Development of Meso-Level Index Based Livestock Insurance in Kenya

The Third International Agricultural Risk, Finance, and Insurance Conference

June 22nd, 2014

Joshua Woodard, Cornell University, USA
Apurba Shee, International Livestock Research Institute, KENYA
Andrew Mude, International Livestock Research Institute, KENYA.

Index Based Livestock Insurance (IBLI) in Kenya

- Pastoral livestock production is the main livelihood of over three million people in northern Kenya's arid and semi-arid lands
- Index Based Livestock Insurance (**IBLI**) is a type of weather-linked insurance to safeguard livestock against drought
- Uses an index based on remotely sensed Normalize Difference Vegetation Index (**NDVI**) observations as basis for indemnification.

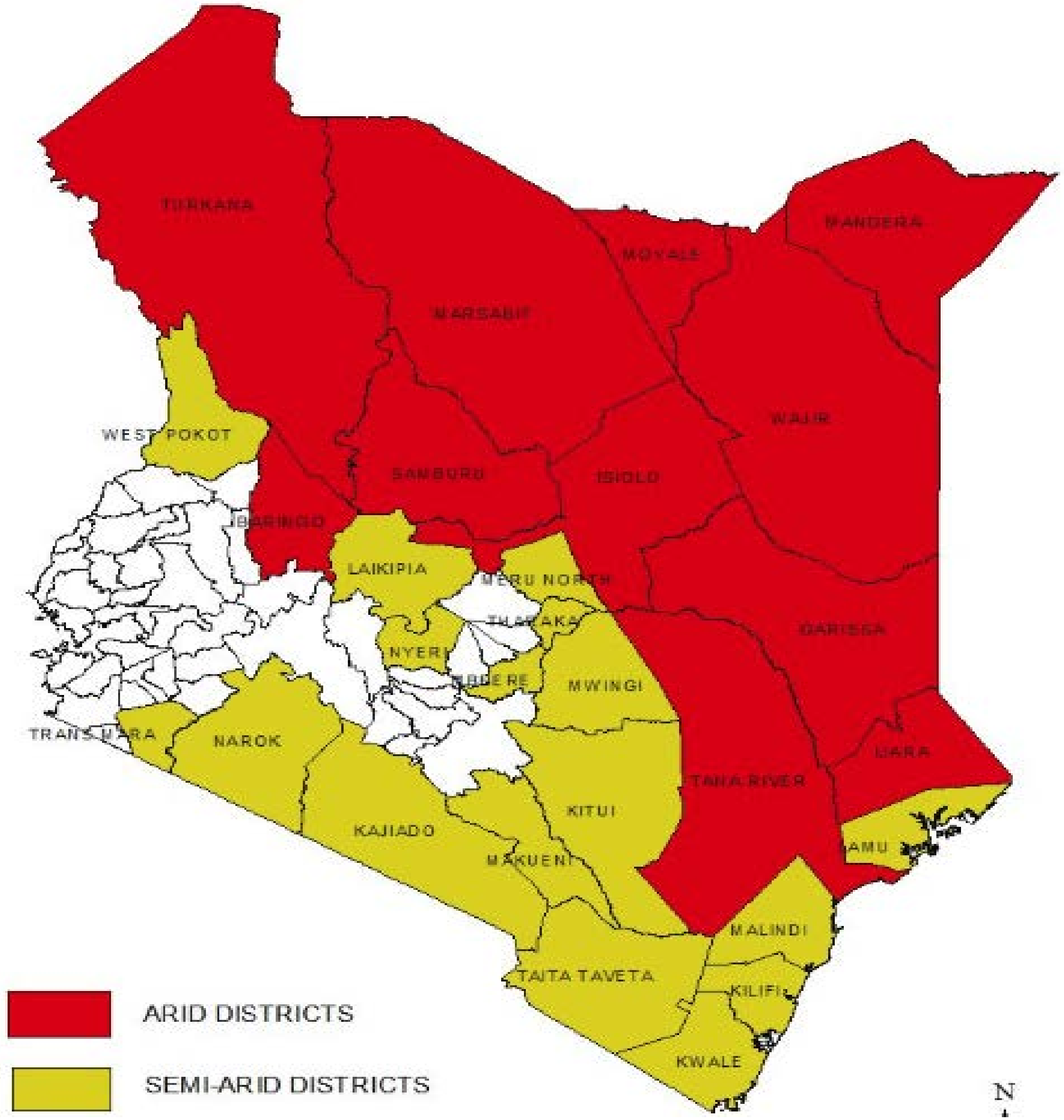
Index Based Livestock Insurance (IBLI) in Kenya

