GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

GRK 2654
Sustainable Food Systems

RTG 2654 Sustainable Food Systems

University of Goettingen

SustainableFood Discussion Papers

No. 14

Developing a Multidimensional Resilience Index for Farm Households: A Food System Approach

Margherita Squarcina
Juliane Hänsch
Florena M. Montoya Cepeda
Magdalena Pallauf
Bruno Paz
Jonas Stehl
Jasmin Wehner
Meike Wollni

November 2024

RTG 2654 Sustainable Food Systems · Heinrich Düker Weg 12 · 37073 Göttingen · Germany
www.uni-goettingen.de/sustainablefood

ISSN (2750-1671)

Suggested Citation

Squarcina, M., J. Hänsch, F. M. Montoya Cepeda, M. Pallauf, B. Paz, J. Stehl, J. Wehner, M. Wollni (2024). Developing a Multidimensional Resilience Index for Farm Households: A Food System Approach. SustainableFood Discussion Paper 14, University of Goettingen.

Imprint

SustainableFood Discussion Paper Series (ISSN 2750-1671)

Publisher and distributor:
RTG 2654 Sustainable Food Systems (SustainableFood) – Georg-August University of Göttingen
Heinrich Düker Weg 12, 37073 Göttingen, Germany

An electronic version of the paper may be downloaded from the RTG website:

www.uni-goettingen.de/sustainablefood

Developing a Multidimensional Resilience Index for Farm Households: A Food System Approach

M. Squarcina[a], J. Hänsch[a], F. M. Montoya Cepeda[a], M. Pallauf[a], B. Paz[a], J. Stehl[a], J. Wehner[a], M. Wollni[a]

[a] *Georg-August Universitat Goettingen*

Abstract

Existing measures of resilience focus on specific food system components, neglecting the complexity of the whole system. We propose a measure of resilience that encompasses three dimensions of a food system: economic profitability, environmental sustainability, and adequate nutrition. To empirically estimate the proposed model, we combine longitudinal household-level data from Malawi, Tanzania, and Nigeria with GIS data and macro-level indicators.We define resilience as a normative condition using a probabilistic moment-based approach following Cissé and Barrett (2018). To aggregate the probabilities across different dimensions into a single index of resilience, we employ and compare two different methods. Our findings indicate an overall increase in resilience of farming households over time, with improvements in Nigeria and Tanzania. Clear trade-offs are evident across the various domains of the food system. Both proposed resilience indexes demonstrate strong performance. They are correlated with improvements in income, vegetation, and dietary diversity, and they partially mitigate the effects of various shocks. The comparison between the two methods indicates a preference for the simpler PCA-based approach to measuring farmers' resilience using a food system approach. Our findings underline the need to broaden our focus beyond individual aspects of resilience to achieve sustainable food systems.

Keywords: Resilience measurement, Shocks, Agriculture, Malawi, Nigeria, Tanzania

JEL: Q12, C43, O13, Q54, Q56

1. Introduction

Food systems repeatedly face political and economic instabilities, environmental challenges, and health crises. These problems are further exacerbated by climate change, making it essential to develop effective strategies for prevention and recovery from such disruptions. In this context, resilience has emerged as a key concept among development and humanitarian organizations, as it offers a multidimensional and interdisciplinary approach to addressing the complexities of risk exposure.

However, until now, resilience has primarily been conceptualized to measure only standards of living, such as poverty and food security. Barrett and Constas (2014) developed a resilience theory specifically focused on poverty, which can be linked to the concept of the poverty trap. Similarly, the Food and Agriculture Organization of the United Nations (FAO) has developed a resilience capacity index to measure the capacity to ensure a certain level of food security in the aftermath of stressors or shocks (FAO, 2016a). However, an emerging body of scholars recognizes the importance of developing and adopting a more holistic and integrative approach to resilience centered on food systems (Béné and Devereux, 2023).

A holistic approach to measure resilience is necessary to simultaneously capture all aspects of a food system and potential trade-offs between them. Neglecting environmental sustainability while focusing merely on poverty reduction can lead to successful economic outcomes but harm the environment. Similarly, maintaining adequate levels of nutrition and food security during crises that affect farmers production requires economic affordability from other income sources.

Although many studies have looked at selected components of food systems, such as agricultural production or other stages in the food value chain (Darnhofer et al., 2010, Milestad et al., 2010, van Apeldoorn et al., 2011, Soane et al., 2012) or selected food security indicators (Pingali et al., 2005), they did not consider their interactions comprehensively. The literature on food system resilience is still relatively scarce (Bizikova et al., 2016, Meyer, 2020, Zurek et al., 2022), and rigorous and reliable methods to measure food system resilience are still missing (Béné et al., 2023). There is therefore a particular need to improve resilience measurement in an integrative and comprehensive way, applying a food system approach that can be empirically measured (Upton et al., 2022).

This study aims to address these literature gaps by proposing an empirical measure of resilience encompassing three dimensions of a food system: economic profitability, environmental sustainability, and adequate nutrition. We compute different resilience indexes by building on previous methods to measure resilience and employing different techniques, including structural equation modeling and principal component analysis, to address the multidimensionality of the resilience measure.

We adopt the definition of resilience as a normative condition (Barrett and Constas, 2014) and we consider the predicted probability that each dimension of the food system is above a defined threshold as the outcome variable, following the methodology that Cissé and Barrett (2018) proposed.

To aggregate the three dimensions into a unique index, we follow two main approaches: a more complex and data-demanding one, which uses structural equation modeling and includes factors selected by the Least Absolute Shrinkage and Selection Operator (LASSO) regression as latent correlates of resilience; and a more simple one, which makes use of principal component analysis to aggregate the three probabilities into a unique indicator.

The two related resilience indexes are empirically measured on a cross-country dataset that combines longitudinal household-level data from Malawi, Nigeria, and Tanzania, with GIS and macro-level indicators. The two resulting indexes are compared and tested over different shocks.

The findings show that over time, farmers in Tanzania and Nigeria, on average, have seen an increase in their resilience levels. Malawi instead shows the lowest and declining levels of resilience during the study period. Clear trade-offs across outcomes exist. Specifically, both resilience indexes are positively linked to economic and environmental outcomes, but negatively associated with the nutritional outcome. When faced with shocks like conflicts or droughts, farmers' level to adapt and recover is positively associated with each dimension of the food system, and partially mitigates the negative effects of the shock. The two indexes are highly correlated and provide similar results. On this basis, preference is given to the simpler PCA-based approach to measuring farmers' food system resilience.

The paper is organized as follows. The next section provides a review of the existing literature on the concept of food system resilience and the related measurements. Section 3 presents the proposed models for measuring food system resilience. Section 4 describes the data and presents some descriptive statistics. In Section 5, the resilience indexes are empirically measured and tested over two shocks. Section 6 concludes.

2. Literature review

Among the first to develop a framework linking resilience and food systems was Tendall et al. (2015), who proposed a theoretical framework that defines the concept of food system resilience, emphasizing the importance of considering multiple levels of the food system and explicitly adopting a whole system perspective. They tie resilience to a functional goal, as suggested by Barrett and Constas (2014), of "ensuring sufficient, appropriate, and accessible food to all, while excluding potentially undesirable outcomes like food and nutrition insecurity or environmental degradation" (Tendall et al., 2015). This framework introduces a novel aspect by linking resilience performance to other social and environmental outcomes.

Similarly, Bizikova et al. (2016) proposed a conceptual framework for assessing the resilience of food systems in the face of climate challenges. The authors emphasise the significant role that supporting systems, institutions and processes play in determining the resilience of food systems. In particular, they identify resource management, essential infrastructure, and policies that promote inclusive decision-making as key factors to increase resilience.

Béné et al. (2023) presented a novel framework for evaluating the resilience of local food systems. This framework distinguishes itself from previous ones by taking into account the interplay of actors and resilience processes at the system level. The framework highlights the importance of considering the mix of actors involved in a food system, thus recognizing that the functioning of the system does not depend merely on individuals and communities, but also on the institutions and their combined actions. Three core functions are considered in the definition of a food system: the achievement of food and nutrition security, the generation of sufficient income, and the protection of the environment. Thus, in the paper, the authors emphasize food security while also highlighting the importance of simultaneously considering livelihoods and recognizing the environmental dimension of food systems.

These studies introduced various conceptualizations and frameworks for food systems resilience in the literature, but did not propose an empirically measurable index. In contrast, this has been done for the concept of development resilience, for which three different approaches exist. As described by Barrett et al. (2021), Barrett and Constas (2014), resilience can be measured as a capacity, as a normative condition, and as a return to equilibrium. The first approach defines resilience as an ex-ante capacity that reduces exposure to shocks, thereby avoiding long-lasting adverse consequences (Constas et al., 2014a). The Resilience Measurement Technical Working Group introduced this definition in 2013, and it has been mainly used by FAO and other UN-affiliated organizations.

Two main indexes of resilience capacity were developed. The first is the Resilience Index Measurement and Analysis (RIMA-II), developed by FAO (2016a). It measures resilience towards food security and is computed using a Multiple Indicator Multiple Causes (MIMIC) model that defines the index from four latent variables, called pillars - Access to Basic Services, Assets, Social Safety Nets, and Adaptive Capacity – and simultaneously from food security indicators, namely food expenditure and dietary diversity. Another Resilience Capacity Index was developed by TANGO International (Smith and Frankenberger, 2018). In this case, resilience is conceptualized as a latent variable that reflects absorptive, adaptive, and transformative capacities. This index is mainly used by the United States Agency for International Development (Henly-Shepard and Sagara, 2018).

The second approach anchors resilience to normative well-being standards, such as a poverty line. This approach, conceptualized by Barrett and Constas (2014) and translated into an econometrically testable measure by Cissé and Barrett (2018), estimates resilience as the probability of achieving at least a minimal standard of living based on observable characteristics and exposure to stressors and shocks (Barrett et al., 2021). The probability measure is estimated via two multivariate regressions, and is based on an assumed two-parameter distribution.

The third approach focuses on the capacity of households to recover from shocks (Constas et al., 2014a,b, Hoddinott, 2014, Knippenberg et al., 2019). It is conceptually similar to the normative condition approach, but without normative anchoring. Based on this approach, households returning to an ex-ante state would be defined as resilient.

All of these measures share a common factor: they focus on a single outcome at a time, neglecting the multidimensional nature of a system.

In this paper, we aim to adapt the existing approaches to measure development resilience to a food system level, translating the proposed conceptual frameworks into a measurable index that accounts for the multidimensional nature of the food systems.

3. Model Conceptualization

The aim of this study is to examine the resilience of various dimensions of food systems simultaneously. Specifically, we consider three dimensions: economics, environment, and nutrition. We adopt the definition of resilience as a normative condition (Barrett and Constas, 2014). According to this definition, a household is considered resilient if it meets a minimum standard in the specific outcome being considered. We then adopt the Cissé and Barrett (2018) approach to measure the conditional probability of being above a certain threshold in each dimension. The predicted probabilities are estimated separately for each dimension and then combined into a single index. For the environmental outcome, we incorporate the concept of sustainability by applying a return to equilibrium definition of resilience, considering the probability that at time t, the outcome is at least at the same level as at time $t-1$.

Our analysis focuses on farmers and their households as the unit of analysis within the agri-food system. Farmers play a critical role in contributing to different outcomes of a food system, e.g. related to environmental sustainability or food security, as they have control over natural resources. However, as highlighted by Maleksaeidi and Karami (2013), it is important to recognize that farm households are integrated into larger systems, including communities and social networks, which can influence their decision-making processes and resilience capacity. The proposed model relies on the framework developed by Béné et al. (2023), where the resilience of food systems builds in processes at both individual and system levels and considers different levels, namely household, plot, and community.

This approach acknowledges that a household's resilience capacity depends not only on its own characteristics but also on the social and economic environment surrounding it. External factors beyond the control of individuals, as well as geographical and environmental variables, impact an individual's ability to cope and recover from shocks, particularly in the context of climate-related shocks.

To address the multidimensionality of resilience at the food system level, we took advantage of some components of the existing approaches currently used to measure development resilience. The probabilistic approach by Cissé and Barrett (2018) alone is not suitable for the purpose of this study, as it does not allow for the measurement of several outcomes simultaneously. In the study by Lee et al. (2024) the multidimensionality of development resilience is computed by estimating the probability of being simultaneously above the threshold in each dimension. However, this method does not consider trade-offs between dimensions. While this approach might be appropriate for measuring development resilience, as the dimensions are expected to be all positively correlated, it is less realistic when examining different dimensions of a food system. The RIMA approach by FAO uses structural equation modeling to combine different indicators into a latent variable of resilience. In this way, different observable variables are used as proxies to define the unobservable variable of resilience. This approach allows trade-offs across the endogeneous indicators. Specifically, the RIMA's framework employs a Multiple Indicators Multiple Causes (MIMIC) model. This model is an extension of confirmatory factor analysis with covariates (Chang et al., 2020). The model comprises two parts: the measurement model and the structural model. The measurement model specifies how the latent variable (of resilience) is measured by the endogenous indicators, while the structural model groups observable variables into correlates (causes) of the latent variable. In our case, the conditional probability of each of the three outcomes to be above a certain threshold can be included as the endogenous indicator in the measurement model.

The correlates of the structural model instead are grouped into three scale components: household, plot, and community. Factor analysis is employed to linearly combine the observable variables into the correlates. Factor analysis is a variable reduction mechanism that finds cross-correlations among observed variables, identifies the number of unobservable factors emerging from these correlations, and predicts the latent variable as a linear combination of these factors (FAO, 2016a). Only factors with an eigenvalue greater than 1 have been considered.

The variables populating the factor analysis for defining each component are selected using LASSO regression, following Knippenberg et al. (2019) and Kshirsagar et al. (2017). LASSO is an approach intended to perform a subset selection of variables while maximizing the fit of the model to obtain a good out-of-sample performance. The key aspect is that it drops highly correlated variables. Their variation is captured by similar variables (which are similarly useful for prediction). The LASSO regression is employed as it selects variables without arbitrariness and it takes into consideration the specific context and type of shock analyzed. In fact, we would expect different factors to be more relevant depending on the shock experienced by farmers and the specific context analyzed.

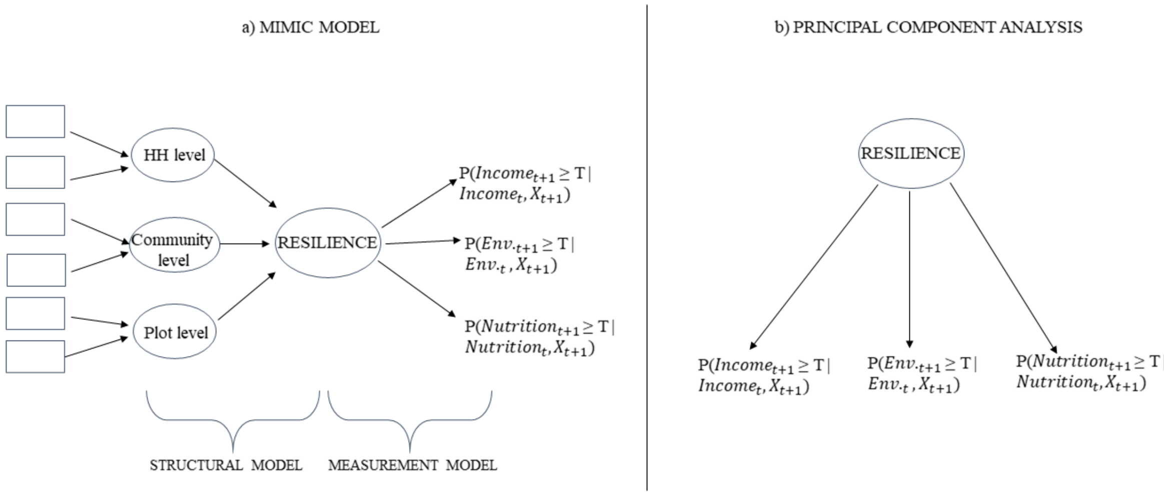
However, this model is complex and computationally demanding. It requires incorporating correlates into the structural part of the model, which can lead to arbitrary decisions in variable selection and issues with interpretability. Additionally, the model may not converge or exhibit poor fit.

To address these challenges, we also consider a simpler approach that utilizes principal component analysis (PCA) to combine the three probabilities into a single index. PCA is a dimensionality reduction technique that extracts linear composites from observed variables. It is often used to consolidate correlated variables into a single index.

Since PCA produces a linear combination of probabilities, no further steps or additional data aggregation are necessary. Moreover, this approach does not require extra variables related to resilience determinants and still allows for identifying the contribution of each probability and assessing the trade-offs among them. The visual representation of the two approaches is shown in Figure 1.

