

A Spatial Econometric Approach to Designing and Rating Scalable Index Insurance in the Presence of Missing Data

Joshua D. Woodard^a, Apurba Shee^b and Andrew Mude^c

^aCharles H. Dyson School of Applied Economics and Management, Cornell University, 351B Warren Hall, Ithaca, NY 14850 U.S.A.

E-mail: joshua.woodard@cornell.edu

^bEnvironment and Production Technology Division, International Food Policy Research Institute, Arusha, Tanzania.

E-mail: a.shee@cgiar.org

^cInternational Livestock Research Institute, Nairobi, Kenya.

E-mail: a.mude@cgiar.org

Index-Based Livestock Insurance has emerged as a promising market-based solution for insuring livestock against drought-related mortality. The objective of this work is to develop an explicit spatial econometric framework to estimate insurable indexes that can be integrated within a general insurance pricing framework. We explore the problem of estimating spatial panel models when there are missing dependent variable observations and cross-sectional dependence, and implement an estimable procedure which employs an iterative method. We also develop an out-of-sample efficient cross-validation mixing method to optimise the degree of index aggregation in the context of spatial index models.

The Geneva Papers (2016) 0, 1–21. doi:10.1057/gpp.2015.31

Keywords: index insurance; spatial econometric models with missing data; NDVI; Kenya pastoralist livestock production; cross-validation; model mixing

Article submitted 19 April 2015; accepted 11 November 2015; advance online publication, 20 January 2016

Introduction

Index insurance has gained considerable interest during the past decade as a tool for transferring weather-related market risks to capital markets.¹ One such product in the marketplace is the Index-Based Livestock Insurance (IBLI) programme. An IBLI pilot was commercially launched in the Marsabit district of Kenya in January 2010, and in the Borana Zone in southern Ethiopia in July 2012 by the International Livestock Research Institute (ILRI), in cooperation with local insurers and global reinsurers. IBLI safeguards pastoralist livestock producers against drought and other vegetation risk. Considering the high correlation between forage loss and livestock mortality, IBLI is designed to insure livestock mortality by compensating pastoralists according to an area-average predicted

¹ Alderman and Haque (2007); Woodard and Garcia (2008a, b); Woodard (2008).

livestock mortality index which is established statistically by fitting household-level livestock mortality data to remotely sensed, normalised difference vegetation indexes (NDVI).

Bolstered by the potential of the product to manage the key source of vulnerability to pastoral communities, ILRI and their partners are currently working to scale up IBLI to 11 arid and semi-arid districts in Kenya (108 divisions). Scaling up a small pilot programme to such an extent is no trivial matter, as area-specific index response functions (i.e. mappings of remotely sensed vegetation indexes to mortality) must be constructed for all areas—many of which only have sparse mortality data available. Thus, this endeavour necessitates developing a methodology that is spatially scalable, feasible and robust in the presence of sparse data. Although the use of spatial econometric methods that capture cross-sectional dependence has become the standard in many branches of applied economics,² its use has been more limited in insurance economics contexts with few exceptions.³ Yet, spatial methods present a very promising route for addressing the challenges that often arise for development economists and actuaries in implementing such programmes.

In contrast to past approaches in the literature on index insurance, the objective of this work is to develop an explicit spatial econometric framework in which to estimate insurable index response functions, which is integrated within a general insurance rating/pricing framework. Building on the lessons learned from the IBLI pilot,⁴ the broader objective is also to improve the product design and actuarial credibility of the contract, and to work towards developing a platform that can be adapted to other contexts. We also address several common interrelated challenges that often arise in such efforts.

To this end, we explore and develop an implementable expectation maximisation (EM) procedure to facilitate the estimation of spatial panel models when there are missing observations for the dependent variable (in our case, livestock mortality survey observations) and cross-sectional dependence by extending the work of LeSage and Pace⁵ to the panel data case, and articulating an estimation method which employs an iterative procedure. Although other methods have been developed recently to estimating spatial panels with missing data,⁶ including a nonlinear least squares method, a generalised method of moments approach, as well as a two-stage least squares procedure, Wang and Lee⁷ find those methods to be less efficient than maximum likelihood estimation implemented under the EM algorithm.

From a design standpoint, our framework is very desirable in the current context, as the spatial parameter can capture the influence of spatial dependence in mortality as well as other feedback processes and unobservables. For example, since herds

² See, for example, Anselin (1988); Elhorst (2003); Kelejian and Prucha (2007); LeSage and Pace (2009).

³ For example, Woodard *et al.* (2012).

⁴ Mude *et al.* (2012).

⁵ LeSage and Pace (2004).

⁶ See, for example, Wang and Lee (2013a, b).

⁷ Wang and Lee (2013a).

migrate during droughts, and livestock mortality is a function of both rangeland carrying capacity and forage availability, these dynamic feedback effects provide valuable information for estimating mortality responses to drought. Spatial methods also present a formal and robust avenue to address the issue of missing data, which can create serious problems for standard non-spatial index construction methods.

Another unique challenge in the development of index insurance is that the remotely sensed data is not always continuously available or consistently measured through time, and so building up a long time-period of observations involves gathering data from multiple satellites, sensors and databases through time. Thus, we also address the issue of properly employing inter-calibration procedures in order to facilitate integration of such varied data within a cohesive insurance rating framework. In doing so, a bootstrapping methodology is developed to appropriately augment the remotely sensed data set for the period in which index estimation is conducted, with data from longer-running remote sensing platforms in order to improve rating efficiency.

As it regards optimal contract design and functional form, the determination of the optimal degree of contract aggregation is also of great importance—that is, determining whether the contract should be designed for a single area in isolation of others, for groupings of areas, or a combination thereof. Unfortunately, this important question has received little treatment in the index insurance economics literature to date. We propose a solution which employs a cross-validation optimal model mixing approach as motivated by Woodard and Sherrick.⁸ The methodology developed in the paper is applied to the IBLI programme in Kenya.

Background and literature review

Livestock are a significant global asset that play an important role in rural livelihoods and typically account for 20–40 per cent of agricultural GDP.⁹ In arid and semi-arid lands in Africa—which comprise about two-thirds of the continent—approximately 20 million pastoralists depend on livestock grazing as their main livelihood. Livestock are the key productive assets in these economies, as intensive crop production is often not possible due to low and risky rainfall, and poor soils. The occurrence of extreme drought threatens the livelihoods of these pastoralists, and much evidence exists to suggest that these uninsured weather risks can push households and communities into poverty traps from which escape is particularly difficult.¹⁰ As a result, producers tend to adopt low-risk, low-return economic activities which hamper investment and economic growth. IBLI has emerged as a promising market-based solution for overcoming some of these problems by transferring correlated weather risk in the community to insurance markets and, ultimately, international reinsurance and capital markets.

⁸ Woodard and Sherrick (2012).

⁹ Steinfeld *et al.* (2006); Herrero *et al.* (2013).

¹⁰ Lybbert *et al.* (2004); Barrett *et al.* (2006).

Motivation for index insurance

Index-based insurance does not insure households directly against individual losses (in our case, realised livestock death), but rather indemnifies them against an index that is correlated with the loss. Optimally, the index should mirror the underlying risk to be insured, should not be subject to moral hazard or adverse selection, and should be easily verifiable at low cost. Although in the best-case scenario, the index would be a perfect mapping to the household's realised losses, this is typically not possible. The resulting error between the index and the true loss to the insured is known as basis risk. Minimising basis risk is essential to enhancing the value proposition of the insurance offering, and thus it is important to have a firm understanding of the degree and form of basis risk.