- **IBLI** was commercially launched in the Marsabit district in Kenya in January 2010 and in the Borana Zone in southern Ethiopia in July 2012 under the direction of the International Livestock Research Institute (**ILRI**)
- Currently, efforts underway to develop in 103 additional divisions
- Given implementation challenges of micro insurance, talks have begun to investigate viability of meso-level structures sold instead to County governments or aid organizations

IBLI Market and Development Model

- The International Livestock Research Institute (ILRI) developed IBLI, but is delivered, reinsured, and underwritten by private insurers and reinsurers
- Partners: USAID, DFID, World Bank, IFPRI, Cornell, Syracuse, University of Wisconsin, various banks, insurance companies, and Kenyan government ministries
- Capacity building:
  - IBLI development and premiums are currently subsidized, but long-term intent is for market to reach a critical mass and operate as a free but regulated unsubsidized market in host country
- Poverty Traps, Basis risk, and Systemic Risk

IBLI expansion areas in Kenya



Why Micro Index Insurance? The Standard Story

- Moral hazard and other asymmetric information considerations
- Systemic risk, self-insurance, and social insurance, poverty traps
- Somewhat predictable performance in long-run, lower operational risk
- Low administrative and claim adjustment costs
- Difficult to insure live animals, and/or otherwise difficult to adjust on-farm claim

Why Micro Index Insurance? The other side of story

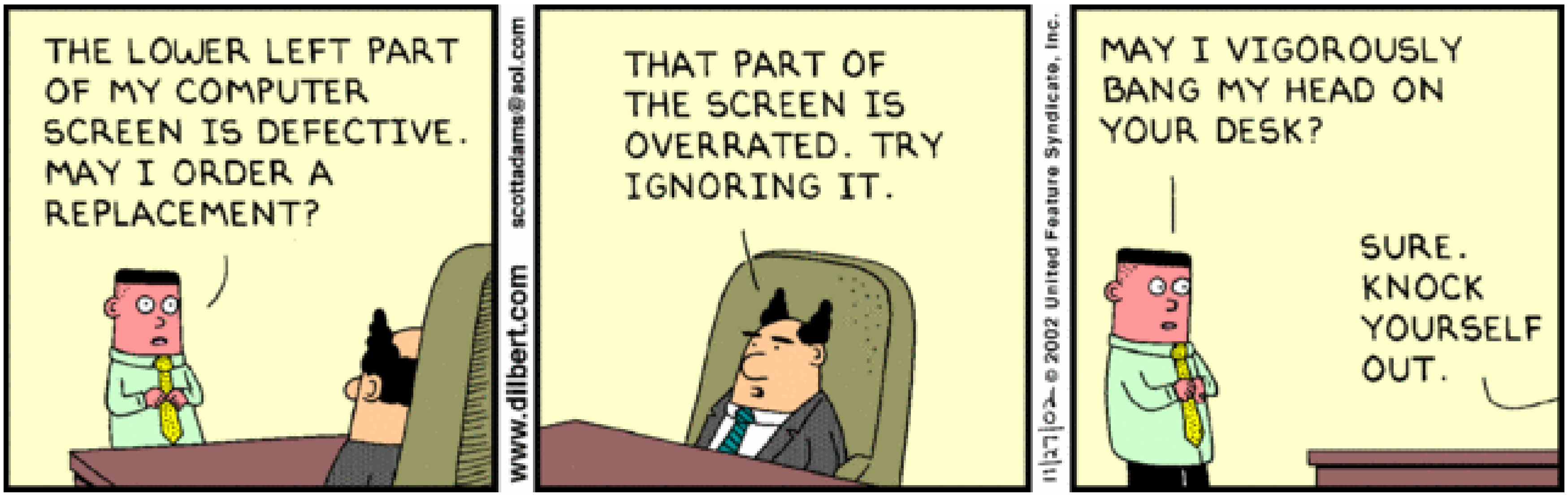
- Very expensive to educate potential insureds
- No delivery infrastructure
- High basis risk for individual covers
- High risk of trust issues as a result of basis risk
- Capital constraints have detrimental effect on uptake

Some Questions

- What types of products are best suited for fitting the developing country operating environment?
- Is **micro index** insurance the only/most viable solution, or is there a way to leverage other systems in the delivery of credit and risk management solutions?
- What is the role of transaction and information costs, etc.?
- How to deal with and communicate **basis risk**?