Comparing the results and performance of the resilience indexes derived from the two approaches will help us identify which method is more effective for calculating a multidimensional index of food system resilience.

Figure 1: Graphical representation of the models.



Source: authors' elaboration.

4. Data

4.1. Data source

Different data sources have been combined, merging household survey data with geographic information on soil, land use, and infrastructure. Conducting a cross-country analysis enables testing the external validity of the framework, while employing different types of shocks allows assessing how the model can predict resilience across various shock scenarios.

The criteria considered to select the data were the following: longitudinal surveys at household level with information on socio-economic characteristics of household members as well as plot information, covering at least three rounds of panel data. Data should also include household coordinates to be merged with geo-referenced data, and should have a harmonized set of variables to be comparable across countries. We identified three countries that satisfy all the criteria: Tanzania, Malawi, and Nigeria. The two main data sources used are the LSMS-ISA survey by the World Bank, and the RuLIS [1] indicators by FAO, which are based on the same surveys.

[1]For more information about the RuLIS project, please visit https://www.fao.org/in-action/rural-livelihoods-dataset-rulis/en/

The LSMS-ISA for these three countries include modified household coordinates[2] and have panel data of three survey rounds. The LSMS-ISA data provides also geographic data based on the actual location of the plot. As the focus of the analysis is on the food system, we only considered farm households. The definition we applied to select those households relies on the one used in RuLIS, which includes any household engaged in crop production in the last 12 months. [3] Based on this, we obtained a balanced sample of 2,949 farm households per year, spanning three time periods in each of the three countries (Table 1).

Table 1: Dataset description

Country	*Survey*	*Year*	*Panel sample size*
Malawi	Integrated Panel Household Survey	2010, 2013, 2016	949
Nigeria	General Household Survey	2012, 2015, 2018	462
Tanzania	National Panel Survey	2008, 2010, 2012	1538

Source: authors' elaboration.

Other data included in the final dataset are taken from The Armed Conflict Location and Event Data Project (ACLED) dataset, the Afrobarometer data, gridded SPEI (Standardised Precipitation - Evapotranspiration Index) data from the Spanish National Research Council (CSIC), and the Normalized Difference Vegetation Index (NDVI) derived from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery. The Armed Conflict Location and Event Data Project data collects information on the dates, actors, locations, fatalities, and types of all reported political violence and protest events around the world. The Afrobarometer provides data on democracy, governance, and quality of life from nationally representative surveys in 38 African countries. Surveys are based on a standard questionnaire that allows comparisons across countries and over time. Survey rounds are collected every two to three years. The geocoded dataset allows the aggregation of data at a low administrative level (BenYishay et al., 2017). In this analysis, we aggregated the data at the administrative level 2. See Appendix 7.4 for more details.

The SPEI is a multiscalar drought index based on monthly precipitation and potential evapotranspiration data from the Climatic Research Unit of the University

[2]For the methodology used to modify the coordinates please refer to the technical documentation of each survey.

[3]The definition provided by RuLIS is as follows: "Dummy variable indicating whether or not the household was engaged in crop production in the last 12 months. The households that have produced any crop and exhibit any related income/ expenditure are considered as being engaged in crop production."

of East Anglia (from January 1901). This SPEI dataset is based on the FAO-56 Penman-Monteith estimation of potential evapotranspiration, which is considered a superior method and recommended for most uses, including long-term climatological analysis. Thus, the SPEI offers long-time and robust information about different drought conditions, depending on the time-scale under analysis, with respect to normal conditions. Data are available from CSIC with a 0.5 degrees spatial resolution and a monthly time resolution (between 1 and 48 months). The NDVI is a measure of vegetation greenness captured in a satellite image and is used to examine vegetation density and changes in plant health.

A higher NDVI value indicates higher vegetation greenness, i.e. healthy and dense vegetation. MODIS provides images with a 500 meter spatial and 16-day temporal resolution. The vegetation index products provided by MODIS MOD13A1 V6.1 are composited images and corrected for e.g. clouds. For the purposes of higher precision and storage, the MODIS NDVI is rescaled to values ranging from -2.000 to +10.000 (healthy greenness is indicated by values from 3.000 to 10.000).

4.2. Outcome variables

Economic outcome. We use the income variable provided by RuLIS, which corresponds to the daily income per capita in the last 12 months[4]. The value was originally expressed in local currency units on an annual basis, and converted in USD in 2017 Purchasing Power Parities (PPPs). The conversion in USD makes it comparable with the international poverty line defined by the World Bank, which is set at USD 2.15 per person per day using 2017 prices. We used this poverty line as the threshold for our analysis, considering all households below it as extreme poor. Although an international poverty line does not take into account country-specific factors, it is preferable for cross-country analysis for reasons of comparability.

Nutritional outcome. We use the dietary diversity score (DDS) as a measure of nutritional quality of the diet. The DDS measures the number of different food groups consumed, with a higher number corresponding to greater diversity[5]. The indicator considers food groups instead of food items to ensure the measurement of diversity in both macro- and micronutrients. Between 59 food items (Tanzania) and 141 (Malawi 2016) are assigned to 10 food groups, namely 1) grains, white roots and tubers, and plantains, 2) pulses (beans, peas and lentils), 3) nuts and seeds, 4) dairy, 5) meat,

[4]Specifically, it is "the quotient of total income, 365 days and the household size" (FAO, 2022)

[5]DDS usually measures food consumed the previous day or night. However, in this study, we consider a recall period of one week to fit the data.

poultry, and fish, 6) eggs, 7) dark green leafy vegetables, 8) other vitamin A-rich fruits and vegetables, 9) other vegetables, and 10) other fruits.

Based on the DDS, we computed the Minimum Dietary Diversity for Women (MDD-W), which indicates the proportion of the population that consumed at least five out of the ten defined food groups. The MDD-W is an indicator developed by FAO to measure the dietary diversity of women in reproductive age (FAO, 2016b). Although it has been validated as a proxy of micronutrient adequacy specifically in the diet of this population group, since women are more vulnerable in terms of micronutrient deficiencies, and the nutritional requirements for women during pregnancy and lactation are higher than those of adult men, the indicator can be extended to the entire population. Several studies indeed used MDD-W to measure dietary diversity in other population groups (Heim and Paksi, 2019, Zhang et al., 2020, Mridha et al., 2020, Gómez et al., 2024).

Environmental outcome. The environmental outcome aims to ensure that the farmers' activities do not negatively impact the surrounding landscape. To achieve this, we adopt a landscape approach to capture the spillover effects of farming on the surrounding area. We use the Normalized Difference Vegetation Index (NDVI) as a proxy for environmental sustainability. While other indicators, such as forest cover, are commonly used to measure environmental sustainability (Moldan et al., 2012), our focus is on understanding the farmers' influence on their cropland and adjacent areas. Moreover, our sample includes farmers from landscapes that are not naturally covered by tropical forests, e.g. the Sahel region in Nigeria. For these reasons, tree or forest cover is not the most suitable indicator for this context. Instead, analyzing overall vegetation levels through NDVI provides a better understanding of the farmer's environmental impact.

We expect that farmers practicing sustainable methods will maintain or increase vegetation indices levels on their plots by e.g. avoiding deforestation or applying practices which positively influence soil and vegetation health. NDVI is an appropriate measure as it reflects vegetation health and can be used for land cover classification and monitoring forest cover and change (Sader and Winne, 1992, Defries and Townshend, 1994, Jia et al., 2014, Liang et al., 2018). It has been used in previous studies to measure changes in environmental quality (Fung and Siu, 2000, Pettorelli et al., 2005).

We calculated the annual mean NDVI as raster images for each country. The modification of the household coordinates in the LSMS-ISA surveys is performed within a radius of 2 kilometers for urban households and 5 kilometers for rural households.

Thus, we used the same approximation to calculate the average annual NDVI in buffer zones of 2 kilometers for urban and 5 kilometers for rural households around the modified coordinates. As we are interested in measuring the change over time, we computed the percentage change of average annual mean NDVI from t to the 5-year average value prior to the first round of the survey in each country. The 5-year average is intended to control for possible extreme values. We then defined 0 as the threshold, meaning that the level of NDVI at time t should be at least equal to the baseline 5-year average.

4.3. Shocks

We consider two types of shock: weather shocks and conflict-related shocks. Thus, we can test the model over different shock typologies.

For the weather shock, we used the Standardised Precipitation - Evapotranspiration Index (SPEI) as a measure of drought. The SPEI has been increasingly used in climate-related studies, as it extends the Standardized Precipitation Index (SPI) by taking into account precipitation as well as potential evapotranspiration. In this way, it captures the effect of temperature on water availability. Our study uses the 3-month time-scale of the main growing season in each country[6] to capture short-term drought effects (i.e. agricultural drought). As for the NDVI, we computed the average SPEI value in buffer zones of 2 and 5 kilometers in urban and rural areas, respectively, from the modified household location. The SPEI values range from -5 to +5. Smaller values indicate stronger degrees of drought, and larger values indicate higher degrees of moisture. Based on Li et al. (2015), Mckee et al. (1993), Paulo et al. (2012), we considered -1.5 as a threshold for drought, which includes measures of severely and extremely dry. The three countries analyzed are particularly exposed to drought, with extreme peaks in some years, as reported in Table 2. Nigeria experienced drought events in 2015 (Eze et al., 2020, Orimoloye et al., 2021, Durowoju et al., 2021), while Tanzania was affected by a major drought in 2010/2011 (Gebremeskel Haile et al., 2019, Mwangi et al., 2014). In Malawi instead, the El Nino-induced prolonged drought occurred at the same time of short periods of repeated floods (Henriksson et al., 2021). For this reason, the effect of this drought is not captured using a 3-month time frame, but it becomes explicit when using annual values. Although SPEI and NDVI could be seen as proxies of the same phenomenon, the two variables capture different aspects. The NDVI indeed is the result of different factors, including exogenous factors, such as the level of precipitation and temperature, and endogenous ones, such as deforestation and agricultural practices. Therefore, the level of precipitation measured through the SPEI is one of the factors

[6]A detailed description is provided in Appendix 7.4

affecting the level of greenness. Additionally, although the two variables are correlated, the use of different time frames (annual NDVI vs 3-month SPEI) allows us to differentiate the two variables. When using annual values of SPEI, the two variables are positively correlated at 1% significance level. However, when considering the 3-month SPEI, the two variables are negatively correlated. Additionally, the NDVI variable is not correlated with the dummy of drought.

For conflict-related shocks, we used the ACLED data to compute a dummy equal to one if any incident occurred in time t. Following Thiede et al. (2020) and Eseosa Ekhator-Mobayode et al. (2022), we used a 10 kilometer buffer zone around the households' modified geographic location to match the event location to the household location. As shown in Table 2, Nigeria is the country most affected, as it was exposed to Boko Haram insurgency since 2009, with 2,378 conflict events by this terrorist group occurring across the country between 2009 and 2017 (Eseosa Ekhator-Mobayode et al., 2022).

Table 2: Percentages of households affected by shocks

Country	Drought shock	Conflict shock
Malawi 2013	0.00%	7.90%
Malawi 2016	0.00%	13.28%
Nigeria 2015	46.32%	20.13%
Nigeria 2018	0.22%	28.14%
Tanzania 2010	9.88%	1.11%
Tanzania 2012	31.51%	7.28%
Total	14.43%	9.38%

Source: authors' elaboration.

4.4. Descriptive Statistics

Table 3 presents the mean values of the main variables used in the analysis, which include the three outcomes and the covariates used for the measurement of the conditional probabilities. The table reports the mean over the pooled sample, and by time period[7]. Generally, we do not observe major changes across rounds. Per capita daily income is on average 1.08 USD, below the international poverty line. DDS remains quite constant over time, with low variation on average, at between 5-6 food groups consumed within a week. The NDVI instead shows a slight decrease

[7] The number of each time period corresponds to the round of data in each country. However, it corresponds to different years among countries.

over time. 20 percent of households are female headed, with three female members in the households on average.

Table 3: Descriptive statistics

Variable	Total	t=1	t=2	t=3
Outcome Variables				
Per capita HH income (2017 USD PPP)	1.08	1.07	1.05	1.13
DDS	5.60	5.59	5.85	5.71
Annual mean NDVI	4789	4820	4778	4769
Shock Variables				
Drought	0.10	0.02	0.12	0.16
Conflict	0.07	0.03	0.06	0.12
Control Variables				
HH head is female	0.21	0.20	0.21	0.22
Age of HH head	48.12	47.18	48.73	48.44
HH head is married	0.75	0.79	0.75	0.71
Educ. of HH head	5.35	5.15	5.23	5.69
HH size	6.25	6.11	6.44	6.21
HH receives food for free	0.49	0.44	0.52	0.52
Ag. Employment	0.31	0.23	0.32	0.38
Land owned (hectares)	1.30	1.25	1.30	1.36
Use of chemicals	0.20	0.17	0.23	0.20
Use of mechanized equipment	0.19	0.16	0.19	0.23
Demographic dependency ratio	1.18	1.22	1.19	1.13
N. of of females in the HH	3.16	3.17	3.23	3.07
HH has electricity	0.09	0.08	0.09	0.10
HH owns house	0.91	0.93	0.92	0.89
HH engages in livestock prod.	0.70	0.75	0.74	0.62
Herfindahl index of income	0.66	0.68	0.67	0.65
Distance to main road (km)	14.30	14.55	14.43	13.93

Source: authors' elaboration. Selection of control variables.

Some positive trends however emerge. First, education of the household head shows a slight increase on average over time. Having at least one HH member employed in agriculture also increased from $t = 1$ to $t = 3$, moving from 23 percent to 38 percent, respectively. At the same time, farmers were also able to own more land on average. Some trends reflect the agricultural and rural transformations that occurred in Sub-Saharan Africa in the last decades (IFAD, 2016, Barrett et al., 2017). For instance, the use of mechanization in agriculture increased from 16 percent in time $t = 1$ to 23 percent in time $t = 3$. As the transformation involved the development of infrastructure, we can see that access to electricity has increased, as well as access to the main roads. The Herfindahl index of income, which measures the degree of concentration of different income sources in the total income of the household, reports a lower level of income diversification over time. This suggests that households tended to specialize into one income source. The increase in land owned and the reduction of households engaged in livestock production suggest that farmers in the three countries tended to specialize in crop production. This sample is composed of households that have some crop production over time and excludes those households that stopped doing agriculture, even temporarily, or those that migrated to urban areas.

4.5. Balancing weights

The original panel sample is composed of households tracked over the three rounds of data, and it is representative at the national as well as urban/rural levels. However, since we are considering only households engaged in agricultural production across all three rounds in each country, our final balanced sample of farmers might differ from the original sample. As shown in Table 4, the subsample of farmers continuously practicing agriculture over the entire period analyzed differs significantly in several dimensions from those households that intermittently engaged in agriculture (Column 1). To account for this, we computed inverse probability weights based on the propensity score of being included in the subsample. This score was derived from a logit regression with the dummy variable indicating inclusion in the selected sample as the dependent variable and a set of household characteristics as regressors. Details of the model are provided in Appendix 7.1.

When applying the weights, the two groups of households become balanced, as shown in Table 4. These computed weights were then used in the subsequent analysis of the paper. However, as we are focusing on farming households, we acknowledge that our sample is not representative of the original panel sample, as it was not designed based on this characteristic. Therefore, we do not claim representativeness for our results, as this is not the aim of the study.

Table 4: Comparison of baseline characteristics with and without weights.