So, why not insure live animals directly? There are several reasons why insuring live animals is typically not feasible, and even in the developed world it is typically not possible due to the inability of the insurer to properly monitor the animal, which opens the door for moral hazard. It would also be expensive to administer such insurance as claims adjustment would necessitate on-site visits from veterinarians and insurance agents. Particularly in regions such as northern Kenya, the institutional infrastructure that would be necessary to accomplish this simply does not exist. Index insurance, on the other hand, avoids most of these problems.

In our case, NDVI data are published freely and regularly by various government agencies, and thus there is a reliable and verifiable data source upon which to structure the insurance, a necessary condition. Second, the claims process is very streamlined and can typically be adjusted and paid relatively cheaply and easily. In Kenya, for example, the use of mobile money payment systems such as M-PESA are already pervasive throughout the country and facilitates an increasing number of financial transfers cost effectively through mobile networks.¹¹ Although the density of cellular telephone coverage and usage is lower in the remote regions of the country where pastoralists reside, it has been rapidly increasing over time. As such, the claims process could eventually utilise such systems, can be mostly automated after the sale, and should require very little effort and cost once the delivery system infrastructure is in place.

Of course, the benefits of index insurance come at the expense of basis risk. As noted, not only does the magnitude of basis risk matter, but also the form. At the most basic level, basis risk can be decomposed into a systemic/community component which arises due to design error, and an idiosyncratic/individual component that is not related to design.¹² Strong informal social insurance systems typically exist in these regions which allow individual households to cope with idiosyncratic losses. Indeed, the index insurance product developed here is not equipped nor intended to replace these informal systems. However, severe problems can arise for the community under systemic shocks as the social insurance systems may begin to break down. It is in covering these systemic shocks that index insurance provides value. We explore this more formally in later sections.

¹¹ Jack and Suri (2011).

¹² Woodard and Garcia (2008a).

Motivating spatial econometric approaches in the design of index insurance

Spatial methods are potentially advantageous in the context of designing and rating/pricing index insurance for several reasons. First, the perils to be insured as well as the covariates upon which the index insurances are structured are typically spatially correlated and/or dependent (in our case, mortality). Ignoring the spatial and dependent nature of the data may result in a less efficient set of contracts as well as biased indexes, from the econometric point of view.

Second, it is often the case—particularly in the developing world—that the data available regarding the peril to be insured (e.g. livestock mortality) may be prone to measurement error and/or have sparse coverage and missing observations. In our case, the mortality data are very rich, but does not contain mortality observations for all locations and time periods. Spatial methods can allow for maximal information extraction in these missing data cases, as information from mortality data in neighbouring locations can be formally utilised in a spatial context for predicting mortality where there are missing observations. The mechanism through which this is accomplished is via the link in the spatial parameterisation which ties information in non-missing dependent and independent variables to the independent variables corresponding to the locations with missing data for the dependent variable. This can be accomplished in a spatial lag-type framework but not typically in standard non-spatial frameworks, as the latter have no mechanism by which to incorporate the spatial information content.

Third, from an actuarial standpoint, a coherent system-wide spatial approach is potentially much more reliable, efficient and credible than would be individual location-by-location estimation of index response functions. Spatial methods also allow for a higher degree of control over the parameterisation across locations to be insured and often result in a natural smoothing of parameter estimates over space. This is important in insurance contexts, since a live contract will be sold in many regions, and so the actual shape and integrity of the estimated index at every location is crucial. In this way, the focus of economical insurance structuring differs somewhat from many typical economic investigations where inference alone may be the only primary output of interest from the model. In insurance economics applications, there is typically additional concern about whether the estimated model is reliable enough upon which to structure financial contracts in each location, and indeed, it is often the abnormal or “fringe” locations where data are poor that cause the most problems in the marketplace.

As we comment on more fully below, in the case of a spatial lag model, there is a natural interpretation for the estimated spatial autoregressive parameter as a type of credibility factor as it relates to the index design at any one location, as the final contract structured in any given location will be a function not only of its own NDVI observations, but also of neighbouring observations. Further, if there is in fact spatial dependence in the natural process of the insured peril, then spatial models allow for formal quantification of such. For example, it is not uncommon for pastoralists to modify their grazing patterns in response to conditions in their own and neighbouring locations. While a pastoralist typically has a home base—which in our case is where the survey data are periodically collected—they can also travel to some degree across location boundaries in search of vegetation and water. This fact provides a strong case for expecting spatial dependence in the livestock mortality data-generating process.

Data

Household-level mortality panel data

The household-level livestock mortality data are sourced from the Kenya Arid Land Resource Management Project (ALRMP) survey. The raw data provide household-level monthly observations for livestock herd size and mortality by animal type (sheep, cow, goat and camel) from January 2001 to December 2012, and contain approximately 900,000 survey observations. Since the different animal types vary significantly in weight and value, the widely used tropical livestock unit (TLU) measure is employed to construct a weighted measure for each household/monthly observation where 1 TLU = 1 cow = 0.7 camel = 10 goats = 10 sheep.

The ALRMP data represent an unbalanced panel. We focus on data that span 108 divisions within 11 districts in the arid lands of northern Kenya (see Figure 1). There are two seasons in northern Kenya, a short rain/short dry (SRSD) season running from October to February, and a long rain/long dry (LRLD) season that runs from March to September. Division-level seasonal TLU mortality rates are constructed by first calculating a seasonal TLU herd mortality rate for each household as the sum of deaths for all months in the season, divided by the maximum beginning herd size at the beginning of the month; second, a TLU weighted average mortality is calculated across households in a division within a season to arrive at the division-level seasonal mortality figure. It is not uncommon to have several seasons and locations with no data, and in fact, about 40 per cent of divisions/seasons in our data has no corresponding survey observations.

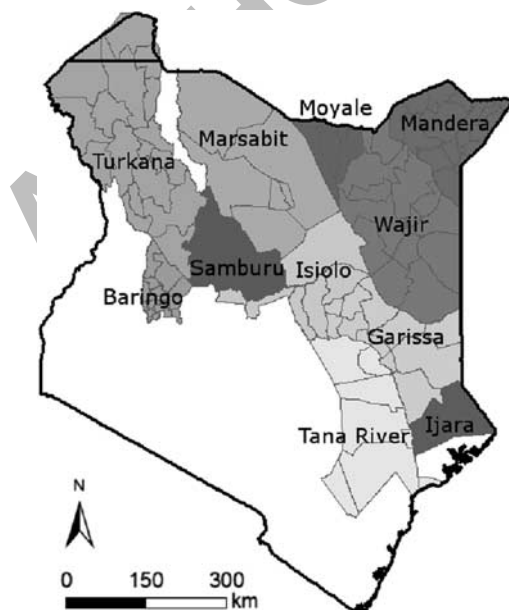


Figure 1. Study area in northern Kenya.

Remotely sensed vegetation data

We employ normalised difference vegetation index (NDVI) data derived from remote sensing platforms—which is an indicator of the level of photosynthetic activity in the vegetation—as our measure of forage scarcity to predict livestock mortality. There are several different NDVI data sets available, many of which come from different satellites/sensors and have different temporal coverage and spatial resolution. We employ data from NASA’s eMODIS data set, which is available from 2000 to the present.¹³ The data are published in raster format (essentially, pixel data) as 10-day, 250-meter composites.

