Does the design matter? Comparing satellite-based indices for insuring pastoralists in Kenya

Technical report prepared for the BASIS Assets and Market Access CRSP

December 31, 2016

Clement Atzberger[1], Michael Carter[2], Francesco Fava[3], Nathaniel Jensen[† 3], Michele Meroni[4], Andrew Mude[3], Quentin Stoeffler[5], Anton Vrieling[6]

Abstract: Index insurance is promoted as a low-cost approach to increasing access to formal insurance products in regions and for individuals that were previously inaccessible for conventional insurance products. These products are now used in multiple nations and by various humanitarian organizations to protect the vulnerable from the sometimes devastating impacts of weather related shocks. Those wishing to support the development or provision of these products generally turn to agronomists, meteorologists and/or remote sensing specialists to help them identify the most accurate and useful index for their products. But, there is no consensus in these communities on which of the many off-the-shelf or pay-for-service indices most accurately track real-world outcomes. Furthermore, household preferences, such as the desire to meet basic needs or an aversion to extremely poor outcomes, make metrics that are commonly used to measure relationships (e.g., the mean error, correlations) less relevant when examining the quality of an insurance product. This study uses economic approaches and the case of the index based livestock insurance (IBLI) product in Kenya to compare the quality of insurance products developed from a variety of satellite -based indices, all of which have either been proposed or are/have been used by insurance or insurance-like products in the region. Although the indices are highly correlated to each other (ρ>0.98), a utility analysis provides insight into how the small differences can lead to larger differences in product quality. In addition, we examine an additional set of indices that aim to predict end of season conditions early in the season, finding that they do so accurately. More generally, this work provides guidance to those working to identify an appropriate index for their product and for index developers in the remote sensing community as they work to improve upon existing products.

[†] Corresponding Author: n.jensen@cgiar.org. [1] Universität für Bodenkultur Wien. [2] UC Davis. [3] International Livestock Research Institute. [4] JRC European Commission. [5] Istanbul Technical University. [6] University of Twente.
Acknowledgements: This research was made possible by generous funding by the Global Action Network on Index Insurance through UC Davis.

Introduction

Index insurance is, by definition, an imperfect type of insurance, as it relies on a common, exogenous index to determine individual payments. The challenge for index insurance design is to offer the best possible protection to households given the available indices that are mostly unaffected by the actions of individual clients, such as rainfall measures from weather stations, remotely sensed measures of vegetation conditions, and the average yields in a region. Protecting households effectively means reducing as much as possible the gap between their actual losses and the insurance payments they receive.

Several sources threaten the value of index insurance, including idiosyncratic risk, index accuracy, product payment schedule and pricing. Indeed, there is a possibility that an individual has a loss but does not receive insurance payments (or insufficient payments) if the index does not "trigger" (or indicate small predicted losses). This means that an insured individual can end up worse-off than she possibly could without insurance coverage, since in some cases she pays insurance premiums in addition to her losses, but does not receive an indemnity payment. The implications of such welfare reducing outcomes are troubling for those promoting index insurance as a tool to fight poverty and spending resources to increase uptake of index insurance among the poor. Low quality index insurance products can not only fail to protect households against adverse shocks, but can increase their vulnerability to shocks, with potentially devastating consequences to physical and human capital. Although there are no studies that explicitly examine the impact of purchasing poor insurance products, there are many studies that highlight the negative impacts of risk and shocks on households (e.g., Carter & Lybbert, 2012; Dercon, 2004; Duryea, Lam, & Levison, 2007; Gong, de Walque, & Dow, 2015; Hoddinott & Kinsey, 2001; Janzen & Carter, 2013; Robinson & Yeh, 2011). One might reasonably expect that purchasing a costly insurance policy that provided little risk protection could leave households less well prepared for shocks (less ex ante risk mitigation) and exacerbate the impact of shocks.

Those few studies that have assessed the coverage that index insurance provides to households often find that insured farmers do continue to face considerable risk (e.g., Barré, Stoeffler, & Carter, 2016; Clarke, Mahul, Rao, & Verma, 2012; Jensen, Barrett, & Mude, 2016).[1] There is also a growing focus on improving the quality of the protection offered by index insurance products. For instance, the initial Index Based Livestock Insurance (IBLI) contracts in Kenya were developed and validated using two sources of household data and statistical methods to minimize basis risk, a level of rigor previously unheard of in the index insurance domain (Chantarat, Mude, Barrett, & Carter, 2013). As another example, Elabed, Bellemare, Carter, and Guirkinger (2013) developed a double-trigger (multiscale) mechanism that could be used to reduce the basis risk associated with an area-yield cotton product in Mali without substantially increasing the cost of offering the product.

[1] Barré, Stoeffler, and Carter (2016) show that the best cotton area-yield product provides valuable protection to farmers in Burkina Faso, but that some design features combined to produce high policy premium rates, which limited its net value for farmers. Clarke, Mahul, Rao, and Verma (2012) found indemnities made by the Weather Based Crop Insurance Scheme in India were poorly correlated to losses (average Pearson correlation was -0.14). Jensen, Barrett, and Mude (2016) found that the Index based Livestock Insurance (IBLI) product in Kenya covered a large share of the covariate risk, but provided limited protection at the individual level due to a high degree of idiosyncratic risk.

Initially indices derived from remote sensing data were generated predominantly by practitioners and researchers, but not always with a sound understanding of the ecological meaning of the index or of required pre-processing steps (Brown, Osgood, & Carriquiry, 2011; de Leeuw et al., 2014). The remote sensing community has recently become more involved in index insurance, bringing advances in data processing, methods for integrating data from multiple sources, and vegetation/crop modeling (e.g., Black et al., 2016; Klisch & Atzberger, 2016; Mann & Small, 2014; Roumiguié et al., 2016; Vrieling et al., 2014, 2016).

This paper is located at the intersection of these economic and remote sensing studies. Generating an index implies several normative decisions regarding the selection of the data source, filtering, smoothing, the temporal and spatial aggregation, and normalization of the raw remote sensing data acquired by satellites. Until now, the study of these parameters has not been related to households' actual losses based on household-level data. This interdisciplinary collaboration is the first attempt to bridge the two literatures from economics and remote sensing, in order to assess the impact of these technical decisions on index insurance quality, measured in terms of household economic wellbeing. In doing so, it assesses the potential of small changes to index processing for the improvement of the design of index insurance products (rather than improvements based on other dimensions such as price), with an objective of guiding future public efforts as well as the design of the growing number of cash transfer and index insurance products which rely on remote sensing technologies.[2]

We do so by studying a set of indices based on the Normalized Difference Vegetation Index (NDVI), applying them to the case of the Index Based Livestock Insurance (IBLI) program in Northern Kenya. The set of indices examined are all either are being used or have been proposed as candidate indices for IBLI or other IBLI-like tools in the region, such as the index used by Kenya's National Drought Management Authority to monitor for drought conditions and to trigger the transfer of Disaster Contingency Funds. Each index is assessed by the degree to which its insurance product improves outcomes for pastoral households in northern Kenya.

The four main indices examined are nearly identical—their correlation ranges between 0.985 and 0.996—over the twelve seasons and across the eleven index regions examined in this research. But, these small differences manifest in variation in the timing and magnitude of indemnity payments when they are used to construct insurance contracts, developing into larger differences in how many households benefit from each index product. Although the majority of the evidence does favor one specific index—the CZ eMODIS index, which we introduce below—there are some discrepancies between analyses and the differences between the indices is quite small. We then examine an additional set of indices that are generated 1-3 months in advance of the first set, making it possible for indemnity payments to be made much earlier than insurance based on the original set of indices. There early indices track full season loss rates as well as the original set, providing a basis for insurance policies that make indemnity payments in advance of coming forage shortages. These findings are broadly consistent with existing remotely sensed research, but also illustrate the importance of including household-level analysis and risk aversion into index analysis.

[2] These products include, for instance, drought-contingent cash transfers in Kenya (HSNP 2), milk producer index insurance in Dominican Republic, livestock protection in Mongolia, or rice index insurance in Tanzania. There is a growing policy interest in designing scalable, integrated social protection systems combining regular transfers and indexed products, especially in Kenya and Ethiopia.

The remaining of the article is organized as follow. Section 2 describes the IBLI project and data. Section 3 presents the satellite indices studied in this paper, and the various design options that are compared. Section 4 introduces the empirical approach and examines the quality of each index using the observed seasons. Section 5 uses simulations to reduce the likely impact of small sample bias on our analysis. Sections 6 and 7 extend the analysis to additional indices and relaxes some of the initial constraints that we placed on the policies. We conclude with a discussion in section 7.

IBLI Project and Data

The index based livestock insurance (IBLI) product first launched in Marsabit, Kenya, in 2010. It has since expanded to five counties in Kenya and one zone in Ethiopia, and has been integrated into Kenya's national social protection program. The product aims to protect households from risk associated with livestock losses due to drought, which is the largest driver of livestock mortality in the area. The IBLI contracts rely on the premise that drought leads to forage depletion, that forage depletion manifests in observable changes is NDVI data, and that forage depletion highly correlates with livestock mortality.

The IBLI contracts provide coverage for twelve month, which are divided into two coverage periods to follow the bi-modal rainfall patterns in the region. The first coverage period, which is meant to provide coverage for the long rainy and following long dry season, starts on March 1st and extends through September 30th. The second period provides coverage for the short rainy and short dry season, extending from October 1st through February 28th. Indemnity payments are potentially made twice each year, according to each coverage period's conditions, as indicated by the index. IBLI contracts are available for purchase twice each year during the two months immediately preceding each coverage period.