	No weights			IPW		
	Other HHs	Subsample	Mean diff.	Other HHs	Subsample	Mean diff.
HH head is female	0.21	0.20		0.20	0.20	
Age of HH head	49.33	47.18	***	48.17	47.76	
HH head is married	0.77	0.79		0.78	0.79	
Educ. of HH head	5.88	5.15	***	5.32	5.38	
HH size	6.36	6.11	**	6.23	6.19	
Free food	0.37	0.44	***	0.43	0.42	
Ag. Employment	0.11	0.23	***	0.20	0.20	
Land owned	0.96	1.25	***	1.28	1.19	*
Use of chemicals	0.17	0.17		0.19	0.17	
Ag. Machinery	0.14	0.16		0.15	0.15	
N. of female HH members	3.26	3.17		3.20	3.20	
Electricity	0.21	0.08	***	0.12	0.12	
House is owned	0.84	0.93	***	0.90	0.90	
Livestock prod.	0.64	0.75	***	0.73	0.72	
Distance to main road	17.13	14.55	***	14.86	15.15	
Income diversification	0.73	0.68	***	0.69	0.69	
Urban	0.21	0.08	***	0.11	0.11	

Source: authors' elaboration.
Note: IPW stands for Inverse Probability Weights. Weights generated from a logit model with robust standard errors. Mean difference is computed from a linear regression. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

5. Empirical analysis

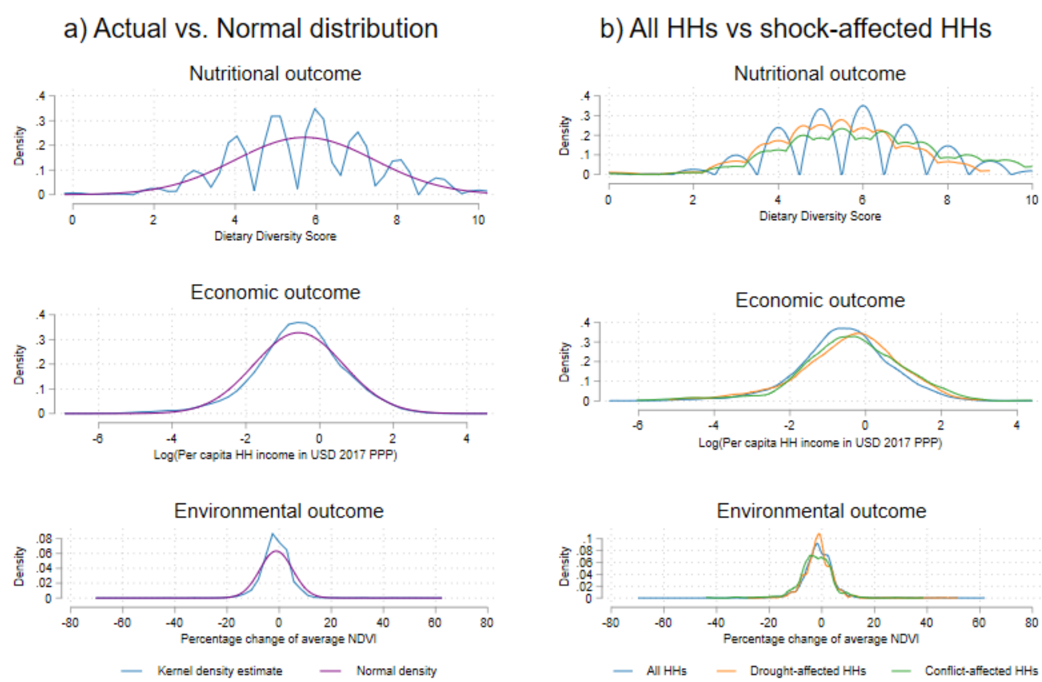
To compute the final resilience indexes, we initially employed the Cissé and Barrett (2018) approach to determine the probabilities that serve as outcomes in the model. Subsequently, LASSO regression was utilized to identify the variables that directly affect the three outcome variables. The same vector of variables is used to define the model's components in the MIMIC model. The latest were then computed through factor analysis. Finally, PCA was used to aggregate the three measures of probability into a unique index, while the MIMIC model was employed to predict the latent variable of resilience in the second approach. We then compared the two indexes of resilience and checked whether different factors are associated with the level of resilience, depending on the approach used. For assessing the performance of the two indexes over various shocks, we ran a regression model with household and time fixed effects. This section presents the results of each phase of the analysis.

5.1. Probability outcomes

To estimate the predicted probabilities of achieving a certain level of environmental, dietary, and economic outcome, we follow the approach proposed by Cissé and Barrett (2018). Since the probability is modeled based on the distribution of the outcome variable, we first need to check and identify the type of distribution for each variable.

Figure 2 provides the Kernel density of the distribution of the three variables used to compute the conditional probability. For the dietary and environmental outcomes, we use the levels of the variables. We transformed the economic variable in logarithm terms.

Figure 2: Distribution of outcome variables



Source: authors' elaboration.

Note: Kernel density of the distribution of the outcome variables. Panel (b) compares the distribution of the entire sample with the distribution of the subsample of households affected by each type of shock analyzed.

Graphs in *Panel a* compare the actual with a normal distribution. The density functions suggest that all three variables follow a normal or lognormal distribution in levels[8].

Panel b of Figure 2 plots the distribution of the overall sample with the distributions of the sub-samples affected by the two shocks to unveil possible differences across them. The two samples of households report a similar distribution in each of the outcomes considered. This graphical test suggests that there are no notable differences in the outcome variables between households affected by a shock and those unaffected[9].

The conditional mean and variance of the outcome of interest are then estimated using a normal distribution. This is accomplished by running an OLS regression, assuming the outcome follows a normal distribution, to estimate the expected outcome at time t defined as a polynomial function of lagged outcome and a set of control variables. After testing for different polynomial specifications[10], only the first-order specification was included in the linear regression. The decision was made by checking the Akaike information criterion (AIC) values and the t-test on the equality of means between the predicted values of the higher-order specifications, following Cissé and Barrett (2018). The t-test shows statistically insignificant differences for everything above the first order. Based on the distributional assumptions and considering the lagged value of order one, we computed the probabilistic measures, using the probabilistic moment-based approach following Cissé and Barrett (2018). Full results of the regressions are reported in Appendix 7.1.

The equation for the model is expressed as follows:

$$Y_{ht} = \hat{\beta}_M Y_{h,t-1} + X_{ht} + \epsilon_{Mht} \tag{1}$$

where Y_{ht} is the outcome variable of the food system (either income, NDVI, or DDS); $Y_{h,t-1}$ is the same variable in the previous survey round; X_{ht} is a set of control variables, including household characteristics, such as household size, level of education of the household head, land owned, access to electricity, dependency ratio, as well as distance to the main road and agroecological zones; and ϵ_{Mht} is the error

[8]The DDS, used as a proxy for the nutritional outcome, has non-negative values, therefore a Poisson distribution could fit better.

[9]Although there is a correlation between the continuous variable of SPEI and NDVI, no significant correlation between the dummy variable for drought and NDVI is observed.

[10]The test was conducted only for income and DDS, since for NDVI the lagged value is not included among the regressors. Tables 12 and 13 in Appendix 7.2 show the results of the test for income and DDS, respectively.

term.

We then predict the residuals from the previous equation and square them to obtain the variance equation. To estimate the conditional variance, we employ a similar model to the previous one, but with the variance as the dependent variable. This model is also estimated via maximum likelihood.

$$\sigma^2_{ht} = \hat{\beta}_V Y_{h,t-1} + X_{ht} + \epsilon_{Vht} \tag{2}$$

Based on the estimated conditional mean and variance, we calculate the probability density function and the complementary cumulative probability beyond the threshold value. The complementary cumulative probability corresponds to the resilience score of household i at time t. Finally, we regress these resilience scores, specific to households and time periods, on the same set of regressors used in the mean and variance equations. Results tables for the estimation of the conditional mean, conditional variance, and conditional probability are reported in Appendix 7.1.

Table 5 presents the conditional probability of each outcome by country and year. Among the three countries analyzed, Malawi exhibits the poorest performance over time, showing a decreased probability of all outcomes from 2013 to 2016. The environmental dimension, in particular, reflects the country's weakest performance. Contrary, despite a decline in 2016, the probability of achieving the nutritional threshold remains the highest in Malawi compared to the other countries. In contrast, Nigeria experienced a decline in the probability of achieving economic outcomes but showed improvements in the other two dimensions, with a significant increase in the probability of the environmental outcome over the years. Lastly, Tanzania saw a slight decrease in the probability of the nutritional outcome, but an increase in both economic and environmental outcomes.

Table 5: Probability outcomes by country and year.

	Prob(Economic)	Prob(Nutrition)	Prob(Environment)
Malawi 2013	0.155	0.878	0.185
Malawi 2016	0.112	0.812	0.096
Nigeria 2015	0.198	0.540	0.121
Nigeria 2018	0.169	0.630	0.558
Tanzania 2010	0.177	0.636	0.525
Tanzania 2012	0.207	0.605	0.557
Total	0.172	0.686	0.382

Source: authors' elaboration.
Note: Prob(X) indicates the conditional probability of the outcome above the threshold. Inverse probability weights applied.

5.2. Approach 1: Principal Component Analysis

PCA is an unsupervised machine learning technique widely used for dimensionality reduction in large datasets, ensuring minimal loss of information. It transforms the original correlated variables into a new set of uncorrelated variables known as principal components, which are linear combinations of the original variables. These principal components are ordered such that the first few retain most of the variation present in the original dataset. In this way, PCA reduces the dataset's dimensionality while preserving as much variability as possible. PCA involves several key steps. First, the data is centered by subtracting the mean of each variable. Next, the covariance matrix is computed to understand how variables vary together. Eigenvalues and eigenvectors of this covariance matrix are then calculated. The eigenvectors (principal components) corresponding to the largest eigenvalues capture the greatest variance in the data. The loadings in PCA represent the correlation between the original variables and the principal components.

In our data, we first verified that all three probability variables are correlated with each other. This is true at the 0.001 significance level. We then computed the PCA and found that the first principal component explains 51 percent of the total variance. As it is the only component with an eigenvalue greater than 1, we discarded the other two components. The loadings indicate a positive correlation between the principal component and the environmental and economic probabilities (0.65 and 0.47, respectively), and a negative correlation with the nutritional probability (-0.60).

5.3. Approach 2: Multiple Indicators Multiple Causes Model

To select the variables to consider for the scale components, we conducted a three-step procedure. First, we manually selected those variables that can be considered predictors of resilience towards the three outcome variables. The selection was conducted based on literature, and in particular, considering the variables used to define the pillars in the RIMA-II model[11]. In this way, we include in the model only those variables that can directly affect the level of resilience of each farmer household. However, since the list of variables was still extensive, we ran a LASSO regression to avoid any subjective imputation regarding which factor should be considered in the model. The second step therefore is the LASSO regression for each outcome separately. The variables selected in each regression are first grouped together and subsequently split into three levels (household, community, and plot). LASSO se-

[11] The full list of variables selected and the comparison with the RIMA-II pillars is reported in Appendix 7.5.

lected 76 percent of the initial list of variables (72 out of 95 variables)[12] Afterwards, we ran the factor analysis to compute the components. The components will be used as the "causes" in the MIMIC model to ensure proper model specification. This step is needed because MIMIC can only handle a limited number of variables in its structural part. Including all observable variables in the model would prevent it from converging. Variables have been standardized before the computation of the factor analysis. Only factors with an eigenvalue greater than 1 have been considered.

Figure 3 shows the contribution of each observed variable to the latent variable that defines the component. Specifically, it reports the factor loadings [13] of the components as well as the correlations between the observed variables and the latent variable.

The component at the HH level reports the highest number of variables as compared to the other two scales. Variables considered include household characteristics, as well as variables linked to agricultural production and income composition. The level of education within the household is the most relevant factor at the household level. Indeed, all the different variables of education report the highest factor loadings and are highly correlated with the latent variable. Receiving social assistance is also highly correlated with the HH-level component variable.

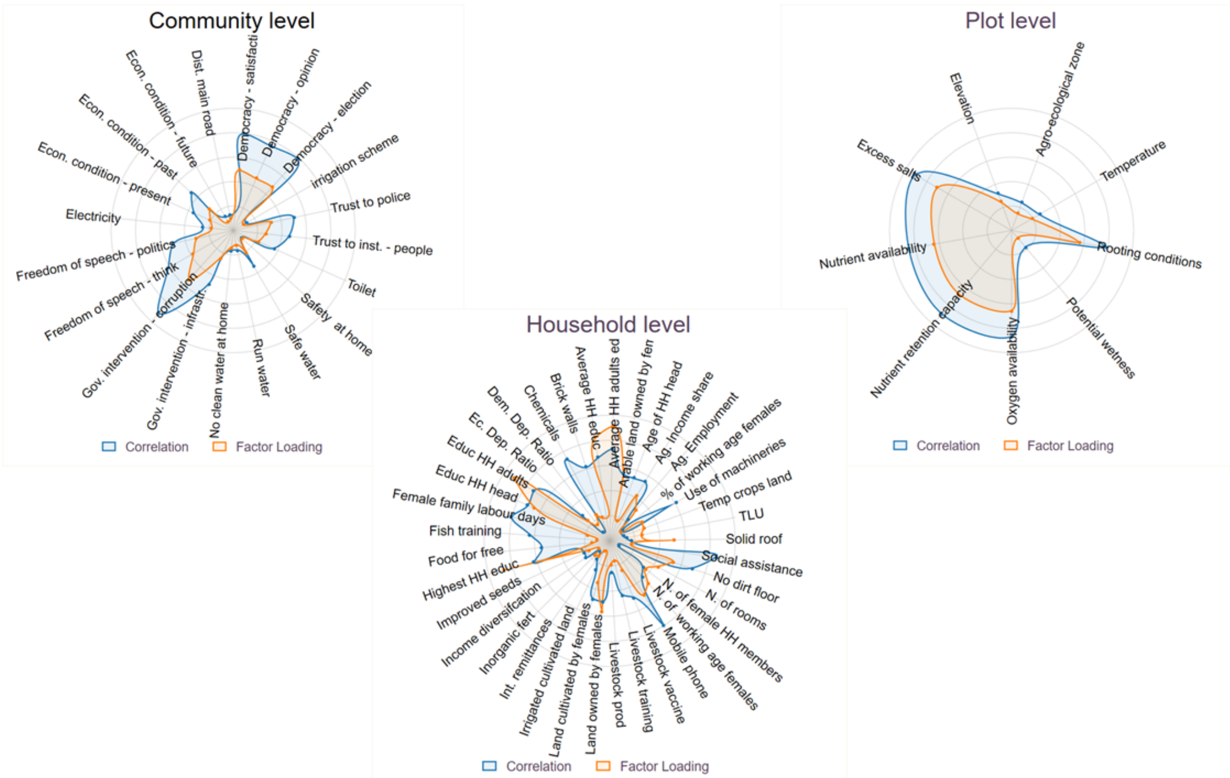
At the community level, the variables considered include proxies for the political stability and level of trust perceived, as well as access to services and infrastructure. Government effectiveness, in particular in handling corruption, and the level of democracy are particularly relevant.

At the plot level, we included variables related to the land cover and soil characteristics. Among them, soil quality, in terms of excess salt, nutrient retention capacity, and oxygen availability, is the most important one.

[12] In the LASSO regression, we used a lambda value within one standard error of the minimum. We recognize the trade-off between model accuracy and the number of variables when choosing between the lambda within one standard error of the minimum and the minimum lambda value. For this study, reducing the number of variables is more relevant than maximizing the accuracy (Hastie et al., 2009, Krstajic et al., 2014).

[13] Factor loadings express the correlations between every single item and the factor. The higher the load the more relevant in defining the factor's dimensionality the item is.

Figure 3: Scale compositions



Source: authors' elaboration.

The Multiple Indicator Multiple Causes (MIMIC) model allows for the simultaneous estimation of the structural and measurement parts (FAO, 2016a). It assumes that variables are measured as deviations from their means and the error terms do not correlate with the latent correlates (Bühn and Schneider, 2008). The error terms of the correlates instead, as well as the error terms of the indicators, are assumed to be correlated.

The corresponding system of equations is the following:

$$\begin{cases} Y_{ht} = \Lambda \eta_{ht} + \epsilon_{ht} \\ \eta_{ht} = \Gamma' X_{ht} + \zeta_{ht} \end{cases} \tag{3}$$

where η_{ht} represents the resilience index of household h in year t; Y_{ht} specifies the indicators, which in the model are the predicted probability for each outcome; X_{ht}

refers to the latent correlates (scale components); Λ and Γ are vectors of regression coefficients. ϵ_{ht} are the residuals, and ζ_{ht} are the disturbances. Thus, the resilience index is at the same time defined as regressand of the components and the common factor between the predicted probabilities of each outcome.

Table 6 presents the estimates of the MIMIC model. RES is the latent variable that identifies the resilience index. Estimated coefficients have been standardized to make them comparable. Estimation is conducted through asymptotic distribution-free method to relax the multivariate normality assumption.

The structural part of the model shows that the component at the household level is negatively associated with resilience. The two other scales instead are positively and significantly associated with the latent variable of resilience. In particular, the component at the community level plays a relevant role in defining the level of resilience.

On the measurement part, we clearly see trade-offs among the outcomes. Specifically, the latent variable of resilience is positively associated with the probability of the environmental[14] and the economic outcomes to be above the threshold, while it is negatively linked to the DDS probability[15].

[14]The coefficient for one of the variables on the measurement side is set to +1 by default. In this case, the one for the prob. of the environmental variable has been constrained. The value different than 1 is the result of the standardization.

[15]The sign persists even when considering different thresholds of DDS

Table 6: MIMIC results.