The data processing employed is explained in more detail in Vrieling *et al.*,¹⁴ but we provide an overview of the processing here. First, the NDVI data are temporally filtered to correct errors from atmospheric effects and residual noise, as is standard in the remote sensing field. The pixel values for each 10-day time period are then aggregated and averaged by division. In general, NDVI is not a standardised measure and cannot be meaningfully compared across locations due to differences in hydrology, elevation, etc. Therefore, in order to create a meaningful measure that can be compared across space and time, some sort of standardisation must be employed. We employ a z -score measure, whereby, for each division and each 10-day period of the year, a 12-year average and standard deviation are calculated and used to arrive at a z -score for each observation. Next, to arrive at a season-level measure, cumulative z -score (C_zNDVI) indexes are constructed as the sum of the 10-day z -scores within each season. We also include lagged NDVI measures in the model to control for the state of rangeland condition at the commencement of the contract. Explicitly, we construct two measures, one for previous season C_zNDVI and one for C_zNDVI within the insured season (i.e. post contract commencement) for each division and time period/season as $pre-C_zNDVI_{ns} = \sum_{p \in (s-1)} zNDVI_{np}$ and $post-C_zNDVI_{ns} = \sum_{p \in s} zNDVI_{np}$, where $p \in s$ are the 10-day periods within the respective season and n is the division/location subscript. It is important to note that these NDVI data are employed as proxies for mortality, as they present a data source that can be independently validated, a requirement of index insurance. It is not the intent of the response function to exactly replicate the bio-physical process of mortality per se, but rather is a proxy model. This is essentially always the case with index insurance.

To obtain a broader range of NDVI measurements for the pricing of the contract, we also employ an inter-calibrated set of data from the GIMMS AVHRR NDVI3g sensor which runs from 1981 to 2012 and is published in 8-km, 15-day composites. GIMMS AVHRR data are processed similarly, and is also temporally filtered using an iterative Savitzky-Golay filter as described by Chen *et al.*¹⁵ Although longer time series of data are typically preferred for insurance pricing purposes for efficiency reasons, care must be taken in the calibration of data from different sensors. Considering the different

¹³ The eMODIS data for East Africa are downloaded from FEWS-NET www.earlywarning.usgs.gov/fews/africa/web/imgbrowse2.php?extent=eazd.

¹⁴ Vrieling *et al.* (2014).

¹⁵ Chen *et al.* (2004).

spatio-temporal resolution of the data sets, it is important to adjust for both time and spatial aggregation differences, since higher resolution sensors provide cumulative z -indexes that result from the aggregation of more pixels and time periods. For each division, regressions between the cumulative z -indexes of both sensors for the period in which the two have common data (2000–2012) are estimated and then used to calculate inter-calibrated AVHRR observations for the out-of-sample period (1981–2000) in which eMODIS data do not exist. Since the fit is not perfect between the two sensors, the volatilities of the fitted/inter-calibrated GIMMS AVHRR $CzNDVI$ values will not be comparable to those of the eMODIS $CzNDVI$ values (although in expectation, the conditional means will be). We address the implications of this and corrections that must be made if the inter-calibrated data are to be used to augment the rating structure further in the rating section below.

Spatial mortality index estimation procedure in the presence of missing observations

Consider the standard spatial panel model with a spatial autoregressive lag,

$$\mathbf{y} = \rho(\mathbf{I}_T \otimes \mathbf{W}_N)\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1)$$

where \mathbf{y} is an $NT \times 1$ vector of mortality observations which is sorted by time then division for N locations (in our case, divisions) and T time periods, \mathbf{I}_T is a $T \times T$ identity matrix, \mathbf{W}_N is an $N \times N$ spatial weight matrix which is row standardised (i.e. all rows sum to one) specifying the relative location of each location,¹⁶ ρ is a scalar spatial autoregressive coefficient that reflects the magnitude of spatial dependence, \mathbf{X} is an $NT \times K$ design matrix, $\boldsymbol{\beta}$ is a $K \times 1$ vector of coefficients, \otimes is the Kronecker product operator, and $\boldsymbol{\varepsilon}$ is a vector of random innovations. Note that we estimate separate models for the SRSD and LRLD seasons. This model requires a balanced panel for estimation (i.e. no missing observations for any location/time period) and, in such cases, can be estimated using maximum likelihood estimators,¹⁷ as well as via other approaches. The $\rho(\mathbf{I}_T \otimes \mathbf{W}_N)$ term acts as a precursor to a spatial filter, and thus interpretation of marginal effects and the calculation of fitted values in the spatial lag model are not as straightforward as in standard regressions. This can be seen by noting that Eq. (1) can be rewritten as $(\mathbf{I}_T \otimes \mathbf{I}_N - \rho(\mathbf{I}_T \otimes \mathbf{W}_N))\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, or $\mathbf{y} = (\mathbf{I}_T \otimes \mathbf{I}_N - \rho(\mathbf{I}_T \otimes \mathbf{W}_N))^{-1}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$. We construct the spatial weight matrix \mathbf{W}_N as a queen contiguity matrix, and hence it is sparse, while the spatial filter $(\mathbf{I}_T \otimes \mathbf{I}_N - \rho(\mathbf{I}_T \otimes \mathbf{W}_N))^{-1}$ is not.¹⁸ The implication of the spatial filter is that each cross-section represents a spatial network, and thus every location is a function of its own explanatory variables and innovations, as well as those of all other locations. This is analogous to the cross-sectional case where $(\mathbf{I}_T - \rho\mathbf{W})^{-1}$ is non-sparse, and so every observation is a function of itself, its neighbours and its

¹⁶ See Anselin (1988).

¹⁷ Elhorst (2003).

¹⁸ A queen contiguity matrix is one that expresses contiguity whereby two locations are considered neighbours if they share a border or vertex (see LeSage and Pace, 2009). See LeSage and Pace (2009) as well for interpretation of the spatial filter and other spatial econometric basics.

neighbours' neighbours (with influence decaying with distance), whereby the magnitude of the spatial dependence is moderated by the spatial lag coefficient ρ .¹⁹ In this sense, the term acts as a "spatial filter".

Estimation with missing dependent variable observations

Before delving into the construction of the design matrix and the implications for index insurance design, we first articulate the estimation procedure with missing data as motivated by LeSage and Pace.^{5,20} Lesage and Pace investigate a similar estimator in the context of a cross-sectional hedonic home pricing model and find that such an approach can improve prediction, increase estimation efficiency for the missing-at-random case, and reduce self-selection bias in the non-missing-at-random case. They develop a solution in the cross-sectional case that provides valuable guidance to extending to the panel data case.

As Lesage and Pace point out, the improved performance of the spatial model with missing data has nothing to do per se with the imputed missing dependent variable values, but rather results from utilisation of additional information in the independent variable values corresponding to the missing values, and the relationship and dependence among them in space. Intuitively, information content among the non-missing dependent variables and the independent variables corresponding to missing observations are linked via the spatial filter, which in turn allows for useful extraction of information embedded in the spatial nature of the data.

Although the estimation approach employed here is similar in concept to that of LeSage and Pace,⁵ we depart in two ways. First and most obvious, we extend to the panel data case. Second, we employ an approach that does not require manipulation of the standard spatial estimator to implement; that is, LeSage and Pace derive the likelihood function for the model by partitioning the data and components of the covariance matrix into their respective missing and non-missing components. They then substitute into the concentrated log-likelihood the missing dependent variable values with their expected values conditional on the observed sample information. A variety of computational techniques are then employed in order to facilitate estimation of the model using either MLE or Bayesian MCMC methods.