Between 2009 and 2015, the IBLI team collected six rounds of a 924 household panel survey in Marsabit. The 2009 survey was collected in October, before IBLI rolled out in Marsabit in January, 2010. Subsequent surveys were collected in October/November 2010, 2011, 2012, 2013, and 2015.[3]

The survey tool included an extensive set of questions on demographic, economic, and social characteristics. The data and code-books are freely available and can be found at https://ibli.ilri.org. This research focuses on the insured risk: seasonal livestock herd mortality rates. These are constructed using the reported herd size at the time of the survey and recall questions concerning the month and details of all livestock intake, offtake, births, slaughter, loss, and deaths. These data are used to construct two seasonal livestock mortality rates for each household in each round of the survey.[4] In cases where the household has zero livestock, their livestock mortality rate is undefined by construction. In addition, we drop observations in which livestock mortality rate that is larger than 100%.[5]

---

[3] No survey was collected in 2014 due to programmatic reasons. Importantly for this work, there were no weather or insurance related reasons for not collecting the 2014 survey.

[4] For consistency throughout the paper, the original IBLI seasonal definitions (March 1st - Sept 30/October 21st - February 28th) are used to define mortality rates.

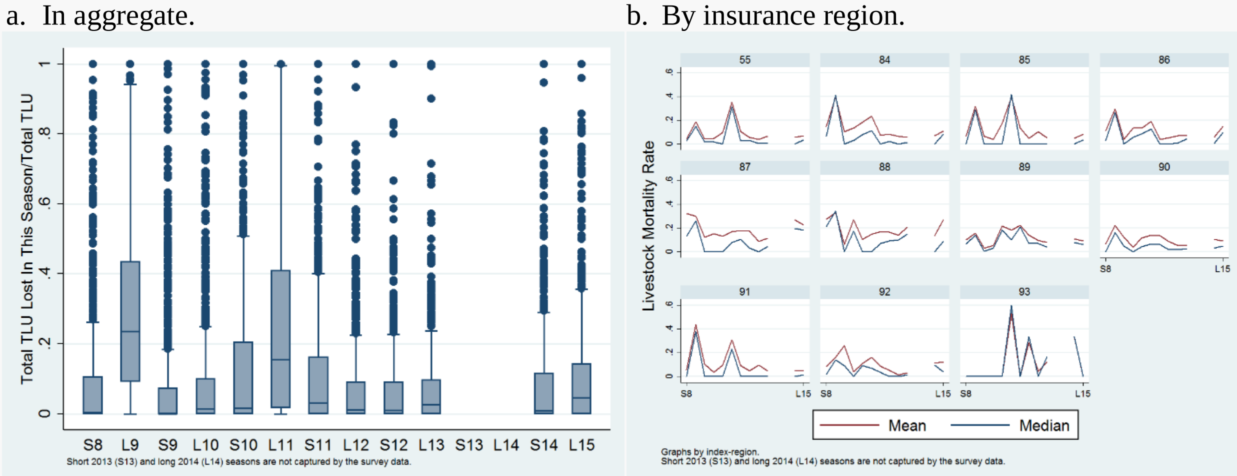
[5] Error in recall can lead to livestock mortality rates that are greater than 100%. Take, for example, a household with one animal in season 1 that purchases two animals in season 2, and then two of their animals die in season 2. If that household misreports the season 2 losses in season 1, the household's reported losses are 200% in season 1.

The attrition rate in the survey is about 4% between rounds. Households that left the survey were replaced when possible.[6] We use the unbalanced panel including replacements, but limit our sample to those for whom there are six or more livestock mortality rate observations over the twelve seasons. This is done at the onset to meet the degrees of freedom requirements of our analysis. The resulting sample is 908 households in Marsabit, all with between 6 and 12 observations per household. As we discuss in more detail in the section on the limitations of this work, we use the livestock mortality data as the benchmark with which to examine accuracy of the remotely sensed indices. But, the livestock mortality data is generated from household survey data, which clearly contains errors itself and those errors may even be larger than the differences between the indices, which are very similar. We acknowledge this shortcoming, but press on because we know of no reason to think that the errors in the survey data systematically favor one index over another.

The observed livestock mortality rates in Marsabit (in aggregate across all the index regions) across each season are depicted in a box plot in **Figure 1.a** and disaggregated by insurance region in **Figure 1.b**. Note the large increase in mortality rates during the long 2009 (L9) and the long 2011 (L11) seasons, when drought occurred in Northern Kenya. Nevertheless, there is a large variation in mortality rate in every season.

Figure 1. Observed livestock mortality rates from the short rain/short dry season of 2008 (S8) through the long rain/long dry season of 2015 (L14):

a. In aggregate.
b. By insurance region.



[6] These attrition figures include households that were missed for one or more rounds, but returned for subsequent rounds.

Satellite Indices and Design Options

Satellite indices for insurance

NDVI time series used in this study were derived from the MODIS (Moderate Resolution Imaging Spectroradiometer) instrument that provides reflectance estimates in the red and near-infrared spectral bands at 250m spatial resolution. MODIS is flown onboard the Terra (2000-present) and the Aqua (2002-present) satellites, both providing daily observations. We used temporally-composited NDVI products, whereby for each pixel the best cloud-free observation is retained during a fixed temporal window of multiple days. To further suppress remaining cloud- and other atmospheric effects, temporal smoothing is applied to NDVI series. Two alternative NDVI products based on MODIS observations are evaluated in this study.

The first is a 10-day constrained maximum value NDVI (i.e. taking into account band quality and view angles) composite product at 250m resolution called eMODIS (Jenkerson, Maiersperger, & Schmidt, 2010), which is currently used in the IBLI program. It is produced by the United States Geological Survey (USGS) from acquisitions by the Terra satellite. Though eMODIS contains six temporally overlapping composites per month, we only used those for day 1-10, 11-20, and 21-last day of each month. Temporal smoothing of the data is performed by USGS using the Swets algorithm; it applies a weighted least-squares regression approach that gives highest weights to local peaks in the NDVI profile, and lowest weights to local valleys (Swets, Reed, Rowland, & Marko, 1999). Use of the Swets-algorithm implies that the data is only release one month after the end of each compositing period. This delayed availability of eMODIS was not accounted for in the present study.

The second NDVI product is generated by the University of Natural Resources and Applied Life Sciences (BOKU) in Vienna, Austria, using as input NASA's 250m-resolution 16-day NDVI composites, both for Terra (MOD13Q1) and Aqua (MYD13Q1). The NASA composites are based on maximum value compositing, but constraining the pixel selection by favouring those recordings with a closer-to-nadir view angle (Huete et al., 2002). Combining Terra and Aqua acquisitions, BOKU's filtering is based on a 8-day product, using the fact that the 16-day composite window for Aqua is shifted by eight days with respect to the Terra composites. The temporal smoothing itself is performed with a modified Whittaker smoother (Atzberger & Eilers, 2011). This B-spline approach incorporates also a 'penalty' criterion regarding the smoothness of the resulting NDVI profile and results in smooth and gap-filled weekly (7-daily) products as well as associated uncertainty information (Klisch & Atzberger, 2016). Contrary to eMODIS, BOKU's filtered data is available directly after the end of each composition period.

Design parameters

Drought indices used for insurance aim to provide a relative measure of environmental conditions with respect to historic conditions. These indices are calculated per IBLI units, defined on the basis of administrative boundaries, adjusted when necessary to better reflect agro-ecological conditions. Even with a single data set, multiple design options exist to translate remote sensing observations into drought indices (de Leeuw et al., 2014). In this study, we selected two main design options that have been used or considered in IBLI **Figure 2**:

1. The first option is based on the original IBLI design as described in Chantarat et al., (2013). For each NDVI composite, pixel-level z-scores are calculated (b1), which express how many standard deviations the NDVI-value is above or below the multi-annual average for that time period (e.g., 1-10 November). The base period for calculating the z-score was October 2002 to September 2015 (13 years) that is available for both NDVI products considered in this study. Subsequently, the z-scores are spatially aggregated for each IBLI-unit (c1). Finally, the temporal series of data is averaged in time to get seasonal values (d1), resulting in a CZ-NDVI value (cumulative value of z-score time series) for the season (e1).
2. The second option is currently implemented in IBLI and aims to first get a proxy of seasonal unit-average primary productivity before z-scoring (Vrieling et al., 2014, 2016). First, NDVI values are spatially averaged at each time period (b2), and subsequently averaged over the season (c2) resulting in a seasonal-average NDVI per unit (d2). Finally, a z-score is calculated comparing this value with the mean and standard deviation based on the same 13 years of data as under option 1; 2002 to 2014 for short rains the seasons, and 2003 to 2015 for long rains, resulting in a ZC-NDVI value (z-score of the cumulative value time series) for each season (e2).

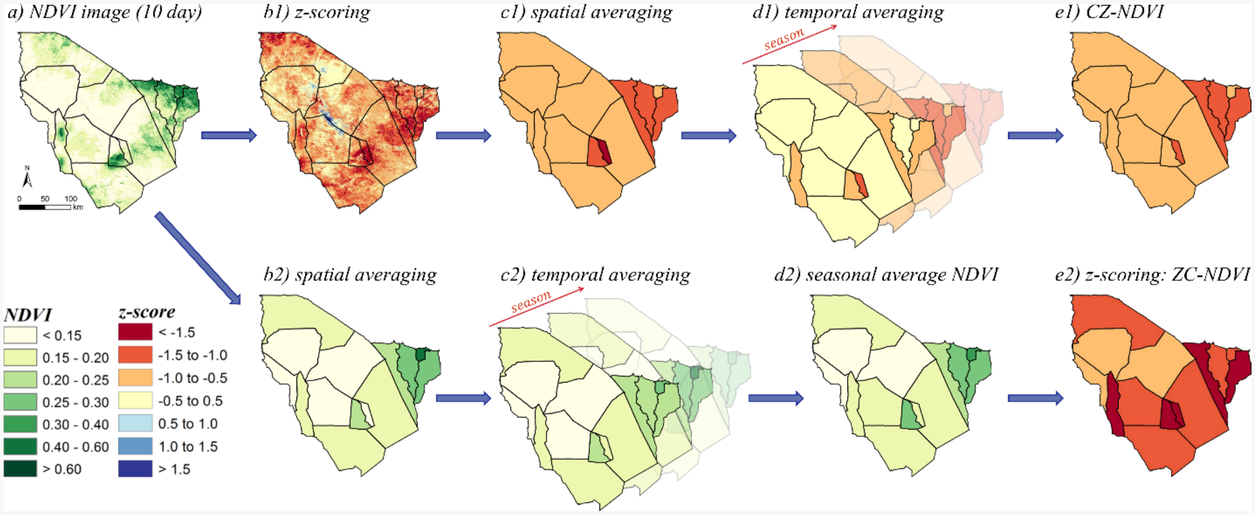
In the spatial aggregation step, we masked out all pixels that had a difference of less than 0.10 NDVI unit between the 95^{th} and 5^{th} percentile for the full NDVI series between October 2002 and September 2015, mostly corresponding to desert areas. We applied each design option to both the eMODIS and the BOKU NDVI series and used for both data sets the same “desert” mask. Note that the above description refers to dekadal eMODIS data. For the 7-daily BOKU data, the same basic ideas were applied. Regarding the seasonal definitions, we examine both the original seasonal definitions of IBLI that cover a full annual cycle (this including the growing periods and the dry ones), a shortened period covering roughly the two vegetation growing periods and a suite of alternative indices that accounts for spatial variations in the season length and timing.