	Stand. Coeff.	Std. Err.	z	P>z	[95% conf. interval]	
Structural						
RES						
Community level	0.448	0.013	34.020	0.000	0.423	0.474
HH level	-0.069	0.013	-5.110	0.000	-0.095	-0.042
Plot level	0.143	0.014	10.320	0.000	0.116	0.171
Measurement						
Prob(Environment)						
RES	0.842	0.016	52.210	0.000	0.810	0.873
Cons	0.073	0.011	6.870	0.000	0.052	0.094
Prob(Economic)						
RES	0.291	0.014	21.010	0.000	0.264	0.318
Cons	0.041	0.013	3.190	0.001	0.016	0.066 5
Prob(Nutrition)						
RES	-0.492	0.013	-37.180	0.000	-0.518	-0.467
Cons	-0.017	0.013	-1.320	0.187	-0.042	0.008
var(e.Prob(Env.))	0.291	0.027			0.243	0.350
var(e.Prob(Economic))	0.915	0.008			0.900	0.931
var(e.Prob(Nutrition))	0.757	0.013			0.732	0.783
var(e.RES)	0.773	0.013			0.749	0.799

Source: authors' elaboration.
Note: all variables are centered. RES is a latent variable. Prob(X) indicates the conditional probability of the outcome above the threshold, as result from the previous step. Number of obs. = 5,556. [Prob(Env.)]RES = 1. Inverse probability weights applied. The first part of the table reports the structural model, while the second refers to the measurement model. The variable aligned on the left corresponds to the dependent variable, while the variables justified on the right are the regressors. Estimation is conducted through asymptotic distribution-free method.

We used different statistics to test the goodness of fit of the model, including the Tucker Lewis Index (TLI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR)[16]. Not all tests indicated an acceptable goodness of fit for the model (TLI=0.863, CFI=0.932, RMSEA=0.069, SRMR=0.034). Modification indices suggested adding two additional paths: one from the plot level component to the environmental prob-

[16]Acceptable values for the goodness of fit tests are: RMSEA <0.05; SRMR <0.08; TLI and CFI >0.9.

ability outcome, and another from the household level component to the nutritional probability outcome. Incorporating these paths significantly improved the model's goodness of fit (TLI=0.979, CFI=0.993, RMSEA=0.032, SRMR=0.012). The two resulting resilience indexes predicted by the two models are highly correlated (98 percent at the 0.001 significance level). As a robustness check, the analysis was conducted using this alternative model, and no significant differences were found in the final results. These findings are detailed in Appendix 7.2.

Since the latent variable RES is unobserved, it does not have a scale. In order to define a reference unit, the MIMIC model imposes one of the coefficients of the measurement part to be restricted to 1. In this way, the predicted variable of RES can be expressed in units of standard deviation (FAO, 2016a). In this model, a one standard deviation increase in RES results in a single unit increase in the standard deviations of the predicted probability of the environmental outcome. A higher value corresponds to a higher level of resilience.

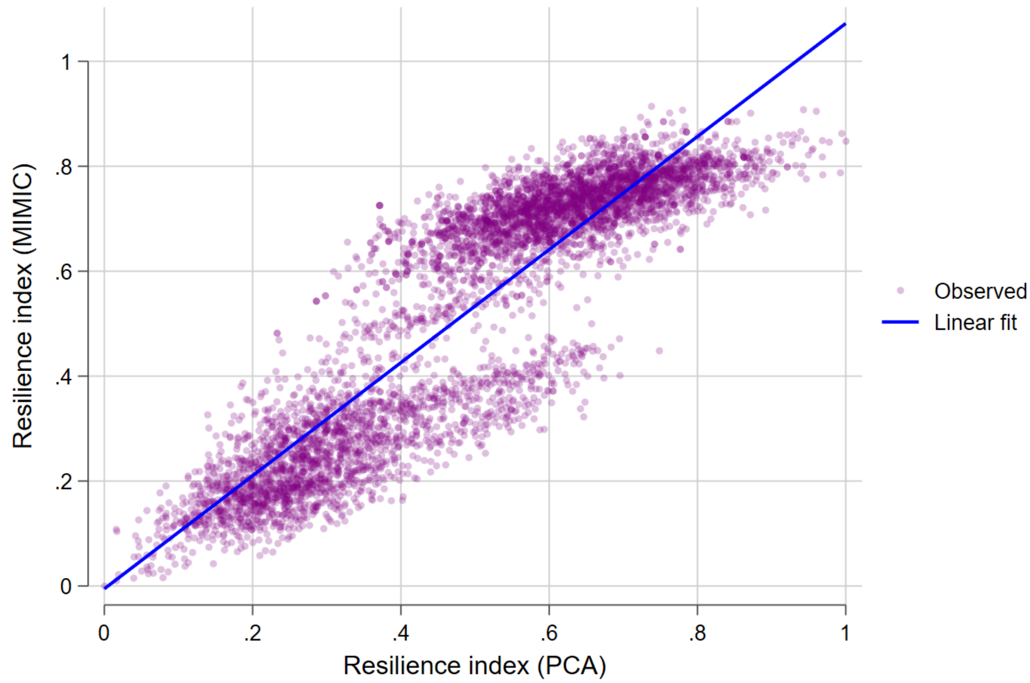
5.4. Comparison of resilience indexes

As the two indexes generated using the two different approaches have different scales, we first normalized them between 0 and 1 to make them comparable[17]. As reported in the scatter plot in Figure 4, the two variables are highly correlated, with a correlation coefficient of 90 percent at the 0.001 significance level.

The coefficients for each probability associated with the predicted resilience variable exhibit consistent signs and similar magnitudes across both approaches. In both methods, resilience is negatively associated with the nutritional probability outcome and positively associated with the other two probabilities, with the environmental dimension showing the largest coefficient. Similar findings emerge when comparing the two indexes across countries, as reported in Table 7. Tanzania exhibits the highest resilience level for both approaches, while Malawi shows the lowest resilience level among the three countries, with a decline over time. Nigeria, on the other hand, managed to increase its resilience level between 2015 and 2018, according to both indexes. Overall, resilience increased by about 3 percentage points over time according to both methods. Therefore, although the magnitude of the resilience level is slightly different when using the PCA and MIMIC approaches, we can conclude that the results are comparable in terms of order and scale.

[17] We divided the values by the minimum value and subtracted the difference between maximum and minimum values.

Figure 4: Relationship between resilience indexes.



Source: authors' elaboration.

Table 7: Resilience indexes by country and year.

	Original		Normalized	
	MIMIC	PCA	MIMIC	PCA
Malawi 2013	-0.777	-1.298	0.308	0.292
Malawi 2016	-1.130	-1.628	0.192	0.238
Nigeria 2015	-0.674	-0.196	0.342	0.471
Nigeria 2018	0.353	0.706	0.680	0.617
Tanzania 2010	0.463	0.629	0.716	0.604
Tanzania 2012	0.500	1.007	0.728	0.666
Total t=2	-0.123	-0.110	0.523	0.485
Total t=3	-0.036	0.111	0.552	0.520

Source: authors' elaboration.
Note: Inverse probability weights applied.

5.5. Testing over different shocks

To test the performance of the two resilience indexes, we test whether farmers with a higher level of resilience were able to cope better in the occurrence of a shock as compared to other farmers. To test this, we ran a two-way fixed effects model based on the following equation:

$$Y_{h,t} = \beta_1 Shock_{h,t} + \beta_2 Res_{h,t} + \beta_3 Shock_{h,t} * Res_{h,t} + \alpha_h + \gamma_t + X_{ht+1} + \epsilon_{h,t} \quad (4)$$

where Y_{ht} is the outcome of the food system (either income, NDVI, and DDS); $Shock_{h,t}$ is a dummy for having experienced the shock at time t;$Res_{h,t}$ is the resilience index previously predicted; $X_{h,t}$ is a set of household and community characteristics; αh and γt are household and time fixed effects, respectively, and $\epsilon_{h,t}$ is the error term. We expect that farmers with higher resilience levels will better manage the adverse effects of shocks, as indicated by the interaction term between resilience and the shock. The two shocks under analysis, namely conflicts and drought, are considered separately.

Table 8 presents the estimates of the individual resilience and shock variables, along with their interaction term, on each outcome variable for the two types of shocks. Overall, the resilience index is shown to be a reliable predictor of the capacity to enhance economic, environmental, and nutritional outcomes in a food system, as it is generally positively associated with each outcome variable. The only exception is the resilience index measured through PCA on income (column 2), where the estimated coefficient is negative for conflict and drought. However, in the case of conflict, the interaction term is positive and significant, indicating that more resilient households can better cope with this shock compared to less resilient households. In contrast, we do not find a significant effect of resilience to mitigate the effect of conflicts when using the MIMIC approach. A similar result is found for drought, where we have a negative effect of resilience when using the PCA approach but a positive estimate for the interaction term. However, in this case, the coefficient is not statistically significant.

A positive and significant interaction term is found for NDVI over conflict shocks. In contrast, the effect on DDS is not significant.

Drought does not appear to significantly affect the three outcome variables. One possible explanation is that we are examining only the short-term impact of drought, specifically focusing on the growing season. The only significant estimate is for NDVI when using the MIMIC model to compute the resilience index, although the signs of both the shock and interaction terms are opposite to what was expected.

Overall, both resilience indexes seem to perform well in being positively associated with the outcomes and mitigating the negative effects of conflicts. Unexpected signs of the estimated coefficients are found for the drought shock over the environmental outcome when using the MIMIC model, and for the economic outcome when using the PCA technique in both shocks.

Table 8: Shocks and farmers' resilience.

	Log(Per capita income)		DDS		NDVI	
	MIMIC	PCA	MIMIC	PCA	MIMIC	PCA
a) Conflict						
Resilience index	0.900***	-2.450***	1.902***	5.625***	673.1***	1,029***
	(0.321)	(0.472)	(0.233)	(0.391)	(33.03)	(58.04)
Shock	-0.634**	-0.766***	-0.0844	0.156	-166.0***	-218.6***
	(0.279)	(0.292)	(0.258)	(0.300)	(38.58)	(52.02)
Resilience index#Shock	0.713	1.067*	0.0892	-0.442	155.3**	312.4***
	(0.493)	(0.564)	(0.442)	(0.584)	(67.15)	(100.4)
b) Drought						
Resilience index	0.641*	-2.594***	1.791***	5.554***	796.2***	1,077***
	(0.388)	(0.494)	(0.292)	(0.409)	(42.67)	(61.47)
Shock	-0.554	-0.237	-0.355	-0.0832	164.2***	-21.21
	(0.534)	(0.530)	(0.363)	(0.358)	(51.21)	(51.75)
Resilience index#Shock	0.922	0.0679	0.873	0.412	-231.4***	6.918
	(0.766)	(0.785)	(0.547)	(0.574)	(77.11)	(81.18)

Source: authors' elaboration.
Note: Control variables, household fixed effects, and time fixed effects included. Inverse probability weights applied. Standard errors clustered at the household level. *** p<0.01, ** p<0.05, * p<0.1

6. Conclusion

In this paper, we empirically measure farmers' resilience over different food system outcomes by embracing the previous literature on development resilience to compute a resilience index with a food system perspective. We define resilience as an index that aggregates the probability of achieving a certain standard over the three dimensions of a food system, namely economic profitability, environmental sustainability, and adequate nutrition. We use and compare two different approaches to aggregate the three probabilities into a unique index. The two related indexes are then measured on a dataset consisting of 8,847 observations of farm households in three African countries spanning three time periods. We considered two different types of shocks, namely drought and conflict, to test the validity of the models.

Considering the individual probabilities for each outcome, we found that nutrition is the dimension in which farmers achieve the highest resilience, having a 67 percent likelihood of being above the threshold on average. In contrast, income is the dimension where farmers are least resilient, with only a 17 percent probability of being above the poverty line. Both dimensions show a declining trend over time, while resilience in the environmental dimension exhibits an overall positive trend. Notably, trends vary across countries, with Malawi experiencing a decline in resilience across all three dimensions between 2013 and 2016.

We then employed two approaches to combine the three probability measures into a single index of food system resilience: a MIMIC model based on the FAO's RIMA approach and a simpler method using PCA. The indexes derived from these approaches are highly correlated and yield similar results. In both cases, resilience is positively associated with economic and environmental outcomes but negatively associated with nutritional outcomes. This finding underscores the presence of trade-offs between food system dimensions and highlights the challenge of achieving improvements across all three outcomes simultaneously.

Both indexes display similar trends over time and across countries. Overall, the resilience level has increased among farmers in the sample, particularly in Tanzania and Nigeria. In Malawi, however, resilience has decreased over the years analyzed. On average, both indexes show that 54 percent of the farm households in the sample reported a score above 0.5 in the second time period, with an increase of the resilient score of 3 percentage points in the later period.

When evaluating the performance of the resilience indexes under different shocks, we found that both indexes generally perform well. They are positively linked to each outcome and can mitigate the negative effects of conflicts and, to some extent, droughts—although the impact on droughts is mostly insignificant. Some inconsistencies however were noted, particularly for the PCA-based resilience index concerning the economic outcome and for the MIMIC model in measuring resilience in response to drought over the environmental outcome.

Based on these findings, we conclude that the PCA-based approach provides results comparable to those of the MIMIC model, but with greater ease of computation and lower demands on data and computational resources.

While the resilience indexes developed in this study demonstrate internal validity, future research is necessary to validate the model across different geographical areas and shocks and to compare it with existing resilience measures.

Acknowledgments

The authors of this article are very grateful to Chris Barrett, who reviewed and commented on an early version of the paper and provided advice on research questions and methodology. We would also like to thank Joanna Upton for reviewing an early version of the paper, as well as Mark Constas, John Hoddinott, Elizabeth Tennant, Seungmin Lee, and Krisztina Kis-Katos for their comments and suggestions. This work has been presented at the 32nd International Conference of Agricultural Economists and at the 5th Global Food Security Conference. This research was financially supported by the German Research Foundation (DFG) through the Sustainable Food Systems Research Training Group (RTG2654).

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the authors utilized ChatGPT from OpenAI to enhance the spelling, grammar, and clarity of the text. Following the use of this tool, the authors thoroughly reviewed and edited the content as necessary.

References

C. B. Barrett and M. A. Constas. Toward a theory of resilience for international development applications. *Proceedings of the National Academy of Sciences*, 111 (40):14625–14630, 2014. doi: 10.1073/pnas.1320880111. URL https://www.pnas.org/doi/abs/10.1073/pnas.1320880111.

C. B. Barrett, L. Christiaensen, M. Sheahan, and A. Shimeles. On the Structural Transformation of Rural Africa. *Journal of African Economies*, 26:i11–i35, 06 2017. ISSN 0963-8024. doi: 10.1093/jae/ejx009. URL https://doi.org/10.1093/jae/ejx009.

C. B. Barrett, K. Ghezzi-Kopel, J. Hoddinott, N. Homami, E. Tennant, J. Upton, and T. Wu. A scoping review of the development resilience literature: Theory, methods and evidence. *World Development*, 146:105612, 2021. ISSN 0305-750X. doi: https://doi.org/10.1016/j.worlddev.2021.105612. URL https://www.sciencedirect.com/science/article/pii/S0305750X21002278.

A. BenYishay, R. Rotberg, J. Wells, Z. Lv, S. Goodman, and D. Kovacevic, L. amd Runfola. Geocoding afrobarometer rounds 1 – 6: Methodology data quality. aiddata. available online. 2017. URL http://geo.aiddata.org.

L. Bizikova, S. Tyler, M. Moench, M. Keller, and D. Echeverria. Climate resilience and food security in central america: a practical framework. *Climate and Development*, 8(5):397–412, 2016. doi: 10.1080/17565529.2015.1064806. URL https://doi.org/10.1080/17565529.2015.1064806.

C. Béné and S. Devereux. *Resilience and Food Security in a Food Systems Context*. 01 2023. ISBN 978-3-031-23534-4. doi: 10.1007/978-3-031-23535-1.

C. Béné, T. Frankenberger, S. Nelson, M. Constas, G. Collins, M. Langworthy, and K. Fox. Food system resilience measurement: principles, framework and caveats. *Food Security*, 11 2023. doi: 10.1007/s12571-023-01407-y.

A. Bühn and F. Schneider. MIMIC Models, Cointegration and Error Correction: An Application to the French Shadow Economy. Technical report, 2008.

C. Chang, J. Gardiner, R. Houang, and Y.-L. Yu. Comparing multiple statistical software for multiple-indicator, multiple-cause modeling: an application of gender disparity in adult cognitive functioning using MIDUS II dataset. *BMC Medical Research Methodology*, 20(1), nov 2020. doi: 10.1186/s12874-020-01150-4. URL https://doi.org/10.1186%2Fs12874-020-01150-4.

J. D. Cissé and C. B. Barrett. Estimating development resilience: A conditional moments-based approach. *Journal of Development Economics*, 135:272–284, 2018. ISSN 0304-3878. doi: https://doi.org/10.1016/j.jdeveco.2018.04.002. URL https://www.sciencedirect.com/science/article/pii/S0304387818303511.