Replicating the approach of LeSage and Pace for the panel data case would increase the level of complexity of the estimator and presentation of such substantially, thus motivating our approach. We also recognise that even in the cross-section case, most applied researchers would likely find it quite a challenge to re-engineer the approach derived in that work, and would likely face great difficulty in operationalising the embedded missing value concentrated likelihood estimator presented in that work. Our approach, on the other hand, is feasible and easily implementable to any researcher who is able to estimate the standard

¹⁹ In typical applications testing is conducted to determine the likely form of the spatial dependence; however, to our knowledge, no such tests exist in the missing data case, so we proceed with the lag model for a variety of credibility reasons. Note that the fitted values from a spatial error type model will not be spatially smoothed necessarily, thus further motivating the lag approach. We comment on this further below in the contract design section.

²⁰ Although other methods exist as developed by Wang and Lee (2013a, b), those works were not published and the code not available when this pilot was designed. We would not anticipate large differences from employing different methods. Nevertheless, further investigation of those alternative estimators is beyond the scope of this work and is left as an area of future investigation.

spatial model without missing data, and simply involves replacing the initial missing dependent variable values with a set of reasonable starting values, and then iteratively replacing those missing values with a specific set of fitted values at each iteration and repeating the process until convergence.²¹ Explicitly, the steps to estimate are as follows:

1. Fill in missing dependent variable observations in \mathbf{y} with reasonable starting values (e.g. the mean of non-missing \mathbf{y} observations, or fitted values using a standard imputation model).
2. Apply a standard spatial panel estimator to the data²² to obtain estimates of $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\varepsilon}}$ and $\hat{\rho}$.
3. Construct a new vector $\hat{\mathbf{y}}$ by replacing the original missing values with their expectations conditioned on the innovation terms which correspond to the non-missing data only, as well as on $\hat{\boldsymbol{\beta}}$ and $\hat{\rho}$ as, $\hat{\mathbf{y}} = (\mathbf{I}_T \otimes \mathbf{I}_N - \hat{\rho}(\mathbf{I}_T \otimes \mathbf{W}_N))^{-1}(\mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\varepsilon}})$, where $\hat{\boldsymbol{\varepsilon}} = \begin{bmatrix} \mathbf{0}_{Miss} \\ \hat{\boldsymbol{\varepsilon}}_{NoMiss} \end{bmatrix}$ is the vector that contains the estimated innovations from the model in Step (2) for those that correspond to non-missing dependent variable observations, and zero for innovations corresponding to the missing observations.
4. Return to Step 2 and employ the new vector $\hat{\mathbf{y}}$ in the estimation in Step (2). Iterate until convergence.²³

Note that in Step 3, the calculation of $\hat{\boldsymbol{\varepsilon}}$ is crucial. If one were to also include in Step 3 the estimated innovation values in the calculation of $\hat{\boldsymbol{\varepsilon}}$ with those that correspond to *missing* values from the estimated model in the previous step, the result would be downward biased final estimates of $\hat{\rho}$. Further, if one were to use only the fitted values at each iteration (i.e. not conditioning on the innovations for the non-missing values and simply replacing the entire $\hat{\boldsymbol{\varepsilon}}$ vector with values of zero in Step 3) a similar problem would emerge.

²¹ LeSage and Pace (2004) also explain how the traditional approach of using the vectorised concentrated log-likelihood can be used to operationalise their approach by iteratively optimising the concentrated log-likelihood over the single spatial autoregressive parameter ρ , then constructing new estimates of β , then estimating a new value for σ (using only non-missing data), and then constructing a new conditional expectation of missing values that is conditional on the last iteration estimation of ρ , β and σ (using only non-missing values in the estimated error terms) to create a “repaired” dependent variable vector, and then iterating this process until convergence. Essentially, what we propose is similar except that our method does not require manipulation of the concentrated log-likelihood function. However, similar to how LeSage and Pace iteratively estimate using only non-missing data when recalculating the σ then constructing the “repaired” dependent variable vector, we simply recalculate the missing dependent variable values at each iteration using the equation for the fitted values conditional on the estimated errors, replacing those that correspond to missing values with zero. To our understanding, our proposed method and the iterative method articulated by Lesage and Pace are mathematically similar but implemented slightly differently computationally. Indeed, the logic behind both approaches is similar conceptually, although we go about it in an albeit more direct and simple manner from an implementation standpoint. It is also not apparent that the methods proposed by LeSage and Pace involve choosing starting values for missing observations in the initial estimation, whereas our approach explicitly does require the analyst to pick a set of starting values for the missing observations.

²² For example, Elhorst (2003).

²³ We conducted Monte Carlo simulations to evaluate the performance of this technique and found results similar to those in LeSage and Pace (2004). We found that the model typically converged to a reasonable level in less than 15–20 iterations. We would caution that while our Monte Carlo results and those in this paper for the mortality models did not appear to have any issues converging, there could be cases in which this might not occur. Monte Carlo results and code are available from the authors on request.

Functional form considerations

The primary risk to livestock in northern Kenya is drought. However, there are also some indications, from speaking with pastoralists and extension agents in the region, that conditions which are too wet or cold can increase mortality incidence. Furthermore, past research on livestock mortality in Kenya suggests the presence of non-linearities in the response of mortality to NDVI.²⁴ We employ a quadratic functional form on the $CzNDVI$ terms to take account of these non-linearities. Past research and field intelligence also suggests the presence of time dependence in the mortality process, as a very dry prior season can weaken the animals and leave them more prone to dying in the next season. Thus, we also include the lagged quadratic terms for previous season $CzNDVI$ (*pre-CzNDVI*) as a regressor, as well as its square.

Determining the optimal level of contract aggregation

The insurance market offering necessitates developing, for each division, an individual index upon which to structure insurance (which may or may not be unique of other divisions). Chief among concerns in selecting an appropriate modelling framework is determining the level of aggregation to be used when constructing the design matrix (explanatory variables) and estimating the mortality index response function, as the existence of heterogeneity in the underlying spatial processes across regions classified, as one spatial unit could lead to biased estimates. On the other hand, if there is a high degree of spatial congruence, then application of a model with location-specific terms will lead to less efficient parameter estimates and resulting indexes than would a model parameterised at a higher level of aggregation. For example, suppose we have one NDVI/weather variable. Should each division get its own response parameter? Should each district get its own? Or, should we just impose the same parameter for the system?

There are many potential variations that one could employ in parameterising the index. This could range from treating all the 108 divisions as if each had the same intercept and mortality response with respect to changes in NDVI, or alternatively we could impose all divisions to have their own response and intercept. For example, one could pool all data analogous to the pooled OLS panel model and structure \mathbf{X} to have one intercept and one parameter for the explanatory variable. This would result in a contract that has the same response function in each division, albeit each division would be conditional on its own data more heavily than on data in other locations. At the other extreme, the design matrix could be structured as a fixed-effects model where each location has its own intercept, and further, could also allow for fixed effects on the explanatory variable responses. If we choose an extremely restrictive model, this would likely lead to biased estimates due to underlying heterogeneity in the system. However, choosing a parameterisation that is too flexible is likely to lead to over-fitting and low predictive efficiency.

We explore two competing models here, one with division-level fixed effects on the intercept and explanatory variable NDVI response terms, and the second which employs district-level fixed effects for the intercept and response terms. Note that we do not intend to canonise any particular functional form as it regards index insurance more generally, and the

²⁴ Chantarrat *et al.* (2012).

choice of functional form in the models presented here is motivated primarily out of expositional and operationalisation considerations. To illustrate for the division level model, for example, we can partition the design matrix into the intercept and NDVI-specific components as $\mathbf{X} = [\mathbf{X}_{NT \times N}^{Int} \mathbf{X}_{NT \times 4N}^{CzNDVI}]$. In the case of $\mathbf{X}_{NT \times N}^{Int}$, the value of the element in row $i \cdot t$, column n , equals zero if the corresponding division mortality observation in \mathbf{y} is not the n th division, and is equal to one if the corresponding division mortality observation in \mathbf{y} is the n th division. The case is similar for the $CzNDVI$ terms, except that the elements in $\mathbf{X}_{NT \times 4N}^{CzNDVI}$ are either equal to zero or the value of the variable, depending on if the observation is in the respective division.