A question to contemplate...

- Basis risk is typically high for most existing micro-index structures...
- So, we should drill down further then and further refine estimates and indexes to make them even more micro to eliminate basis risk, right?
- Put another way...is basis risk typically going to be lower if we are modeling very exact, low levels of aggregation, or modeling at very large scales?


THE LOWER LEFT PART OF MY COMPUTER SCREEN IS DEFECTIVE. MAY I ORDER A REPLACEMENT?
THAT PART OF THE SCREEN IS OVERRATED. TRY IGNORING IT.
MAY I VIGOROUSLY BANG MY HEAD ON YOUR DESK?
SURE. KNOCK YOURSELF OUT.
www.dilbert.com scottadams@aol.com
11/27/02 © 2002 United Feature Syndicate, Inc.


THE Basis risk PART OF MY Micro-index insurance IS DEFECTIVE. MAY I market something _else?
scottadams@aol.com
www.dilbert.com
THAT PART OF Index insurance IS OVERRATED. TRY IGNORING IT.
© 2002 United Feature Syndicate, Inc.
MAY I VIGOROUSLY BANG MY HEAD ON YOUR DESK?
SURE. KNOCK YOURSELF OUT.

Conceptual Model

- Spatial aggregation and index risk.
- Suppose

$$Y_{t,k} = \alpha_k + f_k(\mathbf{W}_{t,k}) + \varepsilon_{t,k} \qquad (1)$$

- Summing across locations, *k*, gives

$$E[\sum_k Y_{t,k}] = \sum_k \alpha_k + E[\sum_k f_k(\mathbf{W}_{t,k})] + E[\sum_k \varepsilon_{t,k}] \qquad (2)$$

and

$$Var[\sum_k Y_{t,k}] = Var[\sum_k f_k(\mathbf{W}_{t,k})] + Var[\sum_k \varepsilon_{t,k}] + Cov[\sum_k f_k(\mathbf{W}_{t,k}), \sum_k \varepsilon_{t,k}] \qquad (3)$$

Conceptual Model (cont.)

- Proposition: If weather effects, $f_k(\mathbf{W}_{t,k})'s$, are relatively more *positively* correlated across space than are other yield effects, $\varepsilon_{t,k}'s$, then more variation in peril may be attributed to the weather effects at larger levels of spatial aggregation.
- Thus, indexes may be more effective in hedging production exposures at larger levels of aggregation.

Conceptual Model (cont.)

- Consider an extreme case for which the following conditions hold:

$$Corr[f_k(\mathbf{W}_{t,k}), f_j(\mathbf{W}_{t,j})] = 1 \qquad \forall j,k$$

$$Corr[\varepsilon_{t,k}, \varepsilon_{t,j}] = -1 \qquad \forall j,k$$

$$Corr[f_k(\mathbf{W}_{t,k}), \varepsilon_{t,j}] = 0 \qquad \forall j,k$$

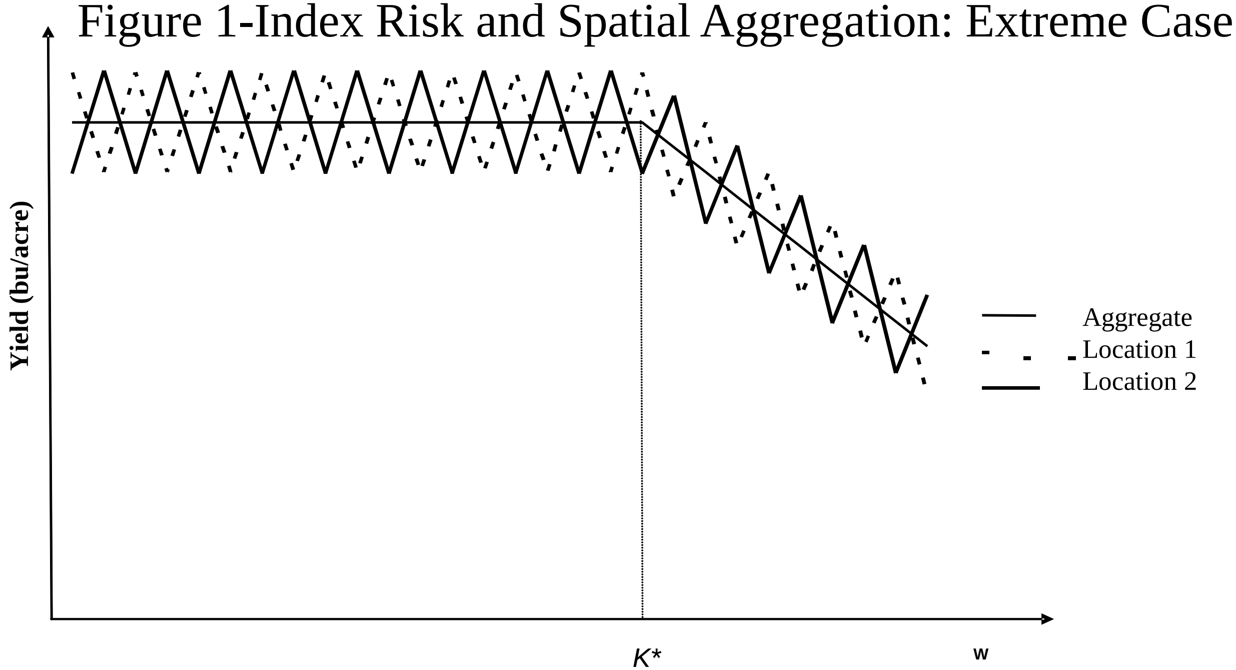
- Thus, equation (3) reduces to:

$$Var[\sum_k Y_{t,k}] = Var[\sum_k f_k(\mathbf{W}_{t,k})] \qquad (4)$$

Conceptual Model (cont.)

- In this extreme case, all variation in some peril (say yields) could be attributed to underlying
- For example, Figure 1 illustrates a case in which all index risk could be hedged using an call option on **W** with strike price *K**.

Figure 1-Index Risk and Spatial Aggregation: Extreme Case



Conceptual Model (cont.)

- Is this likely to be the case in crop production or other perils like mortality?
- Aggregating yields across locations should have a diversifying effect.
  - Farm-level yield risk tends to be greater than county level yield risk on average. Same is true for mortality.
- Weather events, particularly temperature, tend to be persistent (Namias) and highly spatially correlated (Jewson and Brix).
  - Temperature and precipitation tend to interact in extreme events, such as drought, causing a self-perpetuating event across both time and space.

Limits to accuracy

- Given the nature of highly disaggregated exposures, there are likely fundamental limits to the degree that we can reduce basis risk using index insurance schemes
- Many indexes are likely overfit, result in low out-of-sample performance.

Data and Product Dev. Issues

- Typical case for U.S. for product development
  - Extensive farm level and county level yield databases (GB's upon GB's of data), relatively full sets, typically back to at least the 1970's; extensive county, regional and national price, weather, soil, trade, input, cost data, etc.
  - IT staff, lawyers, accountants, sales & marketing people
- Scaling up IBLI in Kenya
  - Mortality survey info: 40% of divisions/time periods had no data; data available were manageable in size, but in some cases difficult to confirm accuracy
  - No on-the-ground weather data...and about 100 GB's of satellite vegetation data
  - No/few agents, mostly uninformed bankers/insurers
- How to adapt product design and estimation strategies to cases with weak or missing data?
- How to cope with lack of infrastructure?

What does this mean for micro-index insurance?

- In some ways a motivating factor...
  - Claim adjustment and delivery (kind of)
  - High systemic risk
- In others not so much...
  - Still hard/expensive to educate and deliver
  - Basis Risk, may be hard limits to precision

  - For example, is basis risk typically going to be lower if we are modeling very exact, low levels of aggregation, or modeling at very large scales?
  - Are we asking the right questions: meso-level products?