M. Constas, T. Frankenberger, and J. Hoddinott. Resilience measurement principles. Toward an agenda for measurement design. Technical report, Food Security Information Network (FSIN), Resilience Measurement Technical Working Group (RMTWG), Technical Series, 1., 2014a. URL https://www.fsinplatform.org/sites/default/files/paragraphs/documents/FSIN_TechnicalSeries_1.pdf.

M. Constas, T. Frankenberger, J. Hoddinott, N. Mock, D. Romano, C. Béné, and D. Maxwell. A common analytical model for resilience measurement: Causal framework and methodological options. Technical report, Food Security Information Network (FSIN), Resilience Measurement Technical Working Group (RMTWG)., 2014b. URL https://www.fsinplatform.org/sites/default/files/paragraphs/documents/FSIN_TechnicalSeries_1.pdf.

I. Darnhofer, J. Fairweather, and H. Moller. Assessing a farm's sustainability: insights from resilience thinking. *International Journal of Agricultural Sustainability*, 8(3):186–198, 2010. doi: 10.3763/ijas.2010.0480. URL https://doi.org/10.3763/ijas.2010.0480.

R. S. Defries and J. R. G. Townshend. Ndvi-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, 15(17):3567–3586, 1994. doi: 10.1080/01431169408954345. URL https://doi.org/10.1080/01431169408954345.

O. Durowoju, T. Ologunorisa, and A. Akinbobola. Characterization of drought in kaduna river basin, kaduna, nigeria. 19:43–59, 06 2021.

U. Eseosa Ekhator-Mobayode, L. C. Hanmer, E. Rubiano-Matulevich, and D. Jimena Arango. The effect of armed conflict on intimate partner violence: Evidence from the boko haram insurgency in nigeria. *World Development*, 153:105780, 2022. ISSN 0305-750X. doi: https://doi.org/10.1016/j.worlddev.2021.105780. URL https://www.sciencedirect.com/science/article/pii/S0305750X21003958.

J. Eze, P. Ibrahim, S. Tiamiyu, and M. Alfa. Assessment of drought occurrences and its implications on agriculture in niger state, nigeria. 6:1–10, 01 2020.

FAO. Resilience Index Measurement and Analysis – II: Analysing Resilience for Better Targeting and Action. . Technical report, FAO, 2016a. URL https://www.fao.org/3/i5665e/I5665E.pdf.

FAO. RuLIS Codebook: Rural Livelihoods Information System. Technical report, Tor Vergata University (ROme) and FAO, 2022. URL https://www.fao.org/3/ca9548en/ca9548en.pdf.

F. FAO. Minimum dietary diversity for women: a guide for measurement. *Rome: FAO*, 82, 2016b. URL https://openknowledge.fao.org/server/api/core/bitstreams/088f944b-d268-4e04-b57d-027a3b6a56eb/content.

T. Fung and W. Siu. Environmental quality and its changes, an analysis using ndvi. *International Journal of Remote Sensing*, 21(5):1011–1024, 2000. doi: 10.1080/014311600210407. URL https://doi.org/10.1080/014311600210407.

G. Gebremeskel Haile, Q. Tang, S. Sun, Z. Huang, X. Zhang, and X. Liu. Droughts in east africa: Causes, impacts and resilience. *Earth-Science Reviews*, 193:146–161, 2019. ISSN 0012-8252. doi: https://doi.org/10.1016/j.earscirev.2019.04.015. URL https://www.sciencedirect.com/science/article/pii/S0012825218303519.

G. Gómez, R. Monge-Rojas, R. Vargas-Quesada, A. N. Previdelli, D. Quesada, I. Kovalskys, M. Herrera-Cuenca, L. Y. Cortes, M. C. Y. García, R. Liria-Domínguez, A. Rigotti, R. M. Fisberg, G. Ferrari, M. Fisberg, and J. C. Brenes. Exploring the fao minimum dietary diversity indicator as a suitable proxy of micronutrient adequacy in men and women across reproductive and non-reproductive ages in 8 latin american countries. *Food and Nutrition Bulletin*, 0(0):03795721241242920, 2024. doi: 10.1177/03795721241242920. URL https://doi.org/10.1177/03795721241242920. PMID: 38807356.

T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)*. 02 2009. ISBN 0387848576.

A. Heim and A. Paksi. Low dietary diversity and its influencing factors among a san group in namibia. *BMC Research Notes*, 12, 12 2019. doi: 10.1186/s13104-019-4408-8.

S. Henly-Shepard and B. Sagara. Resilience Measurement Practical Guidance Note Series: An Overview. Technical report, Mercy Corps, Resilience Evaluation, Analysis and Learning (REAL) Associate Award., 2018. URL https://www.fsnnetwork.org/sites/default/files/GN_Overview_Final_0.pdf.

R. Henriksson, K. Vincent, and K. Naidoo. Exploring the adaptive capacity of sugarcane contract farming schemes in the face of extreme events. *Frontiers in Climate*, 3:578544, 03 2021. doi: 10.3389/fclim.2021.578544.

J. F. Hoddinott. Looking at development through a resilience lens. In S. Fan, R. Pandya-Lorch, and S. Yosef, editors, *2013 Global Food Policy Report*, IFPRI book chapters, chapter 3. International Food Policy Research Institute (IFPRI), 2014. URL https://ideas.repec.org/h/fpr/ifpric/128441.html.

IFAD. Rural development report 2016: fostering inclusive rural transformation. page 375 pp., 2016. URL https://www.ifad.org/en/web/knowledge/-/rural-development-report-2016-fostering-inclusive-rural-transformation.

K. Jia, S. Liang, L. Zhang, X. Wei, Y. Yao, and X. Xie. Forest cover classification using landsat etm+ data and time series modis ndvi data. *International Journal of Applied Earth Observation and Geoinformation*, 33:32–38, 2014. ISSN 1569-8432. doi: https://doi.org/10.1016/j.jag.2014.04.015. URL https://www.sciencedirect.com/science/article/pii/S0303243414001056.

E. Knippenberg, N. Jensen, and M. Constas. Quantifying household resilience with high frequency data: Temporal dynamics and methodological options. *World Development*, 121:1–15, 2019. ISSN 0305-750X. doi: https://doi.org/10.1016/j.worlddev.2019.04.010. URL https://www.sciencedirect.com/science/article/pii/S0305750X19300956.

D. Krstajic, L. Buturovic, D. Leahy, and S. Thomas. Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of cheminformatics*, 6:10, 03 2014. doi: 10.1186/1758-2946-6-10.

V. Kshirsagar, J. Wieczorek, S. Ramanathan, and R. Wells. Household poverty classification in data-scarce environments: a machine learning approach, 2017.

S. Lee, K. A. Abay, C. B. Barrett, and J. F. Hoddinott. Estimating Multidimensional Development Resilience. Working paper, Cornell University, May 2024.

B. Li, W. Zhou, Y. Zhao, Q. Ju, Z. Yu, Z. Liang, and K. Acharya. Using the spei to assess recent climate change in the yarlung zangbo river basin, south tibet. *Water*, 7(10):5474–5486, 2015. ISSN 2073-4441. doi: 10.3390/w7105474. URL https://www.mdpi.com/2073-4441/7/10/5474.

L. Liang, F. Chen, L. Shi, and S. Niu. Ndvi-derived forest area change and its driving factors in china. *PLOS ONE*, 13(10):1–19, 10 2018. doi: 10.1371/journal.pone.0205885. URL https://doi.org/10.1371/journal.pone.0205885.

H. Maleksaeidi and E. Karami. Social-ecological resilience and sustainable agriculture under water scarcity. *Agroecology and Sustainable Food Systems*, 37(3): 262–290, 2013. doi: 10.1080/10440046.2012.746767. URL https://doi.org/10.1080/10440046.2012.746767.

T. B. Mckee, N. J. Doesken, and J. R. Kleist. The relationship of drought frequency and duration to time scales. pages 179–184, Anaheim, California, 1993. URL https://api.semanticscholar.org/CorpusID:129950974.

M. A. Meyer. The role of resilience in food system studies in low- and middle-income countries. *Global Food Security*, 24:100356, 2020. ISSN 2211-9124. doi: https://doi.org/10.1016/j.gfs.2020.100356. URL https://www.sciencedirect.com/science/article/pii/S2211912420300092.

R. Milestad, L. Westberg, U. Geber, and J. Björklund. Enhancing Adaptive Capacity in Food Systems: Learning at Farmers' Markets in Sweden. *Ecology and Society*, 15(3), 2010. ISSN 17083087. URL http://www.jstor.org/stable/26268172.

B. Moldan, S. Janoušková, and T. Hák. How to understand and measure environmental sustainability: Indicators and targets. *Ecological Indicators*, 17:4–13, 2012. ISSN 1470-160X. doi: https://doi.org/10.1016/j.ecolind.2011.04.033. URL https://www.sciencedirect.com/science/article/pii/S1470160X11001282. Indicators of environmental sustainability: From concept to applications.

M. Mridha, A. A. Shamim, M. M. Hossain, M. Hasan, A. A. Mohammad Hanif, M. Hossaine, M. A. Ullah, S. K. Sarker, S. M. Rahman, D. K. Mitra, and M. E. Haque. Dietary practices of men in bangladesh: Evidence from the national nutrition surveillance. *Current Developments in Nutrition*, 4:nzaa046$_0$49, 2020.$ISSN$2475$-$2991.doi : . URL https://www.sciencedirect.com/science/article/pii/S2475299123080678.

E. Mwangi, F. Wetterhall, E. Dutra, F. Di Giuseppe, and F. Pappenberger. Forecasting droughts in east africa. *Hydrology and Earth System Sciences*, 18(2):611–620, 2014. 10.5194/hess-18-611-2014. URL https://hess.copernicus.org/articles/18/611/2014/.

I. Orimoloye, A. Olusola, O. Ololade, and S. Adelabu. A persistent fact: reflections on drought severity evaluation over nigerian sahel using mod13q1. *Arabian Journal of Geosciences*, 14:1–18, 10 2021. 10.1007/s12517-021-08369-5.

A. A. Paulo, R. D. Rosa, and L. S. Pereira. Climate trends and behaviour of drought indices based on precipitation and evapotranspiration in portugal. *Natural Hazards and Earth System Sciences*, 12(5):1481–1491, 2012. 10.5194/nhess-12-1481-2012. URL https://nhess.copernicus.org/articles/12/1481/2012/.

N. Pettorelli, J. O. Vik, A. Mysterud, J.-M. Gaillard, C. J. Tucker, and N. C. Stenseth. Using the satellite-derived ndvi to assess ecological responses to environmental change. *Trends in Ecology Evolution*, 20(9):503–510, 2005. ISSN 0169-5347. https://doi.org/10.1016/j.tree.2005.05.011. URL https://www.sciencedirect.com/science/article/pii/S016953470500162X.

P. Pingali, L. Alinovi, and J. Sutton. Food security in complex emergencies: enhancing food system resilience. *Disasters*, 29 Suppl 1:S5–24, 2005.

S. A. Sader and J. C. Winne. Rgb-ndvi colour composites for visualizing forest change dynamics. *International Journal of Remote Sensing*, 13(16):3055–3067, 1992. 10.1080/01431169208904102. URL https://doi.org/10.1080/01431169208904102.

L. C. Smith and T. R. Frankenberger. Does Resilience Capacity Reduce the Negative Impact of Shocks on Household Food Security? Evidence from the 2014 Floods in Northern Bangladesh. *World Development*, 102:358–376, 2018. ISSN 0305-750X. https://doi.org/10.1016/j.worlddev.2017.07.003. URL https://www.sciencedirect.com/science/article/pii/S0305750X17302371.

I. D. Soane, R. Scolozzi, A. Gretter, and K. Hubacek. Exploring panarchy in alpine grasslands: an application of adaptive cycle concepts to the conservation of a cultural landscape. *Ecology and Society*, 17(3), 2012. ISSN 17083087. URL http://www.jstor.org/stable/26269086.

D. Tendall, J. Joerin, B. Kopainsky, P. Edwards, A. Shreck, Q. Le, P. Kruetli, M. Grant, and J. Six. Food system resilience: Defining the concept. *Global Food Security*, 6:17–23, 2015. ISSN 2211-9124. https://doi.org/10.1016/j.gfs.2015.08.001. URL https://www.sciencedirect.com/science/article/pii/S2211912415300031.

B. C. Thiede, M. Hancock, A. Kodouda, and J. Piazza. Exposure to Armed Conflict and Fertility in Sub-Saharan Africa. *Demography*, 57(6):2113–2141, 10 2020. ISSN 0070-3370. 10.1007/s13524-020-00923-2. URL https://doi.org/10.1007/s13524-020-00923-2.

J. Upton, S. Constenla-Villoslada, and C. B. Barrett. Caveat utilitor: A comparative assessment of resilience measurement approaches. *Journal of Development Economics*, 157:102873, 2022. ISSN 0304-3878. https://doi.org/10.1016/j.jdeveco.2022.102873. URL https://www.sciencedirect.com/science/article/pii/S030438782200044X.

D. F. van Apeldoorn, K. Kok, M. P. Sonneveld, and T. A. Veldkamp. Panarchy Rules: Rethinking Resilience of Agroecosystems, Evidence from Dutch Dairy-Farming. *Ecology and Society*, 16(1), 2011. ISSN 17083087. URL http://www.jstor.org/stable/26268873.

J. Zhang, D. Liang, and A. Zhao. Dietary diversity and the risk of fracture in adults: A prospective study. *Nutrients*, 12:3655, 11 2020. 10.3390/nu12123655.

M. Zurek, J. Ingram, A. Sanderson Bellamy, C. Goold, C. Lyon, P. Alexander, A. Barnes, D. P. Bebber, T. D. Breeze, A. Bruce, L. M. Collins, J. Davies, B. Doherty, J. Ensor, S. C. Franco, A. Gatto, T. Hess, C. Lamprinopoulou, L. Liu, M. Merkle, L. Norton, T. Oliver, J. Ollerton, S. Potts, M. S. Reed, C. Sutcliffe, and P. J. Withers. Food system resilience: Concepts, issues, and challenges. *Annual Review of Environment and Resources*, 47(Volume 47, 2022):511–534, 2022. ISSN 1545-2050. https://doi.org/10.1146/annurev-environ-112320-050744. URL https://www.annualreviews.org/content/journals/10.1146/annurev-environ-112320-050744.

7. Appendix

7.1. Appendix 1 - Regression tables

Table 9: Estimates of probability, per capita income.

	(1) Log(per capita income)	(2) Variance of log(per capita income)	(3) Probability
Log(per capita income)$_{t-1}$	0.160***	-0.0867	0.0174***
	(0.0177)	(0.0622)	(0.000257)
HH size	-0.0266**	-0.120**	-0.0104***
	(0.0134)	(0.0487)	(0.000281)
N. of female HH member	-0.0365*	0.0671	-0.00155***
	(0.0219)	(0.0828)	(0.000422)
HH head is female	-0.250***	-0.444	-0.0657***
	(0.0784)	(0.305)	(0.00131)
HH receives food for free	3.46e-05	0.000317	9.52e-06
	(0.000525)	(0.00207)	(7.08e-06)
Log(Age of HH head)	-0.185**	0.367	-0.0152***
	(0.0744)	(0.306)	(0.00117)
HH head is married	-0.0380	-0.424	-0.0341***
	(0.0721)	(0.288)	(0.00135)
Use of chemicals	0.223***	0.382	0.0562***
	(0.0720)	(0.313)	(0.00113)
Dem. Dep. Ratio	-0.176***	-0.0631	-0.0289***
	(0.0264)	(0.0994)	(0.000626)
Tropic-warm/subhumid	-0.213***	0.608**	-0.00507***
	(0.0619)	(0.261)	(0.00107)
Tropic-warm/humid	-0.728**	2.329***	-0.0173***
	(0.293)	(0.852)	(0.00363)
Tropic-cool/semiarid	0.0464	0.0378	0.0130***
	(0.0809)	(0.367)	(0.00170)
Tropic-cool/subhumid	0.0523	-0.546	-0.0209***
	(0.0766)	(0.338)	(0.00158)

Tropic-cool/humid	0.0856	-1.045***	-0.0426***
	(0.159)	(0.366)	(0.00375)
Land owned (hectares)	0.0659***	-0.216***	-0.00144***
	(0.0102)	(0.0413)	(0.000409)
Country=Nigeria	-0.191*	2.024***	0.0829***
	(0.108)	(0.454)	(0.00153)
Country=Tanzania	-0.169**	1.954***	0.0831***
	(0.0850)	(0.375)	(0.00143)
Year=2012	0.354***	-0.783***	0.0223***
	(0.0599)	(0.251)	(0.000857)
Year=2013	0.164**	0.446	0.0569***
	(0.0671)	(0.306)	(0.000882)
Year=2015	-0.883***	4.887***	-0.00475***
	(0.174)	(0.589)	(0.00168)
Constant	0.691**	0.674	0.314***
	(0.295)	(1.229)	(0.00501)
Observations	5,846	5,846	5,766
R-squared	0.12	0.09	0.91

Note: *** p<0.01, ** p<0.05, * p<0.1. Inverse probability weights applied. Standard errors clustered at the household level.