Obviously, the trade-off relates to one between bias and efficiency. In practice, a common problem is that analysts are often inclined to adopt a model at the lowest-level of space/time aggregation possible given their data, since the more highly parameterised model will always show itself to have better in-sample fit. This results oftentimes, unfortunately, in a very inefficient set of indexes that are of little use. This problem is compounded when there are missing data or many explanatory variables. Consideration of this relationship between bias and efficiency—and the task of finding some sort of optimal trade-off between the two—is an activity that is well known to those working in insurance fields. Determining the optimal level of contract aggregation is important not only for index construction, but also for rating/pricing efficiency. Ideally, an optimal weighting between a model with a specific (local) parameterisation and general (global) parameterisation would be employed, as opposed to a digital choice between one or the other. Estimation of such a weight is difficult, however, as a division-level fixed-effects model will always have more parameters, and thus will always appear to have superior fit in-sample (but potentially be less efficient out-of-sample).

To address this we employ a leave-one-out cross-validation (CV) optimisation procedure to estimate optimal model weight, as motivated by Woodard and Sherrick.⁸ Woodard and Sherrick develop this method in the context of univariate unconditional probability distribution estimation in pricing yield insurance, although the method extends in a straightforward manner to any scenario in which there are sets of competing models, including regression models (which are in essence simply a characterisation of a conditional distribution).

Cross validation optimised model mixing procedure

The CV optimisation estimator is implemented as follows. Each model is successively re-estimated whereby one year worth of observations are left out at each iteration. Thus, each model above is estimated $T = 11$ times, one for each hold-out year. At each iteration, a forecast is calculated for the observations which are held out, conditional on the $CzNDVI$ values in the hold-out year. Last, an optimal weight can be estimated for each model by maximising the out-of-sample log-likelihood (or in the case of normally distributed errors, minimising the out-of-sample sum-of-squared error between the predicted values and the actual values). Note that we only employ data for non-missing values in optimising the weights. Explicitly, we calculate $\hat{\underline{\mathbf{e}}} = [\hat{\underline{\mathbf{y}}}_1 \cdot \omega_1 + \hat{\underline{\mathbf{y}}}_2 \cdot \omega_2] - \mathbf{y}$, where $\hat{\underline{\mathbf{y}}}_m = [\hat{\mathbf{y}}_{m,-1}', \hat{\mathbf{y}}_{m,-2}', \dots, \hat{\mathbf{y}}_{m,-T}']'$ is the stacked vector of the out-of-sample mortality predictions for each model, $m = \{DistrictModel, DivisionModel\} = \{1, 2\}$ and $\hat{\mathbf{y}}_{m,-t}$ is the out-of-sample estimate from the t th CV iteration for model m . In order for the resulting model to be a valid

mixture model, the restriction that $\sum_m \omega_m = 1$ is also imposed. Optimal weights are then estimated as, $\boldsymbol{\omega}^*(\mathbf{y}, \hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2) = \arg \min_{\boldsymbol{\omega}} (\hat{\boldsymbol{\xi}}' \hat{\boldsymbol{\xi}})$.

With the estimate of $\boldsymbol{\omega}^*$ in hand, the final index model is then constructed as the weighted sum of the estimated component models, where each component model is estimated using all data. Note that this framework can extend to any arbitrary number of component/competing models and model weights. The optimisation of weights at the out-of-sample stage does not affect the estimation of the model parameters of the underlying candidate distributions; rather, it optimises the weight given to each model itself. The insight of the approach is that, in the case of a mixture distribution, the weights can be optimised within the out-of-sample likelihood function as a linear weighted sum of the component models, independent of the model parameters themselves. Doing so explicitly takes into account the impact of out-of-sample inefficiency which is ignored by other in-sample and EM-type algorithms. Other objective functions could also be employed (e.g. downside risk measure). As it relates to credibility theory, the optimal weights have a natural interpretation as credibility parameters for each model.

Contract structure and pricing

The IBLI insurance pilot is structured as a call option on the estimated mortality index for each season, where the strike price/deductible is an election of either 10 or 15 per cent. The contract sold covers the next two seasons. For example, if a pastoralist buys insurance in time for LRLD sales closing, he/she is sold a bundled contract that provides coverage for the next season (i.e. LRLD) as well as the following SRSD season. The indemnity function for each season is defined as,

$$\text{Indemnity}_{t,n} = \text{Max}[\hat{y}_n(\mathbf{CzNDVI}_t \mid \hat{\boldsymbol{\beta}}, \hat{\rho}, \boldsymbol{\omega}^*) - \text{Deductible}, 0] \times p \times \text{TLU_Units} \quad (2)$$

where *TLU_Units* are the number of TLUs insured, *p* is the fixed indemnity price per *TLU* insured (20,000 or 25,000 shillings currently), *Deductible* ∈ [10%, 15%], *t* is the time period /season, *n* is the division subscript, and the predicted mortality index in region *n*, $\hat{y}_n(\mathbf{CzNDVI}_t \mid \hat{\boldsymbol{\beta}}, \hat{\rho}, \boldsymbol{\omega}^*)$, is the *n*th element of the predicted mortality vector for the period conditioned on the realised NDVI measurements in period *t* and the estimated parameters that define the index,

$$\begin{aligned} \hat{\mathbf{y}}(\mathbf{CzNDVI}_t \mid \hat{\boldsymbol{\beta}}, \hat{\rho}, \boldsymbol{\omega}^*) &= \boldsymbol{\omega}_1^* \cdot \left((\mathbf{I}_N - \hat{\rho}_1 \mathbf{W}_N)^{-1} [\mathbf{X}_1^{Int} \mathbf{X}_{t,1}^{CzNDVI}] \hat{\boldsymbol{\beta}}_1 \right) \\ &+ \boldsymbol{\omega}_2^* \cdot \left((\mathbf{I}_N - \hat{\rho}_2 \mathbf{W}_N)^{-1} [\mathbf{X}_2^{Int} \mathbf{X}_{t,2}^{CzNDVI}] \hat{\boldsymbol{\beta}}_2 \right). \end{aligned} \quad (3)$$

where subscripts 1 and 2 are district and division models, respectively.

Note that $\hat{y}_{t,n}$ is a function of its own NDVI realisations as well as all other divisions because of the spatial network nature of the model. The actuarially fair premium (or price) for each season is calculated as,

$$\text{Prem}_n = \int (\text{Max}[\hat{y}_n(\mathbf{V} \mid \hat{\boldsymbol{\beta}}, \hat{\rho}, \boldsymbol{\omega}^*) - \text{Deductible}, 0] \times p \times \text{TLU_Units}) f(\mathbf{V}) d\mathbf{V} \quad (4)$$

where $f(\mathbf{V})$ is the joint *pdf* of the *pre-CzNDVI* and *post-CzNDVI* values, which in our case is estimated using the historical eMODIS data and inter-calibrated AVHRR data. For this paper, we focus on unconditional insurance prices, but in practice, if the season being rated is

the first to be insured, then specific conditioning on current NDVI conditions in the rating integration would be prudent. Note that the *premium rate* is $\text{PremRate}_n = \text{Prem}_n / (p \times \text{TLU_Units})$, and that there is a unique contract and premium for each season (i.e. LRLD and SRSD); we drop the subscript for ease of exposition.