1. *LRLD/SRSD*: The original IBLI definitions, i.e. March-September for LRLD (long rains-long dry), and October-February for SRSD (short rains-short dry);
2. *LR/SR*: Definitions currently used in IBLI (for Marsabit/Borana), i.e. March-June for long rains (LR), October-December for short rains (SR);
3. *phenoFull*: Unit-specific season definitions as in Vrieling et al. (2016) where start and end of season are estimated based on the eMODIS NDVI temporal profiles. Per pixel and season, a parametric double hyperbolic tangent model is fitted to the NDVI time series (Meroni, Verstraete, Rembold, Urbano, & Kayitakire, 2014). Season start/end is determined as the moment when the fitted NDVI exceeds/drops above/below 20% of the amplitude between minimum and maximum NDVI fitted values. The start/end dates are subsequently averaged for all years and all pixels within a unit, while half a (temporal) standard deviation is applied to advance the start and delay the end, allowing for

interannual variability in seasonality. This "phenological analysis" is a common remote sensing approach to obtain an estimate of the time period when vegetation is photosynthetically active;

4. *pheno90p*: These unit-specific definitions are similar to *phenoFull*, but the season end date is brought forward in time up to the date at which the index for the shortened season definitions still explains at least 90 percent of the interannual unit-specific index variability (Vrieling et al., 2016). The advantage of shortening the period with respect to *phenoFull* is that a final season index can be calculated earlier allowing more timely indemnity payments.

Figure 2. Graphical representation of the two design options considered in this study. The figure shows the various spatial units considered within Marsabit County, and is based on real eMODIS data for the 2010 short rainy season. Hence, it illustrates how both design options can lead to different outcomes.



To match the 7-day BOKU dataset to the seasonal definitions for temporal integration (e.g. LR, SR), we integrate all (weekly) images falling within a given season. As BOKU's 7-day dataset is referring exactly to the Monday according to the ISO week date definition, we determine the first and the last Monday for the four different seasonal definitions. The temporal integration comprises all 7-day images from the first to the last Monday of the respective integration period. In addition, the product specific uncertainties are used to calculate spatial and/or temporal weighted averages. Weights in BOKU NDVI are inversely proportional to the specific uncertainty that is calculated for each pixel and time step with the near-real-time (NRT) filtering procedure. For a detailed description of BOKU products see Klisch and Atzberger, (2016).

Observed Seasons: Correlations and Utility Metrics

The pairwise correlations between the remotely sensed indices during the 12 seasons used in the analysis above is greater than 0.98 ($p<0.001$) in every case. To see how well they track livestock loss rates, we start by examining the correlations between average mortality rates and each index overall (row 2, **Table 1**), and within each index region (rows 3-13, **Table 1**). Here we see that although the performances are quite similar between indices, they are very different between index regions. In fact, the livestock mortality rates are positively correlated the indices in two regions, but statistically the correlations in those two regions are indistinguishable from zero as are all of those whose correlations are less than 0.5 in magnitude. We also include the correlation between regional average loss rates and individual loss rates to provide an indication of the levels and variation in covariate risk in the data (column 8, **Table 1**).

Table 1. Inter-annual variability in livestock mortality that is captured by each index (2009-2015). The correlation coefficient is calculated in all cases from 12 pairs of index vs. mortality. The number (obs) of individual household panels entering in the calculation of average mortality are also indicated (N). In the last column, we report the correlation between the average mortalities and the individual loss rates

Region (N)	Obs	N	CZ eMODIS	CZ BOKU	ZC eMODIS	ZC BOKU	Individual Mortality
All	908	12	-0.41	-0.40	-0.40	-0.40	0.35
Uran	43	12	-0.49	-0.47	-0.51	-0.49	0.56
Central Marsabit	58	12	-0.85	-0.85	-0.84	-0.85	0.43
Gadamoji	55	12	-0.72	-0.70	-0.72	-0.74	0.48
Laisamis	113	12	-0.62	-0.61	-0.60	-0.62	0.40
Loiyangalani	64	12	0.04	0.03	0.01	0.01	0.27
Mt. Kulal	59	12	0.07	0.10	0.01	0.02	0.36
Kargi	123	12	-0.23	-0.22	-0.17	-0.21	0.34
Maikona	159	12	-0.45	-0.45	-0.47	-0.50	0.31
Turbi	215	12	-0.51	-0.50	-0.53	-0.52	0.56
Dukana	16	12	-0.31	-0.35	-0.28	-0.28	0.39
North Horr	3	12	-0.43	-0.41	-0.42	-0.42	0.80

In **Table 2**, we calculate for each region and season the average livestock mortality and regress those rates onto the index, assuming a linear (columns 2-5, **Table 2**) and a third order polynomial (columns 6-9, **Table 2**) relationship. The coefficient of determination (R^2) and root mean squared error (RMSE) for each regression provide intuitive measures of the variance in mortality rate that is correctly captured by each index. We also include three additional metrics, mean error, mean over prediction, and mean under prediction.[7] Mean under prediction provides a simple metric of downside risk while mean over prediction

[7] The error and over/under statistics are calculated as follows. First, the coefficient estimates from the regressions that produces the R^2 statistics are used to predict the average livestock mortality rates for each region from each index. The error is the difference between the observed and predicted values. The mean error is the simple average of those errors. The mean over (under) prediction is the average error, conditional on the error being positive (negative).

is a measure of importance for the viability of insurance providers. By all these metrics, there is no discernable difference between the indices.

To focus on risk associated with bad seasons, we repeat the regression (row 6-11, **Table 2**), only including those seasons in which the average livestock mortality is above its 12-season average. This restriction reduces the number of index-region-season observations from 131 to 83, although the number of regions remains that same. Notice that the indices are able to explain much more of the observed variation in losses during poor seasons than they were overall. That is, all the indices track losses better during poor seasons than during average or good season. Once again, the indices' performance is nearly identical.

Table 2. Relation between season- and region- specific livestock mortality rates and the corresponding remote sensing index across all sub-counties of Marsabit and the entire data set (2009-2015). Reported are the coefficient of determination and average differences in predicted and observed outcomes

	Linear [$y = f(index)$]				3rd order poly. [$y = f(index, index^2, index^3)$]			
	CZ		ZC		CZ		ZC	
VARIABLES	eMODIS	BOKU	eMODIS	BOKU	eMODIS	BOKU	eMODIS	BOKU
Full Panel								
Observations	131	131	131	131	131	131	131	131
R^2	0.265	0.256	0.259	0.262	0.377	0.349	0.367	0.356
RMSE	0.0892	0.0897	0.0895	0.0893	0.0828	0.0847	0.0835	0.0842
Mean error	0.0629	0.0632	0.0632	0.0630	0.0553	0.0568	0.0562	0.0570
Mean over	0.0542	0.0524	0.0524	0.0510	0.0458	0.0477	0.0490	0.0472
Mean under	0.0749	0.0797	0.0797	0.0826	0.0696	0.0702	0.0657	0.0718
Restricted Panel								
Observations	82	82	82	82	82	82	82	82
R^2	0.6890	0.6890	0.6890	0.6890	0.7030	0.6990	0.6970	0.6970
RMSE	0.0258	0.0258	0.0258	0.0258	0.0256	0.0258	0.0259	0.0259
Mean error	0.0196	0.0196	0.0196	0.0196	0.0194	0.0196	0.0196	0.0196
Mean over	0.0167	0.0167	0.0171	0.0167	0.0176	0.0179	0.0178	0.0178
Mean under	0.0236	0.0236	0.0230	0.0236	0.0215	0.0217	0.0217	0.0217

Note: The regression includes index-region dummy variables to focus on within, rather than between, region variation.

The above analysis provide some evidence as to the Overall, the BOKU index more accurately captures the variation in livestock mortality rates during the survey seasons. These gains result in lower mean under-predictions of losses, which could easily manifest in lower un-indemnified losses, and lower over-predictions of losses. When the data is restricted to only the below average seasons, BOKU's advantage falls but it still remains the more accurate product. The differences between CZ and ZC products are more ambiguous, although the most accurate index is CZ BOKU.

Utility metrics

Although the above analysis provides evidence that the performance of the four indices is quite similar, they may not identically meet the preferences of households. A utility framework allows us to integrate risk aversion–placing greater weight on poor outcomes than good outcomes, *ceteris paribus*–into the quality measures. For example, given two income schemes with identical means, one with periodically high and low incomes and the second with a constant income at the mean, a risk adverse individual would prefer the scheme with no variation that the one with the high variation. Indeed, risk aversion is one of the key tenets of demand for risk reducing financial tools, such as insurance.