Construction and Estimation of Mortality Index Response Function

- Recall, in general we seek to estimate a function of livestock mortality on various measures of weather or drought (in our case NDVI), $\mathbf{X}$:

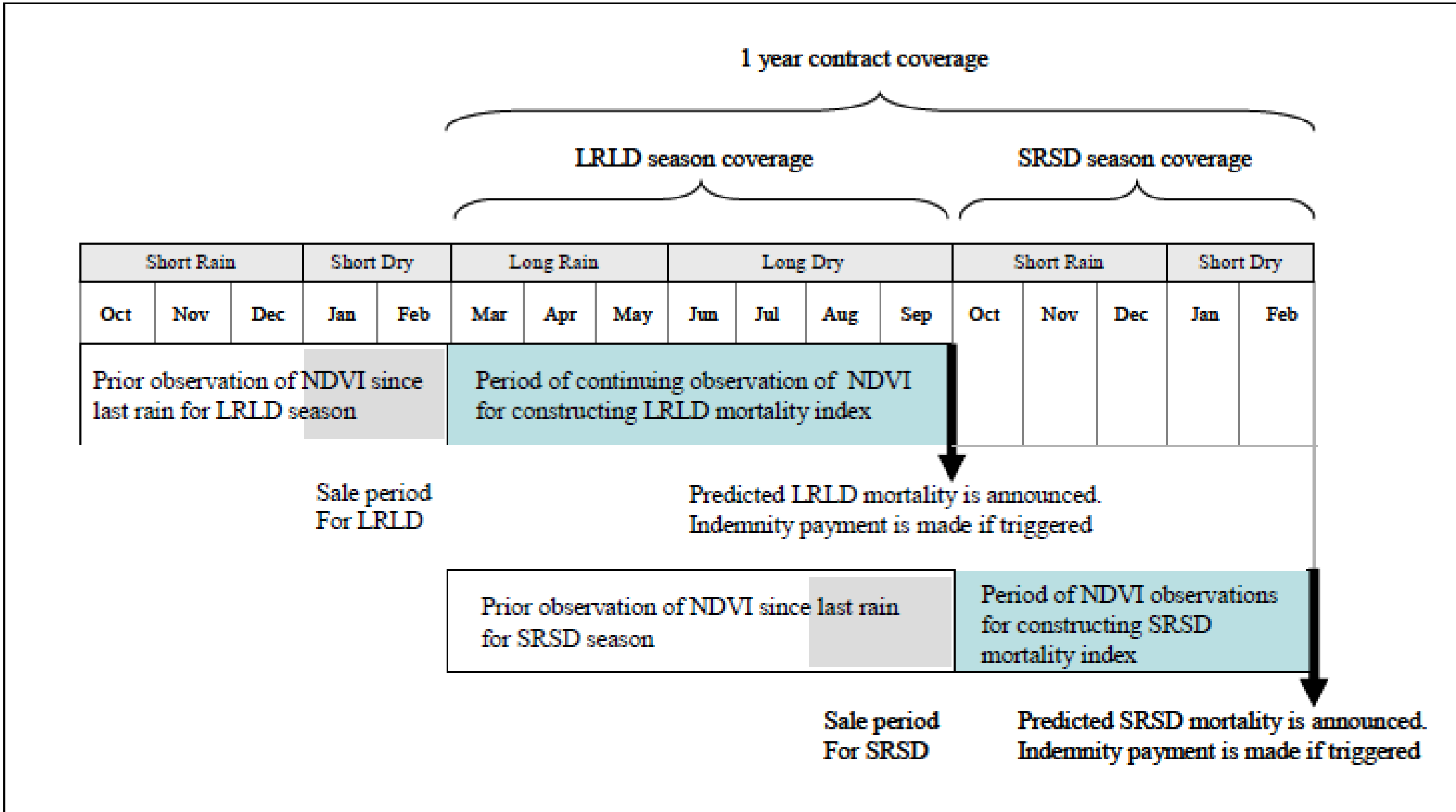
$$M = f(g(X), \theta)$$

- Weather and drought are going to tend to be highly spatially correlated
- Herds can move from location to location as well to some extent in response to drought
- How can we take weather, spatial interactions, and spatial autocorrelation into account in a somewhat systematic manner?