In order to adjust for the downward biased variance of the inter-calibrated AVHRR data, a bootstrapping procedure is employed to numerically integrate over the AVHRR data whereby a vector of residuals is sampled with replacement from the original calibration regression (the common period, which spans 2000–2012) one year at a time, and matched to each respective division and added to the fitted value for the out-of-sample inter-calibrated value (spanning 1981–2000) before passing to the indemnity function. These are then averaged for each year to obtain an unbiased indemnity estimate for each of the historical AVHRR year observations. These are then weighted equally by year with the eMODIS historical indemnities to arrive at the final premium.

Another item that must be taken into account is that, for any given live contract offering, the *pre-CzNDVI* values upon which the index is also conditioned for the first season contract will have already been observed by the insureds prior to purchase. Thus, these realised *pre-CzNDVI* values should be employed to condition the index function (i.e. instead using the actual historical values for *pre-CzNDVI* when performing the integration) when pricing in order to reduce adverse selection.

Results

To focus the presentation of results, we only present those for the SRSD season here. Figure 2 presents average mortality rates constructed from the ALRMP data for the divisions in Kenya under investigation. The maps indicate that average mortality rates vary by region, with regions in the far north typically experiencing higher mortality. A clear pattern of spatial autocorrelation is also observed. Figure 3 displays percentiles for the raw mortality data by year (note, not all years have representation from the same divisions). The percentiles indicate a high degree of catastrophic risk in the underlying process, reflecting severe droughts in 2005 and 2009, and also indicate that the mortality observations tend to be correlated with each other across-time periods, indicating a high degree of spatial autocorrelation. This is not unexpected since the processes driving mortality are related primarily to weather, which is highly spatially correlated.

Table 1 presents regression results for the division level and district level spatial models using the missing data estimator described previously. The results are presented for the final iteration of the procedure. The table is split into panels that contain averages and standard deviations for the fixed parameters of the respective type (i.e. intercept, *pre-CzNDVI*, *post CzNDVI*, and their squares) as well as the number of parameters in each group and the percentage of the parameters in each group that are significant at the 5 per cent level. The *R*-square values are 0.9804 and 0.4050 for the division and district-level fixed-effects models, respectively (note that we calculate these considering only the non-missing values using their fitted values to calculate the residuals). As expected, the more highly parameterised division model has a much higher in-sample *R*-square; however, the out-of-sample model weighting is lower for the division model

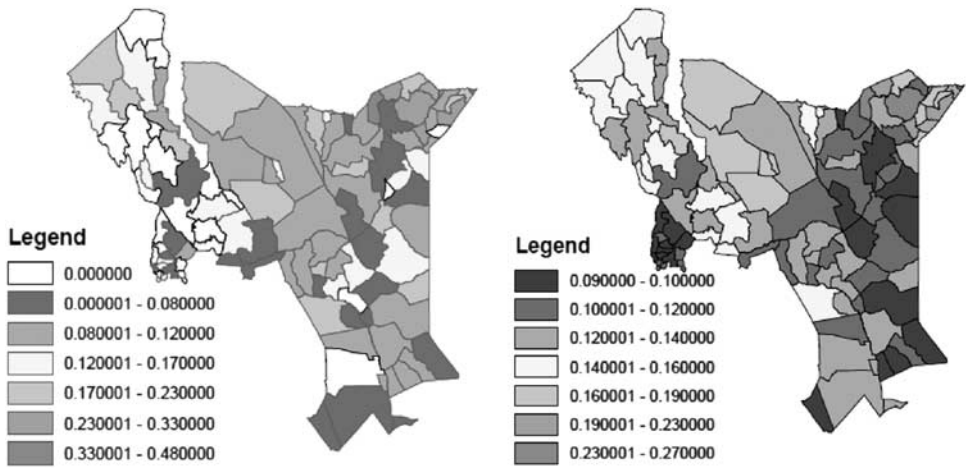


Figure 2. Kenya average livestock mortality rates (SRSD, 108 divisions), 2001–2011.
Note: The left panel displays average mortality for the original data with missing observations, while the right panel displays the average mortality for the fitted index with all year/location missing observations imputed with the model implied index estimate.

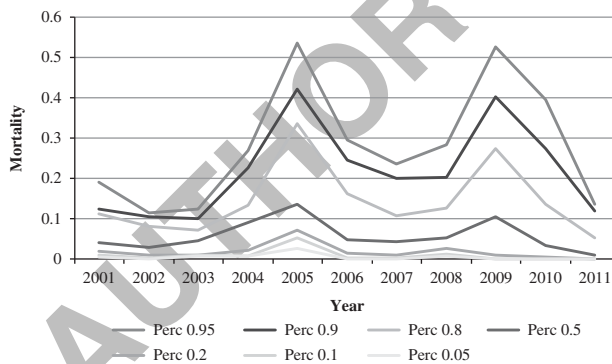


Figure 3. Mortality rates by year, raw data percentiles (SRSD, 108 divisions), 2001–2011.

($\omega_1^* = 0.2776$) vs the district model ($\omega_2^* = 0.7224$) indicating that the out-of-sample properties favour the district model.²⁵ The spatial dependence parameters were also significantly greater than zero for both the division ($\hat{\rho} = 0.4950$) and district ($\hat{\rho} = 0.1800$) level models. Note that if $\hat{\rho}$ were equal to zero (i.e. no spatial dependence or correlation) then the fill-in procedure would not converge.

²⁵ We also investigated a model with division fixed effects for the intercept and district fixed effects for the NDVI terms and this model outperformed both competing models. For clarity and ease of exposition, we present these models though for illustration.

Table 1 Spatial panel regression results

<i>Statistic/Model</i>	<i>Division intercept and NDVI</i>	<i>District intercept and NDVI</i>
Raw <i>R</i> -squared	0.9926	0.5933
Non-missing value <i>R</i> -squared	0.9804	0.4050
σ^2	0.0007	0.0037
ω^*	0.2776	0.7224
ρ	0.4950	0.1800
<i>p</i> -value	0.0000	0.0000
<i>Intercept parameters</i>		
# of parameters	108	11
mean	0.0501	0.0742
st. dev.	0.0745	0.0287
% significant at $\alpha=5\%$	69.4%	100.0%
<i>pre-CzNDVI parameter</i>		
# of parameters	108	11
mean	-0.0005	-0.0011
st. dev.	0.0040	0.0014
% significant at $\alpha=5\%$	50.0%	36.4%
<i>pre-CzNDVI² parameter</i>		
# of parameters	108	11
mean	0.0001	0.0001
st. dev.	0.0003	0.0001
% significant at $\alpha=5\%$	45.4%	63.6%
<i>post-CzNDVI parameter</i>		
# of parameters	108	11
mean	-0.0023	-0.0049
st. dev.	0.0059	0.0028
% significant at $\alpha=5\%$	50.0%	90.9%
<i>post-CzNDVI² parameter</i>		
# of parameters	108	11
mean	0.0001	0.0002
st. dev.	0.0004	0.0002
% significant at $\alpha=5\%$	57.4%	72.7%