Source: authors' elaboration.

Table 10: Estimates of probability, DDS.

	(1) DDS	(2) Variance of DDS	(3) Probability
DDS_{t-1}	0.248***	-0.0219	0.0504***
	(0.0150)	(0.0352)	(0.000577)
HH head is female	0.195***	-0.395**	0.0527***
	(0.0693)	(0.166)	(0.00253)
Log(Age of HH head)	-0.0570	-0.0731	-0.0149***
	(0.0664)	(0.153)	(0.00254)
HH head is married	0.0909	-0.224	0.0250***
	(0.0654)	(0.165)	(0.00188)
Educ. of HH head	0.0521***	-0.0176	0.0107***
	(0.00551)	(0.0125)	(0.000292)
HH size	0.0306***	-0.101***	0.0109***
	(0.0113)	(0.0274)	(0.000427)
HH receives food for free	0.00360***	-0.00394***	0.000897***
	(0.000422)	(0.000935)	(1.43e-05)
Ag. Employment	-0.282***	-0.0393	-0.0521***
	(0.0494)	(0.105)	(0.00174)
Land owned (hectares)	0.00503	-0.0173	0.00228***
	(0.00986)	(0.0242)	(0.000320)
Use of chemicals	0.167***	-0.0568	0.0398***
	(0.0519)	(0.115)	(0.00183)
Use of machineries	0.223***	0.150	0.0384***
	(0.0528)	(0.113)	(0.00200)
Dem. Dep. Ratio	-0.0428*	-0.0414	-0.00743***
	(0.0230)	(0.0523)	(0.000931)
N. of female HH members	0.00589	0.0734*	-0.00248***
	(0.0184)	(0.0412)	(0.000701)
Access to electricity	0.590***	-0.273	0.121***
	(0.0769)	(0.190)	(0.00489)
HH owns dwelling	-0.170**	-0.0106	-0.0209***
	(0.0753)	(0.181)	(0.00344)
HH engages in livestock prod.	0.368***	-0.0834	0.0766***

	(0.0500)	(0.123)	(0.00169)
Tropic-warm/subhumid	0.198***	0.196*	0.0212***
	(0.0546)	(0.115)	(0.00294)
Tropic-warm/humid	0.709***	0.189	0.144***
	(0.153)	(0.348)	(0.00419)
Tropic-cool/semiarid	0.00664	0.964***	-0.0342***
	(0.0945)	(0.239)	(0.00335)
Tropic-cool/subhumid	0.137**	0.187	0.00463
	(0.0697)	(0.160)	(0.00357)
Tropic-cool/humid	0.399***	-0.324	0.0839***
	(0.133)	(0.335)	(0.00653)
Distance to road (km)	-0.00355***	-0.00557*	-0.00067***
	(0.00121)	(0.00302)	(5.08e-05)
Herfindahl index of income	-0.322***	0.0855	-0.0802***
	(0.106)	(0.256)	(0.00361)
Country=Nigeria	-1.096***	0.484**	-0.217***
	(0.115)	(0.247)	(0.00403)
Country=Tanzania	-0.786***	-0.462**	-0.126***
	(0.0816)	(0.184)	(0.00334)
Year=2012	-0.121**	0.150	-0.0299***
	(0.0539)	(0.127)	(0.00115)
Year=2013	0.369***	-0.270*	0.0664***
	(0.0734)	(0.142)	(0.00199)
Year=2015	-0.117	-0.778***	-0.0207***
	(0.110)	(0.217)	(0.00232)
Constant	4.405***	3.711***	0.365***
	(0.313)	(0.736)	(0.0112)
Observations	5,741	5,741	5,732
R-squared	0.29	0.02	0.95

Note: *** p<0.01, ** p<0.05, * p<0.1. Inverse probability weights applied. Standard errors clustered at the household level.

Source: authors' elaboration.

Table 11: Estimates of probability, percentage change of NDVI.

	(1) % change of NDVI	(2) Variance of % change of NDVI	(3) Probability
Educ. HH head	0.00148	-0.0430	-2.99e-05
	(0.0206)	(0.599)	(0.000106)
Ag. Employment	-0.0337	3.459	0.00452***
	(0.172)	(5.979)	(0.00122)
Land owned (hectares)	-0.0384	-2.677**	-0.00244***
	(0.0411)	(1.199)	(0.000292)
Use of chemicals	-0.753***	-7.382*	-0.0654***
	(0.186)	(4.198)	(0.00157)
Use of irrigation	0.125	9.108	0.0390***
	(0.296)	(13.24)	(0.00242)
Use of machineries	0.377*	3.593	0.0255***
	(0.201)	(5.279)	(0.00139)
HH engages in livestock prod.	0.306**	1.390	0.0222***
	(0.153)	(3.489)	(0.00109)
Log (total rainfall) (in mm)	-1.340***	-2.781	-0.101***
	(0.360)	(10.88)	(0.00230)
Log(temperature) (in DegC*10)	-0.0282	20.47	0.00124
	(1.144)	(36.39)	(0.00730)
Tropic-warm/subhumid	0.801***	3.434	0.0839***
	(0.215)	(4.361)	(0.00163)
Tropic-warm/humid	-0.814	0.306	-0.0174***
	(0.836)	(27.91)	(0.00381)
Tropic-cool/semiarid	1.404**	61.88**	0.140***
	(0.658)	(24.49)	(0.00472)
Tropic-cool/subhumid	0.478	-18.40*	0.0637***
	(0.333)	(10.10)	(0.00251)
Tropic-cool/humid	-0.262	-46.18***	0.00910***
	(0.537)	(14.06)	(0.00314)
Distance to road (km)	-0.0157***	-0.206	-0.00121***
	(0.00597)	(0.197)	(3.38e-05)
Herfindahl index of income	-0.449	8.000	-0.0233***

	(0.325)	(8.815)	(0.00246)
Country=Nigeria	-0.0288	-23.68*	0.0383***
	(0.425)	(12.61)	(0.00411)
Country=Tanzania	0.799**	51.08***	0.0741***
	(0.327)	(11.48)	(0.00284)
Year=2010	0.543**	26.32***	0.0432***
	(0.229)	(6.622)	(0.00189)
Year=2012	1.457***	56.24***	0.0866***
	(0.255)	(8.666)	(0.00205)
Year=2013	-2.397***	31.88***	-0.218***
	(0.283)	(7.418)	(0.00349)
Year=2015	-3.955***	66.78***	-0.262***
	(0.376)	(10.57)	(0.00366)
Year=2016	-6.857***	46.97***	-0.329***
	(0.330)	(8.225)	(0.00354)
Year=2018	2.613***	82.50***	0.188***
	(0.418)	(10.95)	(0.00359)
Constant	8.184	-118.1	1.036***
	(6.878)	(211.3)	(0.0465)
Observations	8,500	8,500	8,130
R-squared	0.20	0.03	0.95

Note: *** p<0.01, ** p<0.05, * p<0.1. Inverse probability weights applied. Standard errors clustered at the household level.

Source: authors' elaboration.

7.2. Appendix 2 - Robustness checks

In this section, we conduct robustness tests to assess the validity of the data and methodology used. Specifically, we first ran different polynomial specifications for income and the DDS variables to identify which order of the lagged value to include in the regression to compute the conditional probability. Then, we addressed the limitations related to the environmental variable, such as coordinate´s precision and the effectiveness of NDVI in measuring environmental sustainability. To address the first limitation, we used the EVI variable measured at the plot location and compared it with the EVI variable calculated using modified coordinates. This variable is available only for Tanzania. For the second limitation, we replaced the NDVI with the percentage of forest as an alternative measure.

In this section, we also report the results for the MIMIC model adjusted based on modification indices, as described in Section 5.3. Finally, we performed a sensitivity analysis by calculating the resilience indexes separately for each country.

7.2.1. Polynomial Specifications

Table 12: Estimates of per capita income – Polynomial Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Log(Income)}^1_{t-1}$	0.160***	0.328***	0.328***	0.364***	0.376***	0.347***	0.307***	0.288***
	(0.0178)	(0.0327)	(0.0330)	(0.0412)	(0.0493)	(0.0512)	(0.0611)	(0.0594)
$\text{Log(Income)}^2_{t-1}$		0.0343***	0.0318	0.0283	0.0479	0.0651*	0.0271	0.0727
		(0.00565)	(0.0206)	(0.0223)	(0.0301)	(0.0369)	(0.0452)	(0.0641)
$\text{Log(Income)}^3_{t-1}$			-0.000363	-0.00998	-0.0117	0.00476	0.0237	0.0454
			(0.00267)	(0.0109)	(0.0120)	(0.0196)	(0.0256)	(0.0301)
$\text{Log(Income)}^4_{t-1}$				-0.00122	-0.00402	-0.00513	0.00705	-0.00470
				(0.00121)	(0.00480)	(0.00494)	(0.0114)	(0.0170)
$\text{Log(Income)}^5_{t-1}$					-0.000317	-0.00184	-0.00258	-0.00762
					(0.000474)	(0.00176)	(0.00180)	(0.00543)
$\text{Log(Income)}^6_{t-1}$						-0.000157	-0.000981	-0.000600
						(0.000169)	(0.000697)	(0.000752)
$\text{Log(Income)}^7_{t-1}$							-7.73e-05	0.000204
							(6.46e-05)	(0.000281)
$\text{Log(Income)}^8_{t-1}$								2.35e-05
								(2.34e-05)
Controls	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.12	0.13	0.13	0.13	0.13	0.13	0.13	0.13
AIC	22595	22523	22525	22525	22526	22706	22525	22525
t-test	0.56	0.97	0.97	0.98	0.99	0.99	1.00	-

Source: authors' elaboration.

Note: The dependent variable is income at time t. Income is computed as per capita household income in USD 2017 PPP. The t-test shows the p-value on the equality of means between predicted values from the specific estimation and the 8th order polynomial specification. *** p<0.01, ** p<0.05, * p<0.1

Table 13: Estimates of DDS – Polynomial Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DDS^1_{t-1}	0.248***	0.117*	-0.0767	-0.639**	-1.323***	-1.356	-2.809*
	(0.0150)	(0.0680)	(0.184)	(0.305)	(0.472)	(0.874)	(1.638)
DDS^2_{t-1}		0.0115**	0.0496	0.252***	0.659**	0.689	2.502
		(0.00576)	(0.0336)	(0.0971)	(0.259)	(0.736)	(1.851)
DDS^3_{t-1}			-0.00227	-0.0302**	-0.127**	-0.138	-1.040
			(0.00196)	(0.0130)	(0.0608)	(0.255)	(0.862)
DDS^4_{t-1}				0.00130**	0.0114*	0.0132	0.240
				(0.000607)	(0.00633)	(0.0431)	(0.206)
DDS^5_{t-1}					-0.000381	-0.000534	-0.0308
					(0.000241)	(0.00351)	(0.0268)
DDS^6_{t-1}						4.85e-06	0.00207
						(0.000111)	(0.00178)
DDS^7_{t-1}							-5.64e-05
							(4.80e-05)
Controls	yes	yes	yes	yes	yes	yes	yes
R-squared	0.29	0.29	0.29	0.29	0.29	0.29	0.29
AIC	20619	20615	20615	20612	20611	20613	20611
t-test	0.97	0.99	0.98	1.00	1.00	1.00	1.00

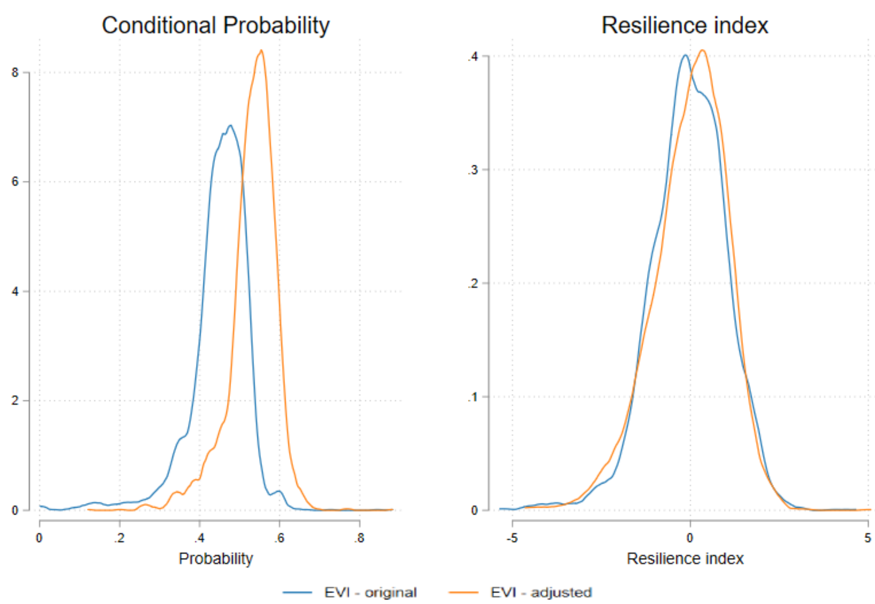
Source: authors' elaboration.

Note: The dependent variable is DDS at time t. The t-test shows the p-value on the equality of means between predicted values from the specific estimation and the 7th order polynomial specification. The estimates based on the 8th order polynomial specification are not reported because the term was omitted due to collinearity issues. *** p<0.01, ** p<0.05, * p<0.1

7.2.2. Tackling measurement error of households location

As discussed in the paper, the NDVI variable used for the environmental dimension relies on adjusted household location coordinates. This potential measurement error could bias our results. To examine this, we compared outcomes computed using the actual location with those based on the adjusted coordinates. This check is feasible only for Tanzania using the EVI variable. In fact, along with the raw data, the World Bank provides some georeferenced data, which are computed based on the right location of the household. Across the countries and rounds analyzed, the only NDVI-like variable consistently available for all three years is the average EVI value at the peak greenness of the growing season in Tanzania. We then retrieved the same variable from MODIS based on the adjusted coordinates. We then calculated the conditional probability and the resilience index for each EVI variable. Results are presented in Figure 5. While a slight difference in the conditional probability is observed between the two variables, this difference disappears in the resilience index, showing nearly identical distributions. This finding indicates that the measurement error does not affect the study's final results.

Figure 5: Distribution of probability and resilience index based on original and adjusted EVI.



Source: Authors' elaboration.
Note: EVI computed on the exact household location and using adjusted coordinates. Resilience index computed using the PCA approach.

7.2.3. Using a different variable of environmental sustainability

Given that NDVI may not be the most suitable variable to assess environmental sustainability, we conducted a robustness check using an alternative proxy that captures forest cover. To implement this, we obtained land cover data for Nigeria from FAOSTAT's Land Cover domain, derived from the MODIS Collection 6.1 Global Land Cover Classification Systems (LCCS) Land Cover Types product (MCD12Q1). The land is categorized into 13 distinct classes[18]. We calculated the percentage of land covered by forest within specific buffer zones (2 km and 5 km for urban and rural areas, respectively) surrounding each household location, using a broad forest definition that includes dense forest, open forest, and forest/cropland mosaics.