We use a simple constant relative risk aversion (CRRA) utility function with net livestock survival rate (*S*) of household *i* from region *d* in period *t*, to examine the impact of purchasing full insurance coverage from each type of insurance policy (equation 1). Alpha represents the household's level of risk aversion and is often found to be in the range of greater than zero but less than three (Saha, Shumway, & Talpaz, 1994). In the case of no insurance, the survival rates are observed or simulated. In the insured case, each net survival rate is calculated as the base survival rate, less the premium rate, plus the indemnity rate (if any). To address observations in which the produced net survival rate is less than zero–the household has a survival rate that is lower than the premium rate, we add 101% of the premium rate to all outcomes before entering them into the utility functions.[8] The result is that S_{idt} is always greater than zero.

$$U(S_{idt}) = \begin{Bmatrix} \dfrac{S_{idt}^{(1-alpha)}}{(1-alpha)} & if\ alpha \neq 1 \\ \ln(S_{idt}) & if\ alpha = 1 \end{Bmatrix} \tag{1}$$

To determine premium and indemnity payments, we develop an insurance policy for each index in each region. Similar to the IBLI policies being sold, the strike for each index product is set at its within-region 20th percentile, so that each product is expected to make indemnity payments every fifth season. The indemnity function draws directly from the IBLI indemnity function (equation 2), but is scaled by δ_d so that all products and regions have identical actuarially fair rates.[9] That rate is set to be equivalent to the benchmark of a sample-level actuarially fair premium rate of a perfect loss-indemnifying insurance product that covers all individual losses beyond 20%. By setting premium rates to be equal, the analysis avoids conflating the impact of the magnitude of gross transfers with the precision of the index defined inter-temporal reallocation of funds. In essence, equating premium rates means that that each index has the same amount of funds to distribute across the seasons; an index is judged by how and when it does so.

$$indemnity_{dt} = \begin{Bmatrix} \delta_d \dfrac{strike_d - index_{dt}}{strike_d - exit_d} & if\ index_{dt} < \ strike_d \\ 0 & if\ index_{dt} \geq strike_d \end{Bmatrix} \tag{2}$$

[8] As will be discussed below, the premium rates are set at 0.0504 so that 0.0509 is added to S_{idt} in all of the following analysis.

[9] The δ_d scaling factor is actually the composite of a weight used in the actual IBLI indemnity function to reflect that the long rain/long dry season is 40% longer than the short rain/short dry season, and the factor developed for this research to equate the actuarially fair premium rates across products.

Each household's expected utility is then computed to determine if the household is better or worse off from purchasing insurance. We use two basic measures to compare the indices. The first examines the ratio of households that are better off with the actuarially fairly priced insurance than without it, while allowing levels of risk aversion to vary between 0.1 and 3. The second is a comparison of the distribution of reservation premium rates, which is equal to the rate at which a household is indifferent to purchasing full insurance and not purchasing insurance at all, while holding level of risk aversion constant. It illustrates the monetary value of the product. Both approaches weigh all households equally, whereas other common approaches, such as comparing the mean reservation rate, are weighted by the relative magnitude of each product's cost or benefit to a household. We believe that our approach of equally weighting households, irrespective of the magnitude of the cost/benefit, is more consistent with the ordinal nature of utility functions and better reflects that calculus by which policies are made.

Results

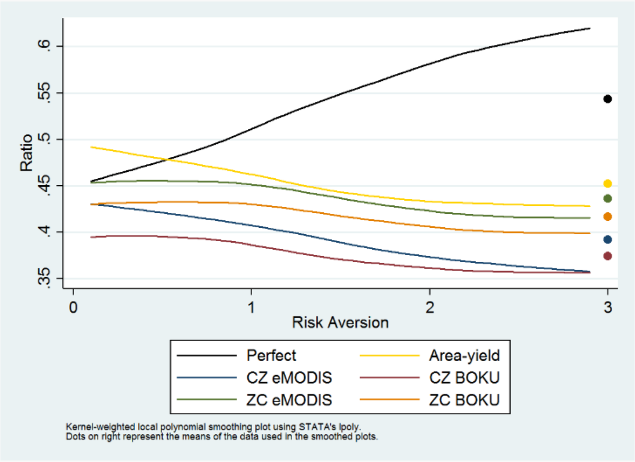
We begin by comparing the impacts of the four insurance products during the observed seasons. An insurance product that perfectly insures all losses beyond 20% is used as a benchmark. The sample-level constant actuarially fair premium rate of the perfect product is then calculated. For Marsabit, that rate is 5.04% of the insured value of the livestock. Index insurance contracts are then constructed for each of the four NDVI indices and for an area-yield product.[10] A strike is set within each index region and for each product at the 20th percentile of the twelve seasons with survey data. In this case, each index insurance product will make indemnity payments in two of the observed twelve seasons in each region. Similar to the IBLI protocol, the exit value is set to the lowest observed index value–the worst observed season. As discussed above, the indemnity payments reflect those used by the IBLI product, but are adjusted by a parameter (δ_d) so that the actuarially fair rate for each region is equal to 5.04% of the insured value (equation 2).

First, the products are compared at identical and actuarially fair premium rates across levels of risk aversion (**Figure 3**). As the level of risk aversion increases, so does the ratio of the sample that benefits from the perfect product.[11] Conversely, the index products fair worse as the level of risk aversion increases. The decrease in the benefits of index insurance coverage as risk aversion increases is a feature of index products in heterogeneous populations; as risk aversion increases, households place more weight on the outcome of high livestock mortality in a season without a high indemnity payment. Note, in this case, the result is due to idiosyncratic risk, not imperfections in the index. Even the area-yield product suffers from this effect. Of the NDVI-index products, the ZC-eMODIS product benefits the greatest ratio of households on average, but the differences between the four indices are always less than ten points.

[10] Although area-survival rate is a more accurate term in this context, we use area-yield because it is more familiar.

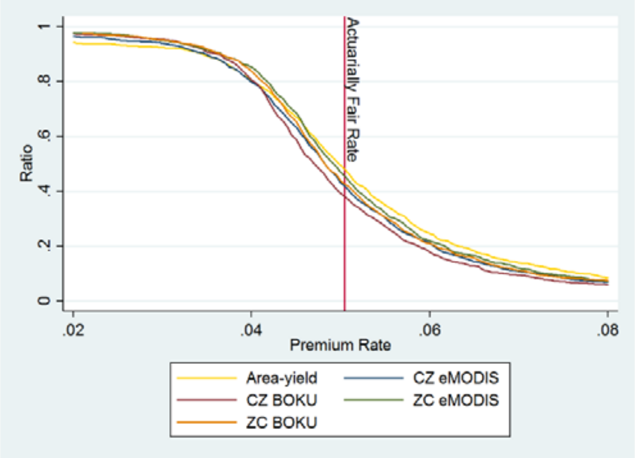
[11] At low levels of risk aversion, the area-yield contracts improve outcomes for more households than does the "Perfect" contracts. This is because the premium levels for the perfect contract is set at the sample-level. Thus, just as is true for most insurance contracts, for any individual household the total indemnities paid can be quite different from the total premiums paid. The same is not true for the index insurance products.

Figure 3. Ratio of sample for whom each actuarially fairly priced insurance product increases expected utility across levels of risk aversion.



Second, setting the level of risk aversion to two (alpha=2.0), we examine the ratio of households that are better off at various premium rates. **Figure 4** illustrates the proportion of the sample for whom their reservation premium rate–the rate at which their utility with insurance is equal to their utility without insurance–is greater than the premium rate on the x-axis.[12] For example, at the actuarially fair premium rate of 5.04% of the insured value, 45% of households are better off with the ZC-eMODIS product than without it, while at the same premium rate 38% of households are better off with the CZ-BOKU product than without it. When the premium reaches 6%, about 20% above the actuarially fair rate, the percentage of households better off with insurance drops off to 22% with the ZC-eMODIS product and 18% with the CZ-BOKU product.

Figure 4. Ratio of the sample for whom each product increases expected utility across premium levels, holding alpha at two.



[12] Note that **Figure 4** intersects **Figure 3** perpendicularly at the actuarially fair premium rate. The localized smoothing performed for **Figure 3** is the reason for apparent differences where they meet.

Setting both the level of risk aversion and the premium rate, allows us to examine the means and their confidence intervals for a specific scenario. For example, at the actuarially fair premium rate (5.04%) and with alpha equal to two, the means suggest that the number of households helped by the ZC products are higher than those helped by the CZ products (**Table 3**). χ^2 tests show that the ZC products are statistically better than the CZ products by this metric. Although the ZC eMODIS product has the highest benefit rate, the rates are only 4% higher than ZC BOKU and the difference is not statistically significant.

Table 3. Ratio of households better off with each actuarially fair product than without it. (Alpha=2)

		H_0: # of households helped is better than the index in the column heading. Person χ^2 (p-value)		
Index	Mean	CZ BOKU	CZ eMODIS	ZC BOKU
CZ eMODIS	0.365	0.12 (0.732)	-	-
CZ BOKU	0.357	-	-	-
ZC eMODIS	0.416	6.77 (0.009)	5.11 (0.024)	0.52 (0.474)
ZC BOKU	0.400	3.56 (0.059)	2.39 (0.122)	-

Overall, the utility approach has allowed us to identify how very small differences in the indices could result in larger differences in the quality of the indices once we take into account risk aversion and model the average outcomes over all 12 seasons. These tests have shown that the ZC products perform better than the CZ products and that differences between the eMODIS and BOKU products are mostly negligible. In addition, this analysis inadvertently illustrated the relative importance of household risk aversion, premium loadings(subsidies) and index filtering. It appears that fairly small changes to premiums can result in extremely large changes in the ratio of households that benefit from a products and that are moderate gains to be had by choosing the appropriate filtering process. In this case, the different filtering processes have different costs associated with them, so policy makers would be wise to weight the dynamics between cost and accuracy carefully and with their objectives in mind.

Simulations

Methods

Although the analysis of observed index-mortality rate data includes about 10,000 observations of household level livestock mortality rate (in Kenya), the index values are time invariant for each season within each region. The result is only 132 (11 regions X 12 seasons) index observations with which to compare indices. Of particular importance, a test of the distribution of index values during the periods with household survey data (2009-2015, N=132) and the periods without survey data (2001-2009, N=154) rejects the null hypothesis of identical means for all four indices with t-statistics at or above 2.43 in every case. The conditions, as measured by the indices, are worse during survey seasons than during non-survey seasons. Such sampling bias will bias our conclusions. In addition, it is unlikely that the indices are independent between regions within a single season or across time within a specific region. Such prospective spatial and temporal correlation could further reduce information, thus increasing the risk of small-sample bias in our estimates.