Figure 4 provides select fitted and historical *CzNDVI* and mortality figures for the Oldonyiro Division (which is a representative division in our sample) in the Isiolo district to illustrate the historical fitted indexes. The first panel displays the estimated marginal index response according to the optimised model over various values of the cumulative *post-CzNDVI* value. Note that the fitted values do not lie perfectly on the plotted line as the fitted impact is a function of not only the division's own *CzNDVI* but also its neighbours through the spatial filter, where the impact plotted for each year equals the element corresponding to the division in the vector $\hat{\Delta}y_t = \sum_m \omega_m^* \cdot ((\mathbf{I}_N - \hat{\rho}_m \mathbf{W}_N)^{-1} \mathbf{X}_{t,m}^{CzNDVI-Post} \hat{\beta}_m^{CzNDVI-Post})$. The second panel provides the historical fitted index according to the optimised model by year (2001–2011), the indemnity payment per unit of shillings insured

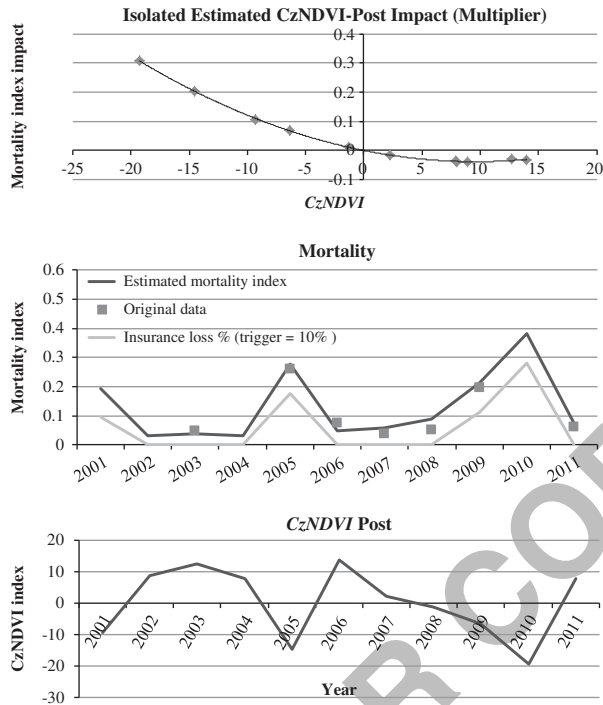


Figure 4. Fitted and historical mortality and $CzNDVI$, Isiolo District, Oldonyiro Division.

Note: The figure provides select fitted and historical $CzNDVI$ and mortality figures for one division. The first panel displays the cumulative $CzNDVI$ post value (i.e. during insured season) and the estimated marginal index response according to the optimised model. The second panel provides the historical fitted index according to the optimised model and the historical non-missing mortality data points. The third displays the corresponding $CzNDVI$ post values.

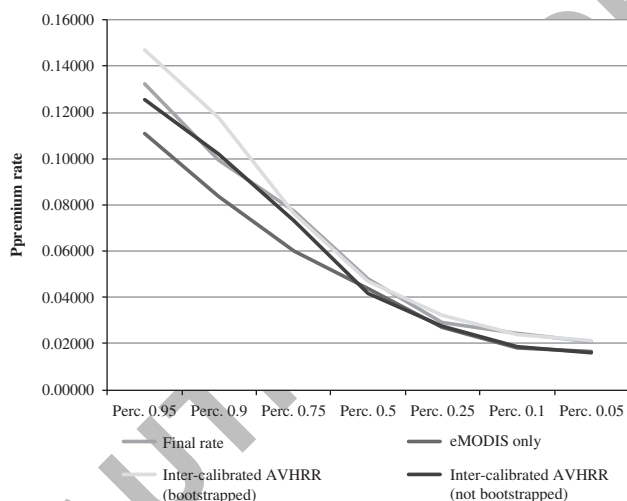
for a contract with a 10 per cent trigger, as well as the original data points for non-missing values. The third provides the corresponding *post-CzNDVI* values by year.

Several observations stand out. First, from panel 2 in Figure 4 it is observed that the fitted values correspond closely to the historical values, indicating that the index provides a reasonable proxy to mortality and also reduce impacts of outliers (note the scales). Second, the estimated index is consistent as it regards response to $CzNDVI$ and illustrates that the model is predictive and stable, which is attractive given the severe lack of data available upon which to fit the models.

Table 2 provides premium rate estimate summary statistics (mean, standard deviation and percentiles) across all divisions for a 10 per cent trigger/deductible contract, for the final premium rate as well as for the rates calculated using only eMODIS and AVHRR data. Recall, the final rate is an average of the eMODIS and AVHRR rates, weighted by the number of years represented by each. Figure 5 presents the same information in graphical format, and Figure 6 presents a map of the final weighted premium rate. There is a high degree of heterogeneity across divisions due to the fact that the volatility of the $CzNDVI$ measures vary widely across districts (note, the cumulative measure does not standardise variance across regions, nor would that likely be desirable in our context) and also due to the

Table 2 Rate results summary statistics, 10 per cent trigger

<i>Statistic</i>	<i>Final rate</i>	<i>eMODIS only</i>	<i>Inter-calibrated AVHRR (bootstrapped)</i>	<i>Inter-calibrated AVHRR (not bootstrapped)</i>
Mean	0.05702	0.04755	0.06249	0.05675
St. Dev.	0.03830	0.02983	0.04840	0.04637
Perc. 0.95	0.13226	0.11068	0.14718	0.12535
Perc. 0.9	0.09940	0.08376	0.11769	0.10204
Perc. 0.75	0.07722	0.06004	0.07651	0.07307
Perc. 0.5	0.04804	0.04387	0.04680	0.04156
Perc. 0.25	0.02907	0.02715	0.03223	0.02749
Perc. 0.1	0.02427	0.01828	0.02397	0.01879
Perc. 0.05	0.02067	0.01641	0.02113	0.01593

**Figure 5.** Estimated premium rates, 10 per cent trigger, percentiles by division.

fact that the index averages relative to the trigger point vary markedly across divisions as well (see Figure 3). In practice, careful determination would need to be made as to which deductible levels should be offered in various regions to maximise scalability.²⁶ The mean final weighted rate across all districts was about 5.7 per cent, and ranged from about 2 to 13 per cent across districts. In our case, the eMODIS rates were lower on average than those using the AVHRR data, indicating that the eMODIS period (2000–2012) had a lower frequency/severity of loss events relative to the time period covering the inter-calibrated

²⁶ We also evaluated within-division basis risk across individuals using a Bayesian random coefficients model (results not presented). The fraction of unexplained variance related to underpayments was approximately 25 per cent.

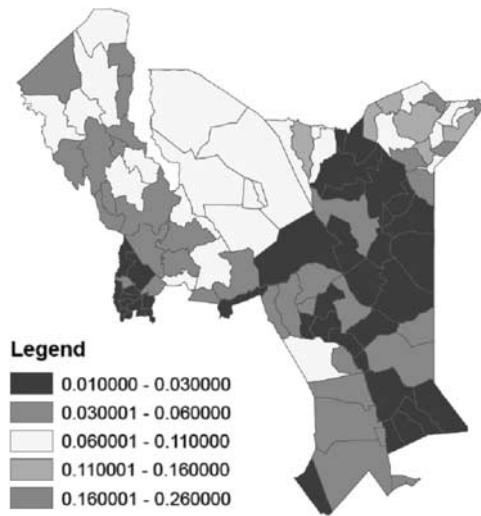


Figure 6. Estimated final premium rates 10 per cent trigger.

AVHRR data (1981–2000). Note also that use of the bootstrapping procedure to correct for variance deflation in the inter-calibrated series also resulted in higher rates, as expected.