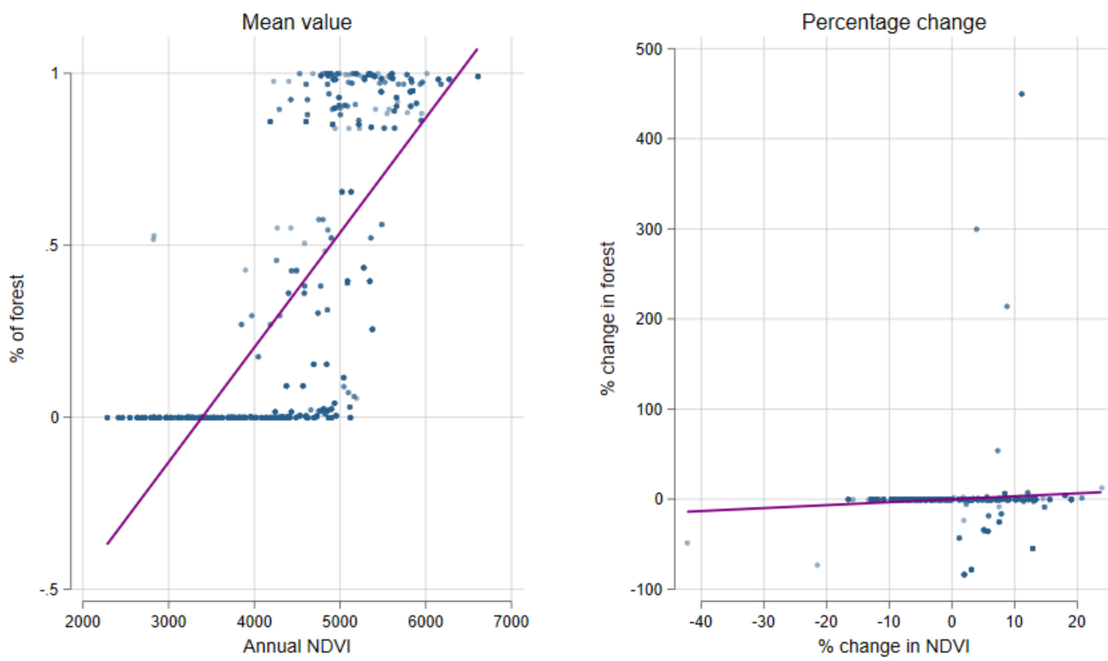
We first assessed the correlation between forest cover and NDVI, finding a strong positive correlation of 74% at the 0.01 significance level. However, examining changes across rounds revealed divergent trends. In particular, many farmers reside in areas without forests, showing no change over time. This is a key reason why we did not use forest cover as the primary variable in the main analysis. NDVI, in contrast, displayed greater variability, as shown in Figure 6. While the overall forest-covered area declined, as seen in the second column of Table 14, the changes, though positive on average, were minimal, ranging from -0.26% to 0.51%.

Next, we estimated the probability of maintaining forest cover (indicating deforestation avoidance) and used this probability—alongside economic and nutritional ones—to calculate a resilience index via the PCA approach. We then compared this resilience index and associated probability with the original index based on NDVI changes from a 5-year baseline average, as well as with an alternative resilience index that tracked changes of NDVI across rounds for better comparability with the forest cover index.

We observed an opposite trend over time when using forest cover compared to NDVI, both in the probability and in the resulting resilience index. This aligns with the distinct temporal patterns of the two variables, underscoring the importance of variable selection as a proxy for each food system dimension in determining the final resilience level.

[18] Based on the combined MODIS-LCCS classification. Full details are available in Appendix 7.4.

Figure 6: Correlation between NDVI and forest cover.



Source: Authors' elaboration.

Table 14: Comparison of NDVI and forest cover measures.

	Mean		Percentage change (%)		Probability			Resilience index		
	NDVI	Forest Cover	NDVI	Forest Cover	NDVI (original)	Forest Cover	NDVI (Lag)	NDVI (original)	Forest Cover	NDVI (Lag)
t=1	4342	0.305	.	.	.	.	.	.	.	.
t=2	4129	0.302	-5.20	-0.26	0.121	0.525	0.139	0.309	0.684	0.339
t=3	4410	0.296	6.82	0.51	0.558	0.502	0.911	0.613	0.579	0.637
Total	4294	0.301	0.81	0.12	0.338	0.511	0.523	0.460	0.620	0.488

Source: authors' elaboration.
Note: Resilience is computed using the PCA approach, values are normalized between 0 and 1. The original NDVI is computed over the 5-year average at the baseline. Inverse probability weights applied. Data for Nigeria only.

7.2.4. Modified MIMIC model

As the original resilience index computed using the standard MIMIC model showed poor goodness of fit in some tests, we ran a modified version of the model by adding two paths, as suggested by the modification indices: one from the plot level component to the environmental probability outcome, and another from the household level component to the nutritional probability outcome. In this way, the model increases its fit and passes all tests.

Table 15: Modified MIMIC results.

	Stand. Coeff.	Std. Err.	z	P>z	[95% conf. interval]	
Structural						
Prob(Environment)						
RES	0.786	0.017	45.960	0.000	0.753	0.820
Plot level	0.142	0.017	8.490	0.000	0.109	0.175
Cons	0.060	0.011	5.520	0.000	0.039	0.081
Prob(Nutrition)						
RES	-0.469	0.013	-36.730	0.000	-0.494	-0.444
HH level	0.116	0.013	8.930	0.000	0.091	0.141
Cons	-0.032	0.013	-2.460	0.014	-0.057	-0.006
RES						
Community level	0.470	0.014	33.900	0.000	0.443	0.497
HH level	-0.040	0.014	-2.730	0.006	-0.068	-0.011
Plot level	-0.014	0.024	-0.570	0.566	-0.060	0.033
Measurement						
Prob(Economic)						
RES	0.309	0.014	21.660	0.000	0.281	0.337
Cons	0.049	0.013	3.860	0.000	0.024	0.074
var(e.Prob(Env.))	0.366	0.027			0.318	0.422
var(e.Prob(Economic))	0.904	0.009			0.887	0.922
var(e.Prob(Nutrition))	0.761	0.012			0.737	0.785
var(e.RES)	0.776	0.013			0.751	0.803

Source: authors' elaboration.

Note: all variables are centered. RES is a latent variable. Prob(X) indicates the conditional probability of the outcome above the threshold, as result from the previous step. Number of obs. = 5,556. [Prob(Env.)]RES = 1. Inverse probability weights applied. The first part of the table reports the structural model, while the second refers to the measurement model. The variable aligned on the left corresponds to the dependent variable, while the variables justified on the right are the regressors. Estimation is conducted through asymptotic distribution-free method.

Comparison of the results with the original model shows very similar coefficients, both in terms of sign and magnitude, and p-values for all associations except for the plot level variable to the latent variable of resilience. In fact, the coefficient using the modified MIMIC model shows a negative sign, although it is not significant. Results of the modified MIMIC model are reported in Table 15. Both indexes show the same results in terms of trends over time, in total and for each country, and sign, with similar magnitude of resilience level.

7.2.5. Sensitivity analysis

To validate our findings, we conducted a sensitivity analysis by running the model separately for each country. We began by calculating the probability for each dimension within each country, followed by computing the resilience indexes using both approaches. However, we faced some issues with the SEM-based approach. In certain instances, the model failed to converge, and even when it did, the model fit was poor. Specifically, the model did not converge for Tanzania, while in Malawi and Nigeria, it converged but did not meet the acceptable thresholds for all goodness-of-fit tests. In contrast, the PCA-based approach was consistently computable, further demonstrating that it is a more flexible and versatile method compared to the SEM approach.

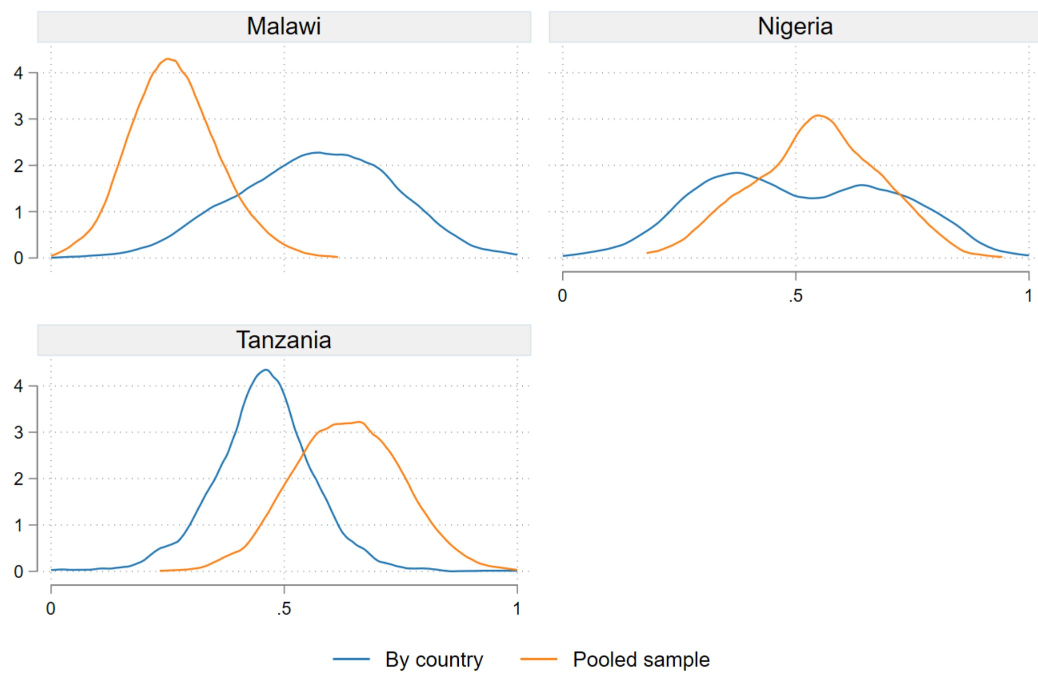
For this reason, we present only the results from the index calculated using the PCA approach. While the overall index—both the original and normalized versions—shows no statistically significant differences between the pooled cross-country sample and the individual country samples, variations in the index values do appear over time and across countries, as shown in Table 16. These differences are further confirmed in Figure 7, which depicts the distribution of the normalized index for the pooled sample and each country sample, by country. Nevertheless, it is important to note that the trend of the index over time remains consistent across all countries and the pooled sample.

Table 16: Comparison of resilience index between the overall sample and by country.

	Original			Normalized		
	All countries	By country	Mean diff.	All countries	By country	Mean diff.
Malawi 2013	-1.298	0.538	***	0.292	0.645	***
Malawi 2016	-1.628	-0.541	***	0.238	0.488	***
Nigeria 2015	-0.196	-0.887	***	0.471	0.361	***
Nigeria 2018	0.706	0.905	***	0.617	0.655	***
Tanzania 2010	0.629	-0.067	***	0.604	0.449	***
Tanzania 2012	1.007	0.068	***	0.666	0.463	***
T=2	-0.110	-0.016	***	0.485	0.496	**
T=3	0.111	0.016	***	0.520	0.504	***
Total	0.000	0.000		0.502	0.500	

Source: authors' elaboration.
Note: Inverse probability weights applied.

Figure 7: Distribution of normalized PCA-based resilience index computed over pooled cross-country sample and for each country.



Source: authors' elaboration.

7.3. Appendix 3 - Balancing panel subsample of farmers

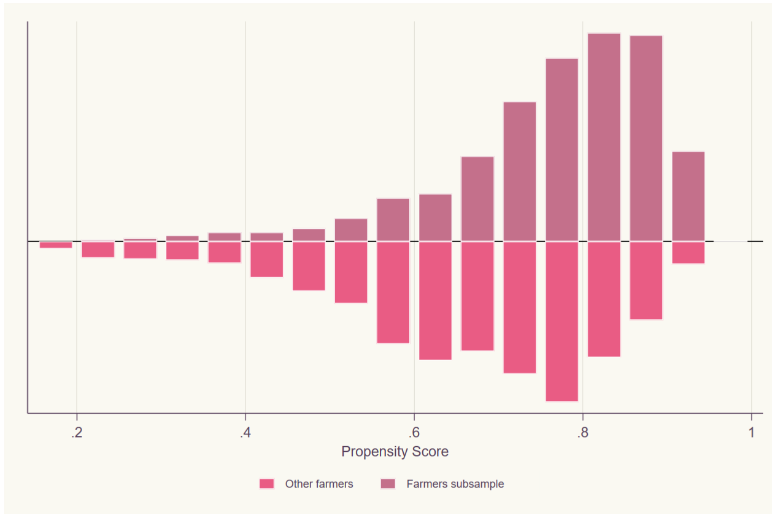
Table 17: Balance test: original rural sample vs. subsample of farming households.

	Original sample	Subsample	T-test
	(1)	(2)	(3)
HH head is female	0.23	0.18	***
Age of HH head	53.32	51.41	***
HH head is married	0.71	0.79	***
Education of HH head	6.67	5.07	***
HH size	6.70	6.77	
N. of female HH members	2.98	3.18	***
Dem. Dep. Ratio	1.03	1.19	***
HH receives food for free	47.19	45.25	**
HH practices agriculture	0.45	1.00	***
Ag. employment	0.04	0.15	***
Land owned (hectares)	0.82	1.14	***
Use of chemicals	0.38	0.39	
Use of mechanized equipment	0.13	0.22	***
Access to electricity	0.44	0.20	***
HH owns dwelling	0.78	0.91	***
HH engages in livestock prod.	0.53	0.73	***
Distance to road (km)	8.52	11.00	***
Herfindahl index of income	0.80	0.72	***
Per capita HH income (2017 USD PPP)	1.87	1.39	***

Source: authors' elaboration.

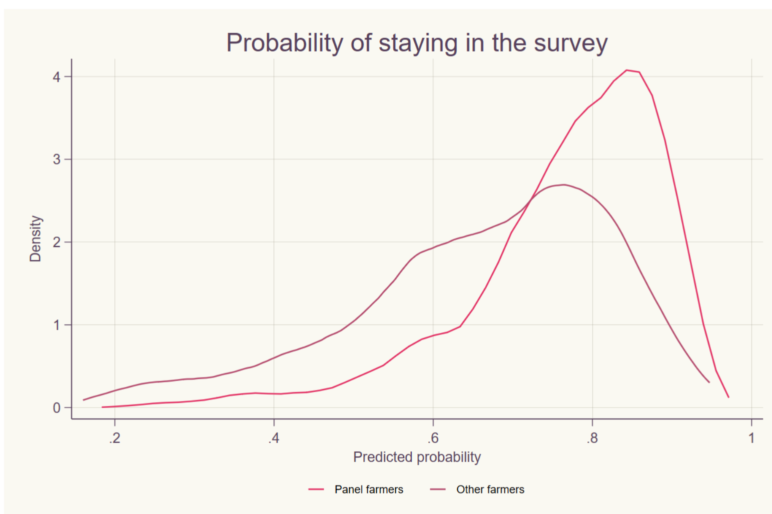
Note: Columns 1 and 2 report the mean value for the original sample of households living in rural areas (Column 1) and the subsample of panel households having some agricultural production in rural areas (Column 2). Column 3 reports the level of significance of the estimates based on a linear regression on the equality of means. Panel sampling weights applied. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 8: Propensity score



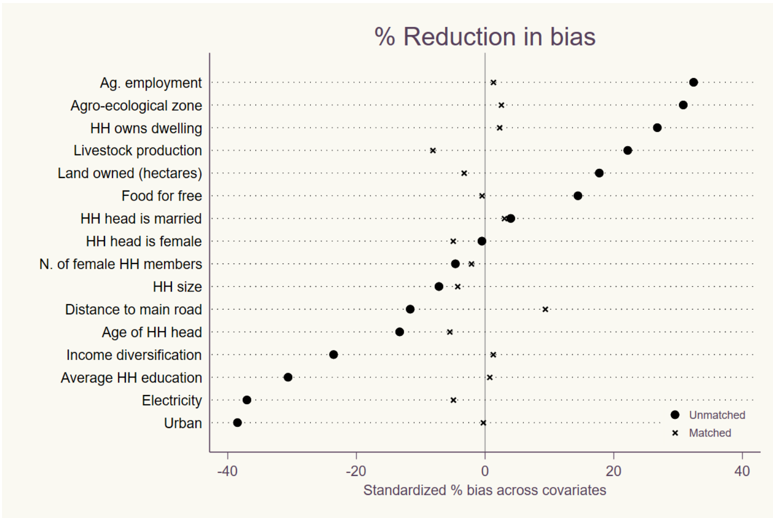
Source: Authors' elaboration.

Figure 9: Common support



Source: Authors' elaboration.

Figure 10: Bias reduction



Source: Authors' elaboration.