We use simulations to better understand each NDVI-based index insurance product in this context and to explore the differences between indices. Our simulation approach allows the analysis to draw from a longer time series of NDVI data than is possible when matching indices to survey data, which should be more representative of each indices' the true distribution, and introduces modeled stochasticity to better understand the relationship between a particular NDVI index value and the outcomes that a household experiences.

The simulation process is as follows.

1. The parameters relating household livestock mortality to index values are estimated using a flexible function of observed survey data (2009-2015).
2. Estimate the distribution parameters for each index using the full set of index data (2002-2015).
3. Index values are drawn from the region specific distribution described by the parameters estimated in (2).
4. A livestock mortality rate is drawn from each household's season-specific distribution of livestock mortality rate, which is a function of the parameters estimates in (1) and the index draws from (3).
5. Comparable area-yield and NDVI-index insurance products are generated within each region.
6. We then examine the ratio of the simulated sample that are better off with each insurance product than without it.
7. Steps 3-6 are repeated 500 times.
8. The distribution of outcomes under each index insurance product are then compared

Each step is discussed in greater detail in the following results section.

Results

To provide the reader with greater detail on the simulation process, we will proceed step by step though the eight steps briefly described in the methods section above, providing greater detail, justification, and/or the outcome of statistical tests when applicable.

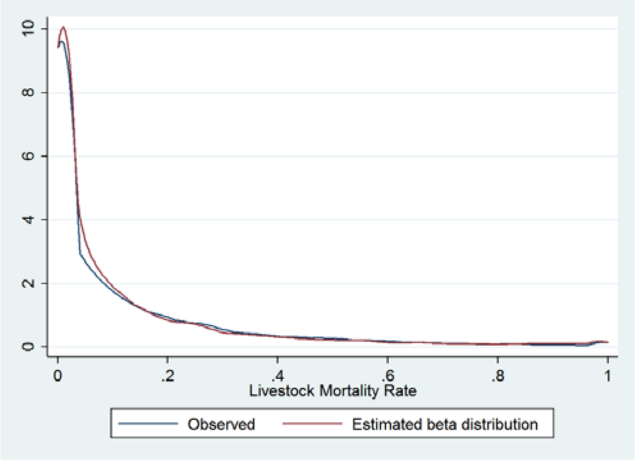
1. The parameters relating household livestock mortality to index values are estimated using a flexible function of observed survey data (2009-2015).

Our simulations are based on the assumption that a household's livestock mortality rate can be accurately modeled as a random draw from a household-specific distribution of livestock mortality rates, and that each household's distribution of livestock mortality rate has moments that are a function of household characteristics and the index. We use the beta distribution, which had the domain [0,10 and is often used to model crops, to model each household's distribution of livestock mortality rate (e.g., Tack, 2013). The beta distribution can be described by two parameters (α, β), which are functions of the mean (μ) and standard deviation (σ) of a variable (equation 3).

$$
\begin{aligned}
LivestockMortality_{idt} &\sim Beta(\alpha_{idt}, \beta_{idt}) \qquad (3)\\
\alpha_{idt} &= -\frac{\mu_{idt}(\sigma_{idt}^2 - \mu_{idt}^2 - \mu_{idt})}{\sigma_{idt}^2}\\
\beta_{idt} &= \frac{(\mu_{idt} - 1)(\sigma_{idt}^2 - \mu_{idt}^2 - \mu_{idt})}{\sigma_{idt}^2}\\
\mu_{idt} &= E[LivestockMortality_{idt}]\\
\sigma_{idt}^2 &= E\left[LivestockMorality_{idt}^2\right]
\end{aligned}
$$

Figure 5 illustrates the probability distribution of livestock mortality rates observed in the survey data and the distribution of livestock mortality rates generated by taking replacing each observed rate with a random draw from the beta distribution with household-level moments estimated directly from the observed data. We proceed assuming that livestock mortality rates are distributed according to the beta distribution.

Figure 5. Kernel density estimation of observed livestock mortality rate and the Beta probability distribution generated using moments estimated from the observed data.



We assume that the household-season specific beta distribution parameters can be described as a second order polynomial of the index in their region (d) in period (t), household fixed effects, and their interactions (equation 4).

$$
\begin{aligned}
\mu_{idt} &= \delta^0 + \delta^{1,k} index_{dt}^k + \delta_i^{2,k} + \delta_i^{3,k} index_{dt}^k + \varepsilon_{idt};\ \varepsilon_{idt} \sim N(0,1) \qquad (4)\\
\sigma_{idt}^2 &= \varepsilon_{idt}^2\\
\sigma_{idt}^2 &= \gamma^0 + \gamma^{1,k} index_{dt}^k + \gamma_i^2 + \gamma_i^{3,k} index_{dt}^k + \vartheta_{idt};\ \vartheta_{idt} \sim N(0,1)\\
k &\in [1,2]
\end{aligned}
$$

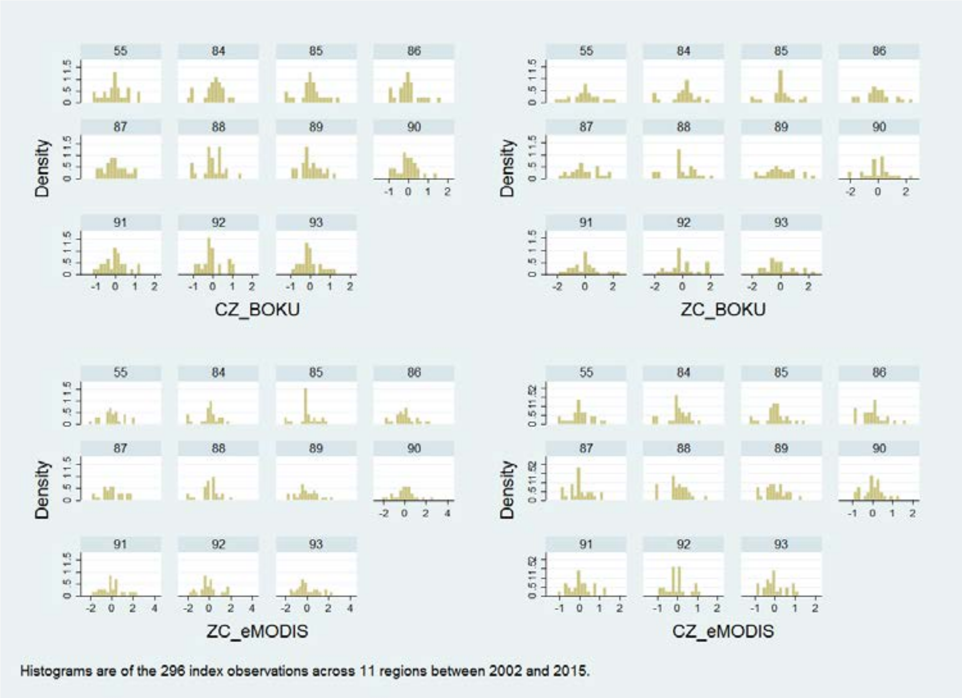
Observed livestock mortality and index data (2009-2015) are used to estimate the parameters (δ^0, $\delta^{1,k}$, $\delta_i^{2,}$ $\delta_i^{3,k}$, γ^0, $\gamma^{1,k}$, γ_i^2, $\gamma_i^{3,k}$) that relate the indices to the distribution parameters. Notice that this procedure estimates three household-specific parameters and three common parameters for each moment. These distribution parameters are then used in step (4) to simulate livestock mortality rates, conditional on draws from the index distributions.

2. Estimate the distribution parameters for each index using the full set of index data (2002-2015).

We observe 27 seasons of NDVI indices for each index in each of the eleven index regions between 2002 and 2015. Although the indices are normalized by design, we begin by testing for systematic differences between LRLD/SRSD seasons and for normality in their probability distribution. A Kolmogorov-Smirnov test for equality of index values between SRSD and LRLD seasons within each region fails to reject the null hypothesis that the index distributions are the same in all 11 observed regions for all four indices. Thus we proceed by pooling the SRSD and LRLD seasons within each region.

Testing for normality within the regions does not provide strong guidance; three of the forty-four (6.8%) Kolmogorov-Smirnov tests reject normality, but the tests do not have a great deal of power within regions because of their small sample size (N=27). **Figure 6** illustrates the distribution of all four indices within each of the eleven Kenyan regions. Visually there are large deviations from what we would expect from a normal distribution; especially troublesome for the purposes of modeling weather shocks are the relative clustering of observations near minimum values, which is where indemnity payments would be made. But, because we do not have sufficient observations to develop an empirically based distribution for each region, we assume normality.

Figure 6. Histograms the four indices in each of the regions.



Thus, we assume that the SRSD and LRLD seasons can be pooled within each region and that the indices are normally distributed within each region. The mean and standard deviation for each index in each region is estimated using the 27 seasons of observed index data (equation 5).

$$
\begin{aligned}
index_{dt} &\sim N(\mu_d, \sigma_d^2) \\
\hat{u}_d &= \frac{1}{27}\sum_{t=1}^{27} index_{dt} \\
\hat{\sigma}_d^2 &= \frac{1}{26}\sum_{t=1}^{27}(index_{dt} - \hat{\mu}_d)^2 \\
t &\in [1,2,\dots,27]
\end{aligned}
\tag{5}
$$

3. An index value is drawn from its distribution described by the parameters estimated in (2).

An index value is randomly drawn from the normal distribution described in equation (5) for twelve seasons in each region. We use the form $\underline{index_{dt}}$ to indicate the index value drawn for region d in period t. We simulate a twelve-season data set, rather than single season or a very large number of seasons, in order to simulate a sample that has properties similar to the original sample within every simulation. Multiple simulation provide information on how our twelve-season sample could have been different.

4. A livestock mortality rate is drawn from each household's beta distribution of livestock mortality rate, each of which is a function of the parameters estimates in (1) and the index draws from (3).

