). Identifying vulnerability to poverty: A critical survey. *Journal of Economic Surveys*, *32*(4), 1074-1105.

Garbero, A., & Letta, M. (2022). Predicting household resilience with machine learning: preliminary cross-country tests. *Empirical Economics*, 1-14.

Gourlay, S., Kilic, T., Martuscelli, A., Wollburg, P., & Zezza, A. (2021). High-frequency phone surveys on COVID-19: good practices, open questions. *Food Policy*, *105*, 102153.

Hastie, T., & Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

Haushofer, J., Niehaus, P., Paramo, C., Miguel, E., & Walker, M. (2022). Targeting impact versus deprivation. *CEGA Working Paper Series*, No. WPS-199. Center for Effective Global Action. University of California, Berkeley. Text. https://doi.org/10.5072/FK22F7TN60

Headey, D., & Barrett, C. B. (2015). Measuring development resilience in the world's poorest countries. *Proceedings of the National Academy of Sciences*, *112*(37), 11423-11425.

Hoddinott, J., & Quisumbing, M. A. R. (2003). *Data sources for microeconometric risk and vulnerability assessments*. World Bank, Social Protection.

Hossain, M., Mullally, C., & Asadullah, M. N. (2019). Alternatives to calorie-based indicators of food security: An application of machine learning methods. *Food policy*, *84*, 77-91.

Huss, M., Brander, M., Kassie, M., Ehlert, U., & Bernauer, T. (2021). Improved storage mitigates vulnerability to food-supply shocks in smallholder agriculture during the COVID-19 pandemic. *Global Food Security*, *28*, 100468.

Ibok, O. W., Osbahr, H., & Srinivasan, C. (2019). Advancing a new index for measuring household vulnerability to food insecurity. *Food Policy*, *84*, 10-20.

Imai, K. S., Gaiha, R., & Kang, W. (2011). Vulnerability and poverty dynamics in Vietnam. *Applied Economics*, *43*(25), 3603-3618.

International Monetary Fund. 2022. World Economic Outlook: War Sets Back the Global Recovery. Washington, DC, April.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790-794.

Knippenberg, E., Jensen, N., & Constas, M. (2019). Quantifying household resilience with high frequency data: Temporal dynamics and methodological options. *World Development*, *121*, 1-15.

Lantz, B. (2019). *Machine learning with R: expert techniques for predictive modeling*. Packt Publishing Ltd.

Lentz, E. C., Michelson, H., Baylis, K., & Zhou, Y. (2019). A data-driven approach improves food insecurity crisis prediction. *World Development*, *122*, 399-409.

Ligon, E., & Schechter, L. (2003). Measuring vulnerability. *The Economic Journal*, *113*(486), C95-C102.

Løvendal CR, Knowles M (2005) Tomorrow's hunger: a framework for analyzing vulnerability to food insecurity. Food and Agriculture Organization of the United Nations, ESA Working Paper No.05-07. Agricultural and Development Economics Division, Rome

Magrini, E., Montalbano, P., & Winters, L. A. (2018). Households' vulnerability from trade in Vietnam. *World Development*, *112*, 46-58.

McBride, L., Barrett, C. B., Browne, C., Hu, L., Liu, Y., Matteson, D. S., ... & Wen, J. (2021). Predicting poverty and malnutrition for targeting, mapping, monitoring, and early warning. *Applied Economic Perspectives and Policy*, 1-18.

McBride, L., & Nichols, A. (2018). Retooling poverty targeting using out-of-sample validation and machine learning. *The World Bank Economic Review*, *32*(3), 531-550.

Montalbano, P. (2011). Trade openness and developing countries' vulnerability: Concepts, misconceptions, and directions for research. *World Development*, *39*(9), 1489-1502.

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87-106.

Mutabazi, K. D., Sieber, S., Maeda, C., & Tscherning, K. (2015). Assessing the determinants of poverty and vulnerability of smallholder farmers in a changing climate: the case of Morogoro region, Tanzania. *Regional Environmental Change*, *15*(7), 1243-1258.