7.4. Appendix 4 - Data sources

- LSMS data

  Data title: **Malawi Integrated Panel Household Survey 2010/2011, 2012/2013, 2016/2017**

  Data source: World Bank Microdata Library

  Malawi Integrated Panel Household Survey 2010/2011:

  - Survey ID: MWI_2010_IHS-III_v01_M
  - DOI: https://doi.org/10.48529/w1jq-qh85
  - Accessed on: 21/04/2021

  Malawi Integrated Panel Household Survey 2012/2013:

  - Survey ID: MWI_2010-2019_IHPS_v06_M
  - DOI: https://doi.org/10.48529/q5q1-2a34
  - Accessed on: 21/04/2021

  Malawi Integrated Panel Household Survey 2016/2017:

- Survey ID: MWI_2016_IHS-IV_v04_M
- DOI: https://doi.org/10.48529/g2p9-9r19
- Accessed on: 21/04/2021

Data title: **Nigeria General Household Survey 2012/2013, 2015/2016, 2018/2019**

Data source: World Bank Microdata Library

Nigeria General Household Survey 2012/2013:

- Survey ID: NGA_2012_GHSP-W2_v02_M
- DOI: https://doi.org/10.48529/kxpy-aa72
- Accessed on: 22/04/2021

Nigeria General Household Survey 2015/2016:

- Survey ID: NGA_2015_GHSP-W3_v02_M
- DOI: https://doi.org/10.48529/7xmj-q133
- Accessed on: 21/04/2021

Nigeria General Household Survey 2018/2019:

- Survey ID: NGA_2018_GHSP-W4_v03_M
- DOI: https://doi.org/10.48529/1hgw-dq47
- Accessed on: 21/04/2021

Data title: **Tanzania National Panel Survey 2008/2009, 2010/2011, 2012/2013**

Data source: World Bank Microdata Library

Tanzania National Panel Survey 2008/2009:

- Survey ID: TZA_2008_NPS-R1_v03_M
- DOI: https://doi.org/10.48529/hz8s-3489
- Accessed on: 27/10/2021

Tanzania National Panel Survey 2010/2011:

- Survey ID: TZA_2010_NPS-R2_v03_M

- DOI: https://doi.org/10.48529/jm20-c742
- Accessed on: 21/04/2021

Tanzania National Panel Survey 2012/2013:

- Survey ID: TZA_2012_NPS-R3_v01_M
- DOI: https://doi.org/10.48529/7vqv-5f71
- Accessed on: 21/04/2021

Data license: All datasets are provided as Public Use Files which are available to anyone agreeing to respect a core set of easy-to-meet conditions. These data are made easily accessible because the risk of identifying individual respondents or data providers is considered to be low. The agreements are made with author Jonas Stehl who processed these data.

Data analysis: These data are used for basic household characteristics such as household size or age of household head. Further, food items of the food consumption module were grouped into food groups of the Minimum Dietary Diversity for Women (MDDW).

- RuLIS

  Data title: RuLIS - Rural Livelihoods Information System

  Data source: data retrieved from RULIS

  Data license: CC BY-NC-SA 3.0 IGO

  Citation: FAO. 2018. Rural Livelihoods Information System (RuLIS) - Technical notes on concepts and definitions used for the indicators derived from household surveys. Rome. 68 pp

- NDVI

  Data title: **MOD13A1.061 Terra Vegetation Indices 16-Day Global 500m**

  Data source: NASA LP DAAC at the USGS EROS Center via Google Earth Engine

  Data license: Terms of use: "MODIS data and products acquired through the LP DAAC have no restrictions on subsequent use, sale, or redistribution."

  Citation: DOI: 10.5067/MODIS/MOD13A1.061

Data title: **FAO GAUL: Global Administrative Unit Layers 2015, Second-Level Administrative Units**

Data source: FAO UN via Google Earth Engine

Data license: FAO grants a license to use, download and print the materials contained in the GAUL dataset solely for non-commercial purposes and in accordance with the conditions specified in the data license. To be mentioned: "Source of Administrative boundaries: The Global Administrative Unit Layers (GAUL) dataset, implemented by FAO within the CountrySTAT and Agricultural Market Information System (AMIS) projects"

NDVI data analysis: Yearly average NDVI values per pixel from MODIS were created using Google Earth Engine. The FAO GAUL dataset was used to set national administrative boundaries for the three analysed countries for analysing the values within the countries and exporting raster files per country using the coordinate reference system WGS 84. The yearly images were downloaded and further analysed in QGIS version3.28.11 to calculate average NDVI values for the buffer zones (2km and 5km) of the household locations.

- SPEI

  Data title: **SPEIbase v2.9**: 1) the CRU TS 4.07 dataset, spanning the period between January 1901 to December 2022. 2) Using SPEI package version 1.8.0.

  Data source: CSIC

  Data license: Open Database License (ODbL 1.0 license).

  Data citation: https://doi.org/10.20350/digitalCSIC/15470

  SPEI analysis: The global 0.5° gridded SPEI dataset was downloaded using the 3-months-timescale as a netCDF file. Separate raster files for the study areas, years and months under analysis were computed using QGIS 3.28.11. We considered the different growing seasons for the different countries: For Tanzania we considered the most intense rainy season from March to May. For Nigeria, we considered the different seasons within the country and focused on March to May in the south and May to July in the north of Nigeria. For Malawi we focused on the main months of the growing season, i.e. from January to March. Average SPEI values were calculated for the buffer zones (2km and 5km) of the household locations.

- ACLED

  Data title: **The Armed Conflict Location & Event Data Project (ACLED) data for Africa**

  Data source: The Armed Conflict Location & Event Data Project (ACLED)

  Data license: "If using ACLED data in any way, direct or manipulated, these data must be clearly and prominently acknowledged. Proper acknowledgement includes (1) a footnote with the full citation which includes a link to ACLED's website (see below for examples), (2) in-text citation/acknowledgement, stating that ACLED is the source of these data and that these data are publicly available, and/or (3) clear citation on any and all visuals making use of ACLED data." Also indicate:

  - The date you accessed these data: 31.01.2024
  - Which data you accessed: countries: Malawi, Nigeria, Tanzania; time period: 01/01/2003 - 31/12/2018
  - no manipulations or changes have been made to the original data

  ACLED data analysis: For each year t (t = 2003, ..., 2018), the distance to the closest conflict event j_t from the modified coordinates of each observation $i_{id,wave}$ of the LSMS dataset was calculated. The calculations were performed using R version 4.2.1 and the R package sf. For Nigeria, the projected coordinate reference system UTM 32N (EPSG 32632) was used and for Malawi and Tanzania UTM 36S (EPSG 32736). Based on the distances, we calculated a binary variable whether the observation is within a certain buffer zone of conflict. Calculations were performed per country to avoid calculating the distance to the nearest events across borders of Tanzania and Malawi.

- Afrobarometer

  Data title: **Subnationally geocoded Afrobarometer data**

  Data source: Afrobarometer

  Data license: Afrobarometer data are protected by copyright. Authors of any published work based on Afrobarometer data or papers are required to acknowledge the source, including, where applicable, citations to data sets posted on this website. Please acknowledge the copyright holders in all publications resulting from the use of Afrobarometer data by means of bibliographic citation in this form: Afrobarometer Data, [Country(ies)], [Round(s)], [Year(s)], available at http://www.afrobarometer.org.

Suggested citations: BenYishay, A., Rotberg, R., Wells, J., Lv, Z., Goodman, S., Kovacevic, L., Runfola, D. 2017. Geocoding Afrobarometer Rounds 4 – 6: Methodology Data Quality. AidData. Available online at http://geo.aiddata.org.

Afrobarometer Data, [Malawi, Nigeria, Tanzania], [Rounds 4-6], [Years 2008, 2012/2013, 2014/2015], available at http://www.afrobarometer.org.

Afrobarometer data analysis: The waves 4, 5, and 6 were used for Malawi (2008, 2012, 2014), Nigeria (2008, 2013, 2015), and Tanzania (2008, 2012, 2015). A spatial join was performed with the provided coordinates of the Afrobarometer dataset and the FAO GAUL data on administrative levels 2. Another spatial join was performed with the provided LSMS household coordinates and the FAO GAUL data on administrative levels 2. Based on the identifiers of the subnational levels, the two datasets (Afrobarometer, LSMS) were joined. Data preparation was done using QGIS version 3.22 and R version 4.2.1 using the R package sf. WGS 84 were used as coordinate reference systems.

- Land cover
  Data title: FAOSTAT Land Cover domain
  Data source: FAO Metadata ; FAO link ; GEE Link ; earthmap MODIS combined land cover
  Data license: CC BY-NC-SA 3.0 IGO; https://creativecommons.org/licenses/by-nc-sa/3.0/igo
  Data citation: FAO, 2023. FAOSTAT Land, Inputs and Sustainability, Land Cover http://www.fao.org/faostat/en/data/LC

  Land classification:
  1 – Barren (At least 60% of area is non-vegetated barren (sand, rock, soil) or permanent snow/ice with less than 10% vegetation.);
  2 – Permanent snow and ice;
  3 – Water Bodies;
  9 – Urban and built up lands;
  10 – Dense forest (Tree cover >60% (canopy >2m). This class includes areas with forests dominated by Evergreen Needleleaf; Evergreen Broadleaf; Deciduous Needleleaf; Deciduous Broadleaf; or Mixed types. The class may include tree crops.);
  20 – Open forest (Tree cover 10-60% (canopy >2m). Class includes forest with open (30-60%) and sparse (10-30%));
  25 – Forest / Cropland Mosaics (Mosaics of small-scale cultivation 40-60% with

$>$10% natural tree cover);
27 – Woody Wetlands (Shrub and tree cover $>$10% ($>$1m). Permanently or seasonally inundated);
30 – Natural Herbaceous (Areas dominated by herbaceous annuals ($<$2m) with at least 10% cover. This class includes areas with Dense Herbaceous (cover at least 60%) or Sparse Herbaceous (10-60% cover));
35 – Natural Herbaceous/Croplands Mosaics (Mosaics of small-scale cultivation 40-60% with natural shrub or herbaceous vegetation);
36 – Herbaceous Croplands (Class dominated by herbaceous annuals ($<$2m), with at least 60% cover. Cultivated fraction $>$60%.);
40 – Shrublands (Shrub cover $>$60% (1-2m). This class includes Dense, Sparse and Shrublands / Grasslands Mosaics. The class may include shrub crops.);
50 – Herbaceous Wetlands (Areas dominated by herbaceous annuals ($<$2m) $>$10% cover. Permanently or seasonally inundated).

7.5. Appendix 5 - Variables description

ABS	**Variable name**	**Definition**	**Scale**	**Source**
Access to water	safewater	Household with access to safe drinking water	Community	RuLIS
Access to electricity	electricity	Household with access to electricity	Community	RuLIS
Nearby markets	dist_road	HH Distance in (KMs) to Nearest Major Road	Community	Geocoded LSMS
	AB_infrastructure	How well or badly would you say the current government is handling the following matters, or haven't you heard enough to say: Maintaining roads and bridges?	Community	Afrobarometer
Sanitation	toilet	Household with access to improved sanitation facilities	Community	RuLIS
Water supply	runwater	Household with piped water connections	Community	RuLIS
	AB_nowaterfamily	Over the past year, how often, if ever, have you or anyone in your family: Gone without enough clean water for home use?	Community	Afrobarometer
Safe houses	brickwalls	Household with brick walls	Household	RuLIS
	nondirtfloor	Household with non-dirt floor	Household	RuLIS
	solidroof	Household with a solid roof	Household	RuLIS
Security in the community	AB_safety_home	Over the past year, how often, if ever, have you or anyone in your family: Feared crime in your own home?	Community	Afrobarometer

ASSETS (AST)	**Variable name**	**Definition**	**Scale**	**Source**
TLU	tlu_total	Total tropical livestock unit owned by the household	Household	RuLIS
Agricultural assets	dliv_prod	Households engaged in livestock production activities	Household	RuLIS
	dvaccine	Household with at least one animal vaccinated	Household	RuLIS
	dmechuse	Household using mechanized equipment for farm activities	Household	RuLIS
	dmechown	Household owning mechanized equipment for farm activities	Household	RuLIS
Non-agricultural assets	cellphone	Household owning a mobile phone	Household	RuLIS
	roomsno	Number of rooms in the main dwelling	Household	RuLIS
Land	temp_crops	Land under temporary crops, in hectares	Household	RuLIS
	crop_land	Crop land, in hectares	Household	RuLIS
	temp_fallow	Land temporarily fallow, in hectares	Household	RuLIS
	cultivated	Cultivated land, in hectares	Household	RuLIS
	farm_area	Farm area, in hectares	Household	RuLIS
	landown	Land owned by the household, in hectares	Household	RuLIS
	arable_land	Arable land, in hectares	Household	RuLIS
	arable_land_own	Arable land owned, in hectares	Household	RuLIS
	dland_cul	Household with cultivated land	Household	RuLIS
Agricultural inputs	chemidummy	Household using chemicals	Household	RuLIS
	impseedsdummy	Household using improved seeds	Household	RuLIS
	infertdummy	Household using inorganic fertilizers	Household	RuLIS
	hiredlab	Total n. of working days that hired workers spent in the household's crop production activities in the last 12 months.	Household	RuLIS

SSN	Variable name	Definition	Scale	Source
Public transfers	covsocass	Social assistance coverage	Household	RuLIS
Remittances	covintrem	International remittances coverage	Household	RuLIS
Food for free	anyrec	Received any foods for free / in-kind	Household	HH LSMS
Credit	creditdummy	Households receiving a credit	Household	RuLIS

AC	Variable name	Definition	Scale	Source
	educhead	Number of years of education of the household head	Household	RuLIS
Years of eduaction	educadult	Average number of years of education of the adult members of the household	Household	RuLIS
	educave	Average number of years of education	Household	RuLIS
	educave15_64	Average number of years of education of the household members aged 15-64	Household	RuLIS
	educhigh	Highest number of years of education in the household	Household	RuLIS
	hhi_income	Herfindahl index of income components	Household	RuLIS
Income diversification	ag_part	Type of household (based on income from agriculture)	Household	RuLIS
	p_agrwge	Household with at least one employee in agriculture	Household	RuLIS
Dep. ratio (inverse)	ec_dep_ratio	Economic dependency ratio	Household	RuLIS
	dem_dep_ratio	Demographic dependency ratio	Household	RuLIS
Irrigation	irrigation	Irrigation schemes in the community	Community	RuLIS
	itotlandcul	Irrigated cultivated land, in hectares	Household	RuLIS
	dtrain_ext_crop	Household farms with access to crop production trainings or extensions	Household	RuLIS
Training	dtrain_ext_lvst	Household farms with access to livestock production trainings or extensions	Household	RuLIS
	dtrain_ext_fish	Household farms with access to fishery production trainings or extensions	Household	RuLIS

Absorptive capacity	Variable name	Definition	Scale	Source
Agroecological zones	ssa_aez09	Agro-ecological Zones	Plot	Geocoded LSMS
	twi_mwi	Potential Wetness Index	Plot	Geocoded LSMS
	sq1	Nutrient availability	Plot	Geocoded LSMS
	sq2	Nutrient retention capacity	Plot	Geocoded LSMS
	sq3	Rooting conditions	Plot	Geocoded LSMS
Soil quality	sq4	Oxygen availability to roots	Plot	Geocoded LSMS
	sq5	Excess salts	Plot	Geocoded LSMS
	sq6	Toxicity	Plot	Geocoded LSMS
	sq7	Workability (constraining field management)	Plot	Geocoded LSMS
	af_bio_1	Annual Mean Temperature (degC * 10)	Plot	Geocoded LSMS
	srtm_1k	Elevation (m)	Plot	Geocoded LSMS

Public goods	Variable name	Definition	Scale	Source
% of women in agriculture	femhead	Female-headed household	Household	RuLIS
	m_fem_temp_crops	Land under temporary crops cultivated by females in the household, in hectares	Household	RuLIS
	wfamdays	Number of family labour days in crop production activities, females (>=15 years)	Household	RuLIS
	flabor	Number of female household members aged 15-64	Household	RuLIS
	flaborshare	Share of household female members in working age	Household	RuLIS
	fsize	Number of females in the household	Household	RuLIS
	fem_arable_land_own	Arable land owned by females, in hectares	Household	RuLIS
	fem_landown	Land owned by females, in hectares	Household	RuLIS
Average age of farmers	agehead	Age of household head	Household	RuLIS

Other community level var.	Variable name	Definition	Scale	Source
Corruption	AB_corruption_government	How well or badly would you say the current government is handling the following matters, or haven't you heard enough to say: Fighting corruption in government?	Community	Afrobarometer
Country econ. condition	AB_economic_country	In general, how would you describe: The present economic condition of this country?	Community	Afrobarometer
	AB_economic_past	Looking back, how do you rate economic conditions in this country compared to twelve months ago?	Community	Afrobarometer
	AB_economic_future	Looking ahead, do you expect economic conditions in this country to be better or worse in twelve months time?	Community	Afrobarometer
Freedom of speech	AB_FOS_think	In this country, how free are you: To say what you think?	Community	Afrobarometer
	AB_FOS_politcs	In this country, how often: do people have to be careful of what they say about politics?	Community	Afrobarometer
Democracy	AB_demo_election	On the whole, how would you rate the freeness and fairness of the last national election	Community	Afrobarometer
	AB_demo_opinion	In your opinion how much of a democracy is your country today?	Community	Afrobarometer
	AB_demo_sat	Overall, how satisfied are you with the way democracy works in the country? Are you: Satisfaction with democracy	Community	Afrobarometer
Trust to institutions	AB_trust_officials	In your opinion, how often, in this country: Do officials who commit crimes go unpunished?	Community	Afrobarometer
	AB_trust_ordinary	In your opinion, how often, in this country: Do ordinary people who break the law go unpunished?	Community	Afrobarometer
	AB_trust_police	How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: Police?	Community	Afrobarometer
	AB_trust_president	How much do you trust each of the following, or haven't you heard enough about them to say: The president	Community	Afrobarometer