The new index values and parameter estimates from equation (4) are then used to predict a household-period-specific expected livestock mortality rate mean and variance (equation 6). Those are then used to define household-period specific beta distributions of livestock mortality rate, from which a mortality rate draw ($\underline{LivestockMortality_{idt}}$) is made.

$$
\begin{aligned}
\hat{u}_{idt} &= \hat{\delta}^0 + \hat{\delta}^{1,k}\underline{index_{dt}^k} + \hat{\delta}_i^2 + \hat{\delta}_i^{3,k}\underline{index_{dt}^k} \\
\hat{\sigma}_{idt}^2 &= \hat{\gamma}^0 + \hat{\gamma}^{1,k}\underline{index_{dt}^k} + \hat{\gamma}_i^{2,k} + \hat{\gamma}_i^{3,k}\underline{index_{dt}^k} \\
\underline{LivestockMortality_{idt}} &\sim Beta\left(\hat{\alpha}_{idt}, \hat{\beta}_{idt} \,\middle|\, \underline{index_{dt}}, \hat{\delta}, \hat{\gamma}\right)
\end{aligned}
\tag{6}
$$

5. NDVI-index insurance products are generated within each region.

The index products for each simulation are generated using a process that is similar to those used to generate the index-products using from the observed data. Importantly, the strikes are reset in each simulation to ensure a 20% strike rate and the indemnities are scaled to maintain a 5.04% premium rate.

6. We then examine the ratio of the simulated sample that are better off with each insurance product than without it.

Net outcomes are then calculated and each household's twelve season expected utility is estimated for each product and without insurance. Each household's expected utility with insurance is compared against its expected utility without insurance to determine if the household is made better or worse off from the insurance coverage. We then calculate the sample-level ratio of households that are better of with insurance than without it.

7. Steps 3-6 are repeated 500 times.

It is important to note that the seasons captured in the Kenyan IBLI household survey were especially dry, which resulted in two of the largest herd die-offs in recent history. Across all the four indices, the mean index values are lower during the survey seasons than during non-survey seasons (t-statistic ranges from 2.43 to 3.37). The simulated seasons generally reflect the less severe distribution of outcomes represented by the full 27 seasons of index values, rather than the more severe 12 seasons that are observed in the survey data and used in the section on observed data. A reduction in the expected incidence of extreme values results in fairly large differences between the simulated and observed outcomes.

8. The distribution of outcomes under each index insurance product are then compared.

The simulated results are presented in the same manner as the results generated form the observed values. In this case, the reported ratios are the mean outcomes from across the simulations. This time, we can use the multiple simulated responses to also learn about the precision of the mean estimates, which are reported using standard errors. **Figure 7** illustrates the ratio of households better off with each insurance product than without. The mean outcomes of the four products are nearly identical.

Figure 7. The ratio of households better off with each insurance product than without, across levels of risk aversion.

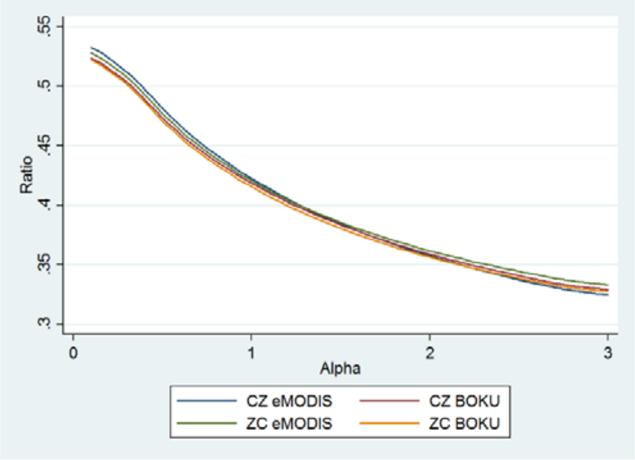
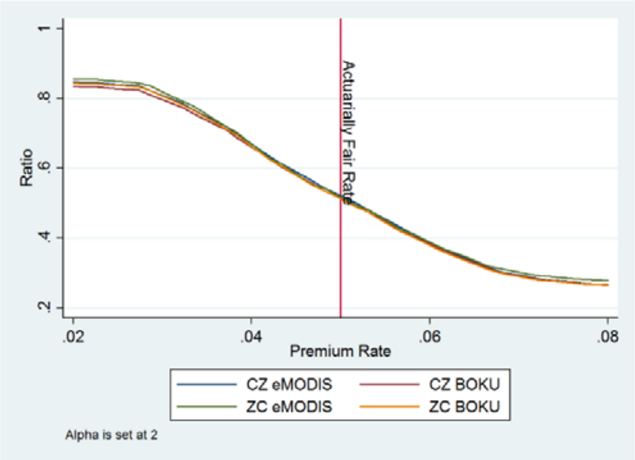


Figure 8 illustrates how the ratio of households that benefit from each product change across premium rates, holding the level of risk aversion at two. Once again, we see that there is very little difference in how well the products perform.

Figure 8. The ratio of households that prefer insurance to no insurance as the premium rate increases.



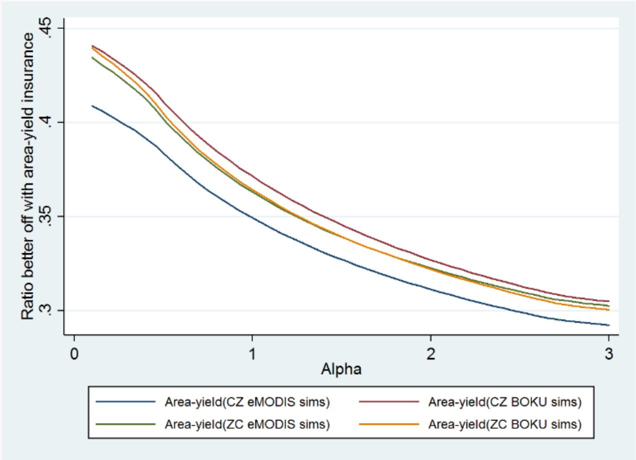
Similar to **Table 3**, **Table 4** examines the ratio of individuals that are better off with coverage from each product. The premium rates are set to 5.04% and alpha is set at two. In this case, each observation represents the mean ratio from each simulation and we compare the distribution of ratios over the simulations using a t-test. This analysis finds that the product based on the CZ eMODIS index benefits the most households, but the differences are quite small. On average, the CZ eMODIS product benefits 0.6% more households than the second best, the ZC eMODIS product, which is equal to about 5 households in our sample of 908. The difference between the worst and best index is equal to about 10 households.

Table 4. Ratio of households better off with each actuarially fair product than without it. (Alpha=2)

		H_0: Ratio of households helped is better than the index in the column heading. t-statistic (p-value)		
Index	Mean	ZC BOKU	CZ BOKU	ZC eMODIS
CZ eMODIS	0.495	4.41 (<0.001)	2.83 (0.003)	2.39 (0.009)
CZ BOKU	0.488	1.52 (0.065)		
ZC eMODIS	0.489	2.03 (0.022)	0.489 (0.313)	
ZC BOKU	0.484			

One issue with the above simulated analysts is that for each simulation round, we are actually generating four index-specific livestock mortality rates. Specifically, each index draw is used by equation (6) to generate a household livestock mortality rate. The result is that the distribution of simulated losses within a particular simulation are not the same between CZ-eMODIS, CZ-BOKU, ZC-eMODIS, and ZC-BOKU. And, these differences can be systematic because the index distribution parameters from equation (5) and coefficient estimates in equation (6) rightly vary across indices. The result is that there is the potential for variation in the simulated distribution of losses to be the driving factor in the results presented here, rather than variation in the precision of each index, as we would like. To illustrate this point, **Figure 9** includes the ratio of households better off under actuarially fair area-yield products developed from the index-specific simulated household mortality rates used above. There are clearly differences in the distribution of losses simulated for each product. For example, the ratio of households that benefit from coverage from an actuarially fair area-yield contract are lower for the CZ-eMODIS distributions across all levels of risk aversion. Another way to say this is that covariate risk plays a relatively smaller role in the CZ-eMODIS simulations than in the distributions associated with the simulations for the other three indices. We should expect such findings to have implications for the relative magnitude of basis risk and thus product quality. To be clear, this variation between indices shown in **Figure 9** is a figment of the simulation process and does not reflect index quality.

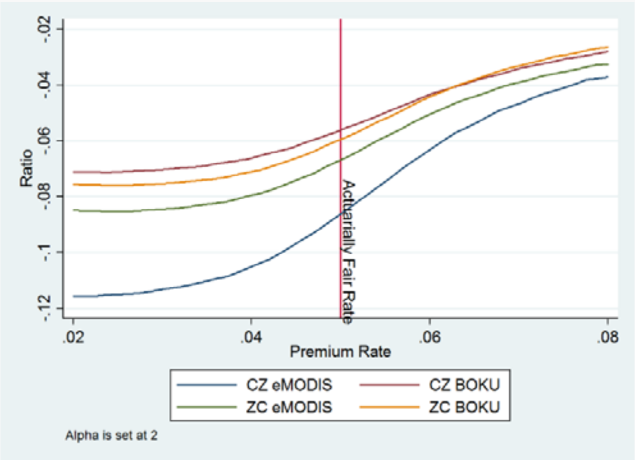
Figure 9. Variation in area-yield products across simulated indices.



One approach to controlling for variation in the simulation is control for each product's area-yield contract as its benchmark. That is, compare the difference in the ratio of households better off with the index products from each of their area-yield contracts. **Figure 10** shows those differences. Surprisingly, in all of these cases, the simulation average index product benefits more households than does their related area-yield contract. As briefly mentioned above, area-yield contracts are very susceptible to outliers, which draw the mean losses (indemnities) away from the median losses, and our data is characterized by a large number of outliers.[13] For example, the mean mortality rate in the data is 12.3% while the median is 3.1%. By this metric, the CZ eMODIS product performs the best as long as premium are not too high; as the premium increases the products become relatively more similar in the performance, which is decreasing in the premium rate.