Conclusion

Index insurance products have great potential for risk management in low-income populations. Furthering the appropriate development of such products will have significant developmental impacts. IBLI in particular has gained popularity in pastoral areas where it provides protection against drought-related mortality risk, and there is tremendous potential to expand such livestock insurance throughout the Horn of Africa and elsewhere. This study develops a methodology for product design which facilitates the scaling-up of such products in the presence of missing data and cross-sectional spatial dependence, and provides an application to the IBLI programme in Kenya. A cross-validated model mixing approach is also developed to optimise the degree of index aggregation. We also discuss some key issues in the design and rating of index insurance.

Insurance offerings derived from this methodology have been operationalised in the arid areas in Kenya since the fall of 2013. While the methodology for scalable index construction presented here pertains only to Kenya, it provides a framework that should be easily adaptable to other index insurance programmes around the world that typically face similar data, spatial and cross-sectional dependence issues.

A key takeaway is that the use of spatial econometric methods is attractive in such contexts as they allow for better information extraction—particularly in common missing data cases—and also provides a natural framework for scalability in index construction and market development. Incorporation of precise information regarding space arguably allows for more efficient index estimation which improves not only the utility of the insurance, but

also allows for more efficient contract pricing. The improved pricing efficiency should increase the confidence of insurers and reinsurers in the resulting products and thus motivate them to decrease risk-loadings for model uncertainty. Ultimately, this will help enable insurers to offer more competitive insurance rates and expand delivery to wider areas.

Some qualifications are in order. While it is our intent to propose a scalable method to index insurance design that we believe has broad appeal and potential, we do not put it up as the be-all-end-all solution. In practice, care must be taken to carefully assess the performance of estimated indexes for each offering/location. Indeed in our case, a thorough individual review of each division where insurance was to be offered was conducted to ensure plausibility and evaluate the behaviour of the insured indexes. Thus, we caution practitioners from naively applying these methods—or any modelling approach, for that matter—without carefully validating the empirics of the end product.

Future research could investigate alternative functional forms, animal specific contracts, or the use of other weather variables and remotely sensed data in index construction. Future research could also investigate further the conditioning of the underlying NDVI dynamics in the pricing/rating of the contracts, which could include evaluating effects of ENSO, climate change, or other space-time dynamics to further improve pricing efficiency.

Acknowledgements

This work was funded under International Livestock Research Institute Cooperative Work Agreement “ILRI-Cornell Collaborative Work Agreement for Special Joint Research Projects”. We would like to thank seminar participants at the 2014 International Agricultural Risk, Finance and Insurance Conference (Zurich) for helpful comments and suggestions. All errors are our own.

References

- Alderman, H. and Haque, T. (2007) *Insurance against covariate shocks: The role of index-based insurance in social protection in low-income countries of Africa*, World Bank Working Paper 95, World Bank, Washington, DC.
- Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Studies in Operational Regional Science, vol. 4, Dordrecht, MA: Kluwer Academic Publishers.
- Barrett, C.B., Marenya, P.P., McPeak, J.G., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J.C., Rasambainarivo, J. and Wangila, J. (2006) ‘Welfare dynamics in rural Kenya and Madagascar’, *The Journal of Development Studies* 42(2): 248–277.
- Chantararat, S., Mude, A.G., Barrett, C.B. and Carter, M.R. (2012) ‘Designing index-based livestock insurance for managing asset risk in northern Kenya’, *The Journal of Risk and Insurance* 80(1): 205–237.
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B. and Eklundh, L. (2004) ‘A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter’, *Remote Sensing of Environment* 91(3–4): 332–344.
- Elhorst, J.P. (2003) ‘Specification and estimation of spatial panel data models’, *International Regional Science Review* 26(3): 244–268.
- Herrero, M., Grace, D., Njuki, J., Johnson, N., Enahoro, D., Silvestri, S. and Rufino, M.C. (2013) ‘The roles of livestock in developing countries’, *Animal* 7(S1): 3–18.
- Jack, W. and Suri, T. (2011) *Mobile money: The economics of M-PESA*, NBER working paper no. 16721, National Bureau of Economic Research, Cambridge, MA.
- Kelejian, H.H. and Prucha, I.R. (2007) ‘HAC estimation in a spatial framework’, *Journal of Econometrics* 140(1): 131–154.

- LeSage, J. and Pace, R.K. (2004) 'Models for spatially dependent missing data', *The Journal of Real Estate Finance and Economics* 29(2): 233–254.
- LeSage, J. and Pace, R.K. (2009) *Introduction to Spatial Econometrics*, Boca Raton, FL: CRC Press.
- Lybbert, T.J., Barrett, C.B., Desta, S. and Coppock, D.L. (2004) 'Stochastic wealth dynamics and risk management among a poor population', *The Economic Journal* 114(498): 750–777.
- Mude, A.G., Chantarat, S., Barrett, C.B., Carter, M.R., Ikegami, M. and McPeak, J. (2012) 'Insuring against drought-related livestock mortality: Piloting index-based livestock insurance in northern Kenya', in E. Makaudze (ed.) *Weather Index Insurance for smallholder Farmers in Africa – Lessons Learnt and Goals for future*, conference proceedings, Stellenbosch: African Sun Media, pp. 49–72.
- Steinfeld, H., Gerber, P., Wassenaar, T., Castel, V., Rosales, M. and de Hann, C. (2006) *Livestock's Long Shadow: Environmental Issues and Options*, Rome, Italy: FAO.
- Vrieling, A., Meroni, M., Shee, A., Mude, A., Woodard, J.D., de Bie, K. and Rembold, F. (2014) 'Historical extension of operational NDVI products for livestock insurance in Kenya', *International Journal of Applied Earth Observation and Geoinformation* 28(May): 238–251.
- Wang, W. and Lee, L.F. (2013a) 'Estimation of spatial autoregressive models with randomly missing data in the dependent variable', *The Econometrics Journal* 16(1): 73–102.
- Wang, W. and Lee, L.F. (2013b) 'Estimation of spatial panel data models with randomly missing data in the dependent variable', *Regional Science and Urban Economics* 43(3): 521–538.
- Woodard, J.D. (2008) "Three essays on systemic risk and rating in crop insurance markets", PhD Dissertation, University of Illinois.
- Woodard, J.D. and Garcia, P. (2008a) 'Basis risk and weather hedging effectiveness', *Agricultural Finance Review* 68(1): 99–117.
- Woodard, J.D. and Garcia, P. (2008b) 'Weather derivatives, spatial aggregation, and systemic insurance risk: Implications for reinsurance hedging', *Journal of Agricultural and Resource Economics* 33(1): 34–51.
- Woodard, J.D., Schnitkey, G.D., Sherrick, B.J., Lozano-Gracia, N. and Anselin, L. (2012) 'A spatial econometric analysis of loss experience in the U.S. crop insurance program', *The Journal of Risk and Insurance* 79(1): 261–286.
- Woodard, J.D. and Sherrick, B.J. (2012) 'Estimation of mixture models using cross-validation optimization: Implications for crop yield distribution modeling', *American Journal of Agricultural Economics* 93(4): 968–982.

About the Author

Joshua Woodard is assistant professor and Zaitz Faculty Fellow, Dyson School of Applied Economics and Management, Cornell University. He was responsible for primary manuscript drafting, model and technical concept development, and data analysis.

Apurba Shee is a researcher at the International Food Policy Research Institute. He was IBLI programme manager and contributed to the drafting, review and editing of the manuscript and IBLI concept development, and provided livestock mortality data.

Andrew Mude is a researcher at the International Livestock Research Institute. He was IBLI programme manager, and contributed to the review and editing of the manuscript and IBLI concept development, and provided livestock mortality data.