Povel, F. (2015). Measuring exposure to downside risk with an application to Thailand and Vietnam. *World Development*, *71*, 4-24.

Reddy, V. R., Singh, S. K., & Anbumozhi, V. (2016). Food supply chain disruption due to natural disasters: Entities, risks, and strategies for resilience. *ERIA Discussion Paper*, *18*.

Sen, A. (1982). *Poverty and famines: an essay on entitlement and deprivation*. Oxford university press.

Sileshi, M., Kadigi, R., Mutabazi, K., & Sieber, S. (2019). Analysis of households' vulnerability to food insecurity and its influencing factors in East Hararghe, Ethiopia. *Journal of Economic Structures*, *8*(1), 1-17.

Tang, B., Liu, Y., & Matteson, D. S. (2021). Predicting poverty with vegetation index. *Applied Economic Perspectives and Policy*, 1-16.

Upton, J., Constenla-Villoslada, S., & Barrett, C. B. (2022). Caveat utilitor: A comparative assessment of resilience measurement approaches. *Journal of Development Economics*, 157, 102873.

Upton, J., Tennant, E., Florella, K., & Barrett, C. B. (2021). Household Resilience, and Rural Food Systems: Evidence from Southern and Eastern Africa. *Applied Economics and Policy Working Paper Series,* No.2021-17. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3992671

Vu, K., Vuong, N. D. T., Vu-Thanh, T. A., & Nguyen, A. N. (2022). Income shock and food insecurity prediction Vietnam under the pandemic. *World Development*, *153*, 105838.

World Bank. 2020. Nigeria Development Update, June 2020: Nigeria in Times of COVID-19 - Laying Foundations for a Strong Recovery. World Bank, Washington, DC. © World Bank.

Zhou, Y., Lentz, E., Michelson, H., Kim, C., & Baylis, K. (2021). Machine learning for food security: Principles for transparency and usability. *Applied Economic Perspectives and Policy*, 1-18.

Appendix A: Variables description and descriptive statistics

Table A.1. Definition of the variables included in the analysis

Variable name	Definition	Time period	Source
	Vulnerability analysis (VTFI) – Machine learning analysis		
Total income	Log of the total household gross income (non-farm, wage, farm, and other income)	2011 – 2019	LSMS - ISA
Death shock	= 1 if the household has been affected by the death of a working member of the household	2011 – 2019	LSMS - ISA
Illness shock	= 1 if the household has been affected by the illness of a working member of the household	2011 – 2019	LSMS - ISA
Food price shock	= 1 if the household experienced an increase in the price of major food items consumed	2011 – 2019	LSMS - ISA
Drought shock	= 1 if the household experienced a drought shock	2011 – 2019	LSMS - ISA
Heavy rains/Flood shock	= 1 if the household experienced a flood shock	2011 – 2019	LSMS - ISA
Age of the household head	Age of the household head, in years	2011 – 2019	LSMS - ISA
Female household heads	= 1 if the household head is female	2011 – 2019	LSMS - ISA
Household head can read and write	= 1 if the household head can read and write	2011 – 2019	LSMS - ISA
Household size	Number of members of the household	2011 – 2019	LSMS - ISA
Urban households	= 1 if the household live in an urban area	2011 – 2019	LSMS - ISA
Non-agricultural wealth index (*)	Index of non-agricultural assets held by the household	2011 – 2019	LSMS - ISA
TLU large ruminants	Tropical Livestock Unit - large ruminants held by the household	2011 – 2019	LSMS - ISA
TLU small ruminants	Tropical Livestock Unit - small ruminants held by the household	2011 – 2019	LSMS - ISA
TLU other animals	Tropical Livestock Unit - other animals held by the household	2011 – 2019	LSMS - ISA
Year of the interview	Year in which the household has been interviewed	2011 – 2019	LSMS - ISA

(*) A factor analysis has been used to construct the variable indicating the non-agricultural wealth index, it includes 34 durable assets held by the household.