[13] The ratios are lower in **Figure 9** than in **Figure 3** because of the oversampling of high loss events in the original survey seasons.

Figure 10. Ratio better with an area yield contract than with the index product.



Extension: Phenologically defined seasons and early index readings

A study by (Vrieling et al., 2016) found that end of season index reading could be reliably predicted much before the end of season. Simply put, the vegetation that forms during the early portion of each insurance season, which is the precipitation period, is an accurate predictor of the vegetation at remains by the end of the following dry seasons. In response to that research, and an appeal made by pastoralists to provide indemnity payments before losses are incurred, the IBLI product adjusted it parameters to make payments 1-4 months earlier than the original contract. Specifically, indemnity payments are made in January for the coverage during the October - February season and in May for coverage during the March -September seasons. These contracts have been aptly called asset protection contracts, as opposed to the earlier contracts, which are now referred to as asset replacement contracts.

In addition, the same research highlighted the heterogeneity in the timing of the true LRLD and SRSD seasons across space. Vrieling et al., (2016) redefine the start of each rainy season according when sudden increases in NDVI values are usually observed, allowing for the start and end dates to vary across space. This analysis has been used to develop index periods that more accurately reflect the variation in expected seasonal periods than the currently used definition, which relies on a single start/stop dates across all index regions. Although these phenologically defined seasons are likely to be more accurate, they have yet to be adopted by insurance agencies. We use them here to examine differences in what appears to us to be the ideal index and contract structure–one that makes payments as early as possible while also being context specific. The approach outlined by Vrieling et al., (2016) is used to develop the four phonologically defined indices, which are then processed at the early season end dates: CZ NDVI, CZ NDVI, CZ BOKU, and ZC BOKU. The pairwise correlations between the indices continues to be quite high–ρ>0.96 for all pairs.

Table 5 is made using the *pheno90p* indices and the approach used for **Figure 2**. The performance of the *pheno90p* indices are nearly identical to that of the original indices (Table 2 vs Table 3). The CZ eMODIS product performs the best but statistically the CZ/ZC and eMODIS/BOKU variations are indistinguishable from each other.

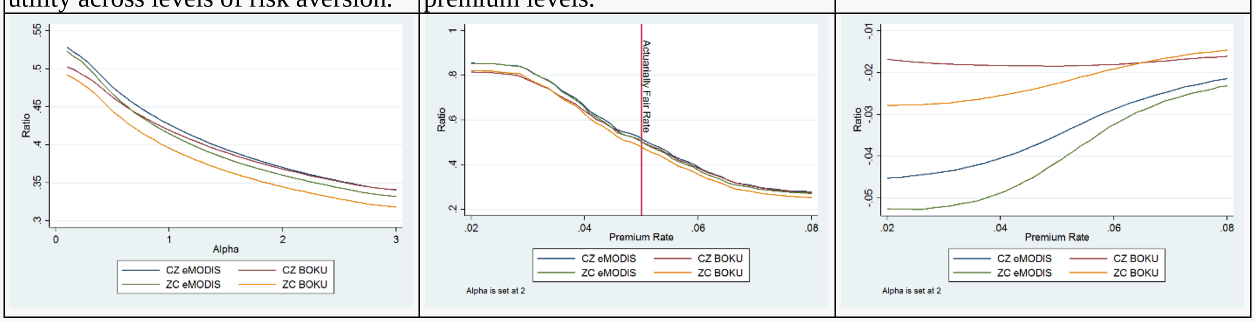
Table 5. The coefficient of determination and average differences in predicted and observed outcomes between observed mortality rates and the *pheno90* index values.

	Linear $[y = f(index)]$				3rd order poly. $[y = f(index, index^2, index^3)]$			
	CZ		ZC		CZ		ZC	
	eMODIS	BOKU	eMODIS	BOKU	eMODIS	BOKU	eMODIS	BOKU
Full Panel								
Observations	131	131	131	131	131	131	131	131
R^2	0.250	0.224	0.249	0.230	0.390	0.339	0.360	0.297
Mean error	0.063	0.064	0.064	0.064	0.055	0.058	0.057	0.060
Mean over	0.053	0.053	0.052	0.052	0.047	0.049	0.049	0.049
Mean under	0.078	0.081	0.082	0.082	0.067	0.070	0.069	0.077
Restricted Panel								
Observations	83	83	83	83	83	83	83	83
R^2	0.693	0.692	0.692	0.692	0.698	0.692	0.698	0.694
Mean error	0.020	0.020	0.020	0.020	0.019	0.019	0.019	0.019
Mean over	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.015
Mean under	0.025	0.026	0.026	0.026	0.023	0.026	0.025	0.026

The simulation process described in above is repeated for the *pheno90p* indices (**Figure 11**). The outcomes are nearly identical to those produced when using the original data (**Figure 7** vs. **Figure 11**a; **Figure 8** vs. **Figure 11**b; **Figure 10** vs **Figure 11**c). Specifically, the indices are nearly identical in their performance and the ratios of those that benefit from insurance coverage are quite similar. These findings support the claims by Vrieling et al. (2016) the early indices are accurate predictors of end-of-season conditions.

Figure 11. The performance of *pheno90p* based indices in simulations.

a. Ratio for whom actuarially fairly priced insurance increases expected utility across levels of risk aversion.	b. Ratio for whom each product increases expected utility across premium levels.	c. Ratio better with an area yield contract than with the index product

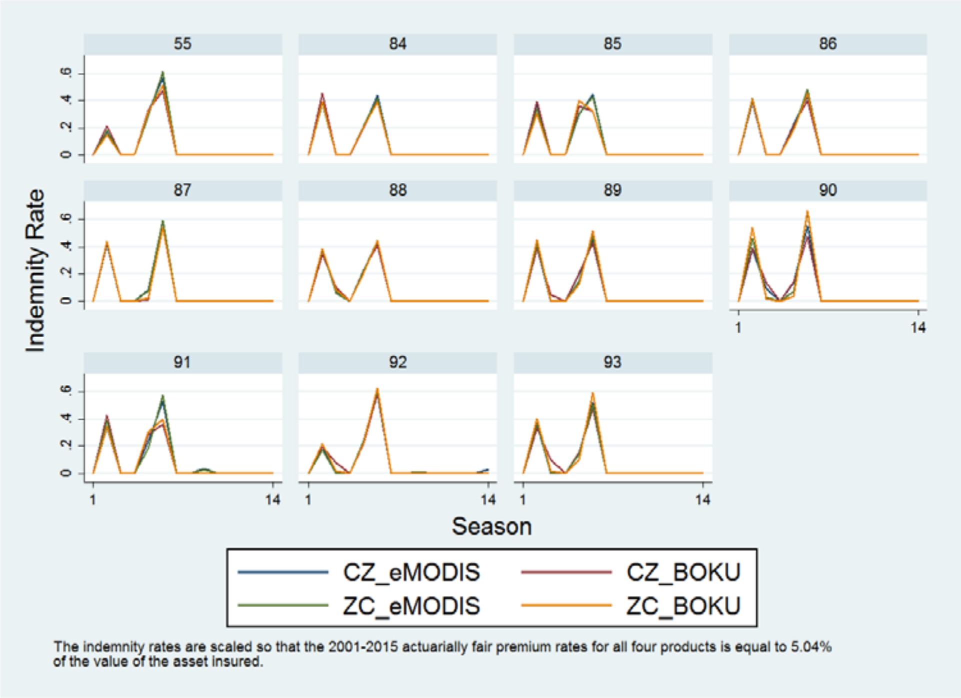


Extension: Relaxed contract constraints

In the above analysis, the index products are all actuarially fair and make indemnity payments at identical rates within the survey periods. Although maintaining equality in contract structure is necessary if we are to focus on the quality of the index, an alternative approach is to set the contract parameters using the full 14 years of data. The result is that during the survey period, the contracts will not necessarily may payments the same number of times nor will the sum of those payments be equal to the premiums paid. Using the *pheno90p* indices, we set the frequency of payouts over the 28 seasons at 20% and adjust the indemnities so that the 28-season actuarially fair premium rate is 5.04% (the same as above).

Figure 12 describes the indemnity payments, focusing only on the survey period. Small deviations in which periods indemnities are made and the magnitude of indemnities are apparent in the figure, butt all in all, the contracts continue to be quite similar.

Figure 12. Indemnity payments across seasons with contract parameters generated using the full 28 index seasons.

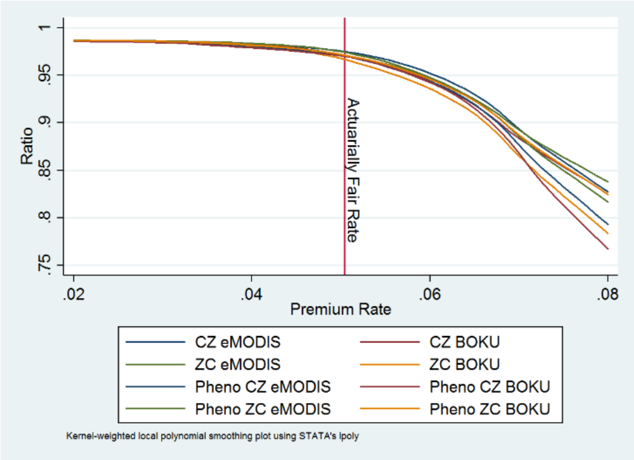


Running the same analysis as described above produces very similar results as above. The utility outcomes associated with all four indices are nearly indistinguishable, and this result is mostly robust across premium levels and levels of risk aversion. At the actuarially fair rate and setting alpha equal to zero, the CZ eMODIS product has the highest expected outcome, but that outcome is only statistically significantly different than the lowest performing index (p=0.076) and the difference between the two indices is about 3% of the sample, or 27 households.

We should note that the 2009 and 2011 droughts represent the lowest two index values in the series across all four indices and in most regions. The result is that within the survey seasons, insured household receive more in indemnity payments than they pay in premium rates. Thus, although the indices perform in a very similar manner, nearly all households are better off using this approach to premium calculation than the within survey period actuarially fair approach used for most of this paper.