Table A.2. Descriptive statistics

Variable name	Obs	Mean	SD	Min	Max
		Vulnerability analysis (VTFI)			
Total income (log)	4,973	7.599	2.92	0	13.802
Non-farm income (log)	4,973	4.597	4.126	0	13.159
Farm income (log)	4,973	4.422	3.939	0	12.269
Death shock	5,000	0.072	0.258	0	1
Illness shock	5,000	0.044	0.204	0	1
Food-price shock	5,000	0.099	0.299	0	1
Drought shock	5,000	0.211	0.408	0	1
Heavy rains/Flood shock	5,000	0.242	0.429	0	1
Age of the household head	5,000	52.33	14.272	18	100
Female household head	5,000	0.182	0.386	0	1
Household head can read and write	5,000	0.675	0.469	0	1
Household size	5,000	6.834	3.497	1	33
Urban household	5,000	0.293	0.455	0	1
Non-agricultural wealth index	4,986	0.044	0.544	-0.647	5.127
TLU large ruminants	4,648	0.015	0.91	-2.195	20.534
TLU small ruminants	5,000	0.42	3.274	0	163.75
TLU other animals	5,000	0.223	0.636	0	14.4
Year of the interview	5,000	0.058	0.779	0	50
		Targeting analysis			
Total income loss & Food Insecurity	316	0.671	0.47	0	1
		Machine learning analysis			
Train sample (pre-COVID-19)					
FIES	3,715	0.145	0.352	0	1
Test sample (post-COVID-19)					
FIES	1,311	0.781	0.414	0	1

Appendix B: Methodological Annex

- *Confusion matrix*

After estimating our models, we gauge their performance on out-of-sample data in the testing set by employing confusion matrices. These matrices enable a comparison between predicted and actual values for our binary outcomes, serving as a widely recognized analytical tool for evaluating the effectiveness of predictive models in classification tasks. In the context of binary classification problems, the confusion matrix is a simple two-way table, as illustrated below:

Table B.1. Example of confusion matrix for a binary classification

		Real status		
		Y = 0	**Y = 1**	
Predicted status	**Y = 0**	True negatives	False negatives	
	Y = 1	False positives	True positives	
	Correctly Predicted	**Specificity**	**Sensitivity**	**Accuracy**

The *True negatives* cell contains those negative cases (Y = 0) that were correctly identified. The *True positives* cell includes the positive cases (Y = 1) correctly identified. The other two cells contain the observations erroneously classified by the model: *False negatives (Type II error)* in which households are predicted to be non-vulnerable to food insecurity before the shock but proved to be food insecure after the shock, and *False positives (Type I error)* in which households are predicted to be vulnerable to food insecurity before the shock but proved to be food secure after the shock. The total accuracy of the predictive model is given by the sum of the *True negatives* and *True positives* cells, divided by the total number of observations.

The *Specificity* of the model, and thus its ability to correctly classify negative cases, is given by the number of observations in the *True negatives* cell divided by the number of negative observations. In the same way, the *Sensitivity* of the model, its ability to correctly predict positive cases only, is the number of units in the *True positives* cells divided by the total number of positive cases. Finally, the *Accuracy* is a general measure of how well the model performs overall, considering both positive and negative predictions. It is calculated as the ratio of correct predictions (true positives + true negatives) to the total number of observations.

Acknowledgments: We are grateful to Luc Christiaensen and Nicholas Sitko for their excellent comments on an earlier version of this work. We also thank Gero Carletto, Andre Croppenstedt, Chris Barrett, Cesar Calvo, Giovanni Andrea Cornia, Stefan Dercon, Donato Romano, and participants at the Sapienza DiSSE 2022 Scientific Meeting, the IFAD 2022 Conference on 'Jobs, Innovation and Value Chains in the age of Climate Change', the 2022 DevEconMeet Workshop, and the 62nd SIE Annual Conference. We also thank Luca Tiberti for the valuable support and the data on restrictive measures. We gratefully acknowledge funding by FAO project "Guidelines to increase the resilience of agricultural supply chains: Getting on the Right Track to Stabilize Production and Markets" – GCP/GLO/588/JPN-F financed by the Ministry of Agriculture, Forestry and Fisheries (MAFF), Japan.