As a final illustration of the similarity of these indices and their impacts on utility, we include **Figure 13**, which includes the outcomes of all eight indices using the relaxed constraints. Once again, we find that the indices are nearly indistinguishable until the premium rates become quite high.[14] At the actuarially fair rate, the *pheno90p CZ NDVI* product has the highest estimated ratio of those helped by the product, but is not statistically better than the second best product. Note that the ratio of individuals that benefit is very high due to the relaxed constraints, which results in all the products netting positive payouts during the survey periods.

Figure 13. Comparison of the eight indices with relaxed contract parameters.



Discussion and Limitations

Limitations

The livestock mortality data itself is both the key strength and key limitation of this research. The data is a strength because these types of ground truthing exercises are extremely rare for index insurance products because the validation data is so rare. This makes this paper unique in that it is comparing indices to actual losses on the ground. But, our approach has been to use the livestock data as though it were error free in order to examine errors across indices. In reality, the errors in the livestock data are probably much larger than the differences in the indices. These errors can be due to recall errors (e.g., month of mortality, number of animals), differences in cultures (e.g., definitions of ownership and households), and even input errors. In addition, the reported livestock mortality is attributed to the base-camp location of the household, even though the livestock death may have taken place well outside of that index region.

Although the indices suffer from their own set of errors, they do not face any of these spatial, temporal, or cultural errors. Rather, the indices are only different in how they process and filter data to reduce error in those data and to aggregate them over time and space. This is all to say that it is possible that "noise" in the mortality rate is larger than the differences in the indices, which could easily lead to spurious conclusions. One option for future research is to run the simulations to include resampling of households as well.

[14] The outcomes (ratios of those that benefit) of the four full season LRLD/SRSD indices studied in sections four and five are within less than one half of one percent of each other in this scenario.

Discussion

In this research we examined a number of different indices for their relative suitability as the basis of an index insurance product for livestock mortality. Our hope was that accounting for standard consumer preferences—namely risk aversion—would highlight differences in the indices that appear nearly identical under other commonly used approaches, helping us to identify which would be most useful for an index product. Although our analysis was able to distinguish between the indices to a greater degree than those based on correlation and mean errors, the differences continue to be quite small, even in simulations that allow for variation in the seasonal conditions.

The key implications of our findings are twofold. With respect to which index performs best, the analysis is not entirely consistent between the observed and simulated seasons. The analysis of the observed seasons point towards the ZC products, but cannot distinguish between the BOKU and eMODIS products. In all of the remaining analysis, the CZ eMODIS index as the most accurate. The discrepancy is likely due to bias towards extreme events in the initial sample caused by the two droughts that took place during the survey period.

The second set of implications are mostly with respect to the non-index characteristics of insurance contracts. That is, insurance providers should focus their attention on the non-index parameters of the contracts specifically, identifying the temporal cycle of risk and working to reduce premium rates. **Table 1** illustrates the importance of the first. Assuming that drought is a covariate event and causes livestock mortality, then the fairly low correlations in livestock mortality rates observed in some of the index regions points towards a potentially inaccurate index season. Livestock health and mortality may operate at a frequency that is not identical to the seasonal definitions employed by the IBLI product. For example, livestock mortality can take place early in a season if the rains are late, but such late-onset of precipitation can be missed by indices calculated mid-way or at the end of a season. In theory, the CZ indices should be more sensitive to late onset precipitation than the ZC indices, but good mid-season conditions can still easily off-set early conditions. In this example, developing contracts that explicitly provide coverage for late onset precipitation may be at least as important as choosing an index. The move to early indices, allowing indemnity payments to come early is another approach to ensuring that the timing of payouts is a beneficial as possible. Our analysis of early indices supports those of Vrieling et al. (2016) and IBLI's recent decision to move to the early index contracts, in that contracts made from the pheno90p indices perform as well as the full season indices.

Figures 4, 8, and 13 illustrate the importance of the second non-index parameter—premium rates. Changing the premium rates by even small amounts can have quite large impacts on the ratio of households that benefit from the index product. For example, using the estimates from section 3, increasing the premium rates from the actuarially fair rate of 5.04% to 7.04%, reduces the ratio of households that benefit from *any* of the insurance products from about 0.4 to less than 0.2, a 50% reduction.

Interestingly, the above two points illustrate a common conflict between timeliness and cost. One main advantage to the BOKU processing approach is that it produces indices more quickly than does the eMODIS process, but the BOKU data is fee-for-use while the eMODIS data is freely available. If those additional fees are passed on to consumers, they are faced with higher premiums for more timely indemnity payments. Pastoralists' preferences in this regards are unknown and should be studied.

References

Atzberger, C., & Eilers, P. H. C. (2011). A time series for monitoring vegetation activity and phenology at 10-daily time steps covering large parts of South America. *International Journal of Digital Earth*, *4*(5), 365–386.

Barré, T., Stoeffler, Q., & Carter, M. (2016). *Assessing index insurance: conceptual approach and empirical illustration from Burkina Faso*.

Black, E., Tarnavsky, E., Maidment, R., Greatrex, H., Mookerjee, A., Quaife, T., & Brown, M. (2016). The Use of Remotely Sensed Rainfall for Managing Drought Risk: A Case Study of Weather Index Insurance in Zambia. *Remote Sensing* . https://doi.org/10.3390/rs8040342

Brown, M. E., Osgood, D. E., & Carriquiry, M. A. (2011). Science-based insurance. *Nature Geosci*, *4*(4), 213–214. Retrieved from http://dx.doi.org/10.1038/ngeo1117

Carter, M. R., & Lybbert, T. J. (2012). Consumption versus asset smoothing: testing the implications of poverty trap theory in Burkina Faso. *Journal of Development Economics*, *99*(2), 255–264.

Chantarat, S., Mude, A. G., Barrett, C. B., & Carter, M. R. (2013). Designing index-based livestock insurance for managing asset risk in northern Kenya. *Journal of Risk and Insurance*, *80*(1), 205–237.

Clarke, D. J., Mahul, O., Rao, K. N., & Verma, N. (2012). Weather based crop insurance in India. *World Bank Policy Research Working Paper*, (5985).

de Leeuw, J., Vrieling, A., Shee, A., Atzberger, C., Hadgu, K. M., Biradar, C. M., … Turvey, C. (2014). The potential and uptake of remote sensing in insurance: a review. *Remote Sensing*, *6*(11), 10888–10912.

Dercon, S. (2004). Growth and shocks: evidence from rural Ethiopia. *Journal of Development Economics*, *74*(2), 309–329.

Duryea, S., Lam, D., & Levison, D. (2007). Effects of economic shocks on children's employment and schooling in Brazil. *Journal of Development Economics*, *84*(1), 188–214.

Elabed, G., Bellemare, M. F., Carter, M. R., & Guirkinger, C. (2013). Managing basis risk with multiscale index insurance. *Agricultural Economics*, *44*(4–5), 419–431.

Gong, E., de Walque, D., & Dow, W. H. (2015). Coping with Risk: The Effects of Shocks on Sexually Transmitted Infections and Transactional Sex in Rural Tanzania.

Hoddinott, J., & Kinsey, B. (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, *63*(4), 409–436.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*(1), 195–213.

Janzen, S. A., & Carter, M. R. (2013). *After the drought: The impact of microinsurance on consumption smoothing and asset protection*.

Jenkerson, C., Maiersperger, T., & Schmidt, G. (2010). *eMODIS: a user-friendly data source*.

Jensen, N. D., Barrett, C. B., & Mude, A. (2016). Index Insurance Quality and Basis Risk: Evidence from Northern Kenya, (5), 1450–1469.

Klisch, A., & Atzberger, C. (2016). Operational Drought Monitoring in Kenya Using MODIS NDVI Time Series. *Remote Sensing*, *8*(4), 267.

Mann, B. F., & Small, C. (2014). Providing a Spatial Context for Crop Insurance in Ethiopia: Multiscale

Comparisons of Vegetation Metrics in Tigray. In *AGU Fall Meeting Abstracts* (Vol. 1, p. 6).

Meroni, M., Verstraete, M. M., Rembold, F., Urbano, F., & Kayitakire, F. (2014). A phenology-based method to derive biomass production anomalies for food security monitoring in the Horn of Africa. *International Journal of Remote Sensing*, *35*, 2472–2492.

Robinson, J., & Yeh, E. (2011). Transactional sex as a response to risk in Western Kenya. *American Economic Journal: Applied Economics*, 35–64.

Roumiguié, A., Sigel, G., Poilvé, H., Bouchard, B., Vrieling, A., & Jacquin, A. (2016). Insuring forage through satellites: testing alternative indices against grassland production estimates for France. *International Journal of Remote Sensing*, 1–28. https://doi.org/10.1080/01431161.2016.1230288

Saha, A., Shumway, C. R., & Talpaz, H. (1994). Joint estimation of risk preference structure and technology using expo-power utility. *American Journal of Agricultural Economics*, *76*(2), 173–184.

Swets, D. L., Reed, B. C., Rowland, J. D., & Marko, S. E. (1999). A weighted least-squares approach to temporal NDVI smoothing. In *Proceedings of the 1999 ASPRS Annual Conference: From Image to Information, Portland, Oregon* (pp. 17–21).

Tack, J. (2013). The effect of climate on crop insurance premium rates and producer subsidies. In *Selected Paper. Agric. Appl. Econ. Assoc. Annual Meeting, Washington, DC*.

Vrieling, A., Meroni, M., Mude, A. G., Chantarat, S., Ummenhofer, C. C., & de Bie, K. C. (2016). Early assessment of seasonal forage availability for mitigating the impact of drought on East African pastoralists. *Remote Sensing of Environment*, *174*, 44–55.

Vrieling, A., Meroni, M., Shee, A., Mude, A. G., Woodard, J., de Bie, C. K., & Rembold, F. (2014). Historical extension of operational NDVI products for livestock insurance in Kenya. *International Journal of Applied Earth Observation and Geoinformation*, *28*, 238–251.