Data

- Mortality: ALRMP data from Kenya (N=~ 900k)
  - Monthly Household survey, up to 30 HH's per month/location
  - Many missing months/locations/years in data (~40% of division/time observations, 2001-2012
- Normalized Difference Vegetation Index (NDVI)
  - Measure of "green-ness"
  - NASA's eMODIS satellite data (Moderate Resolution Imaging Spectroradiometer), 250 meter/10 day resolution, processed to account for outliers/cloudcover
  - Period: 2001-2012; Augment rating with inter-calibrated data for 1981-2001 GIMMS/AVHRR data (Vrieling, Meroni, Shee, Mude, Rembold, and Woodard)
- Two insured easons in Kenya:
  - Short Rain-Short Dry (SRSD): Oct - Feb
  - Long Rain Long Dry (LRLD): Feb - Sept

Temporal/Seasonal Contract Structure



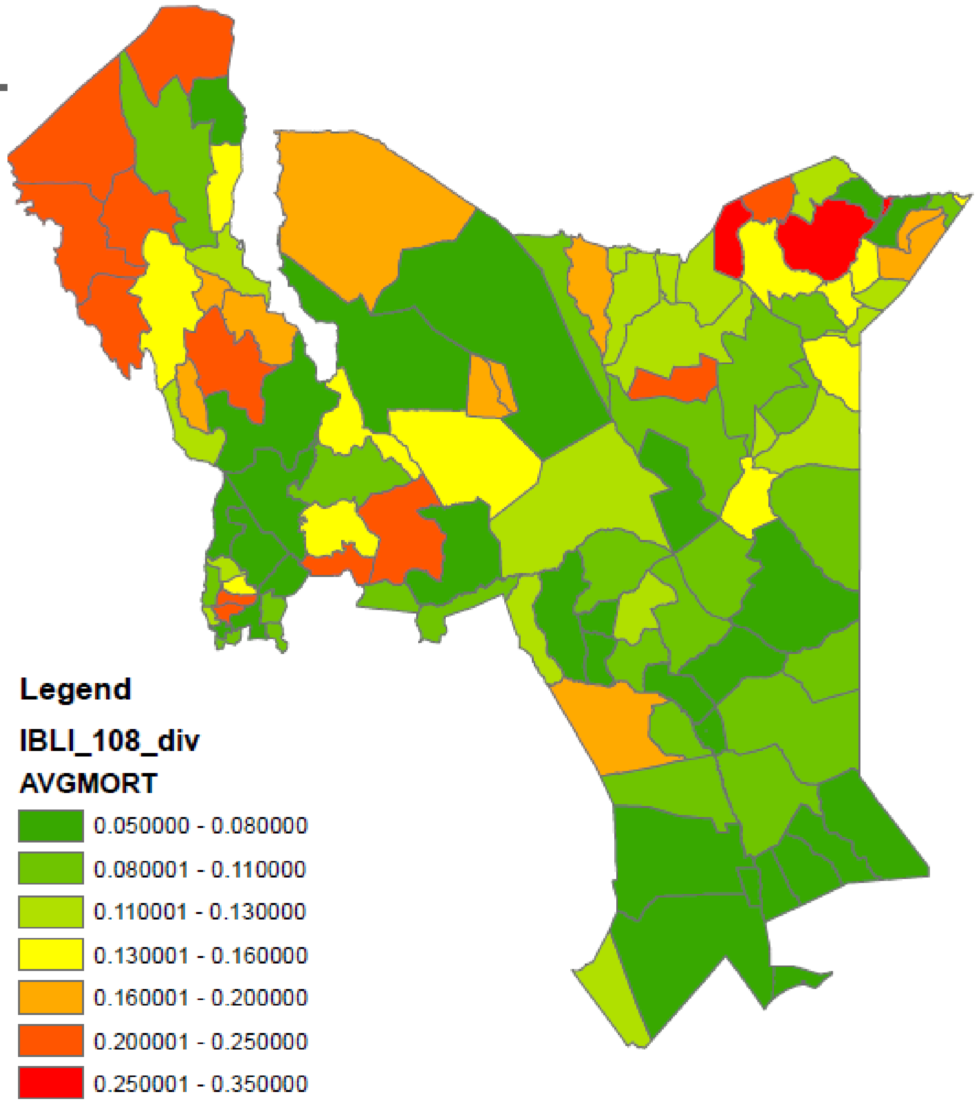
Source: Chantarrat, Barrett, and Mude (2012, JRI)

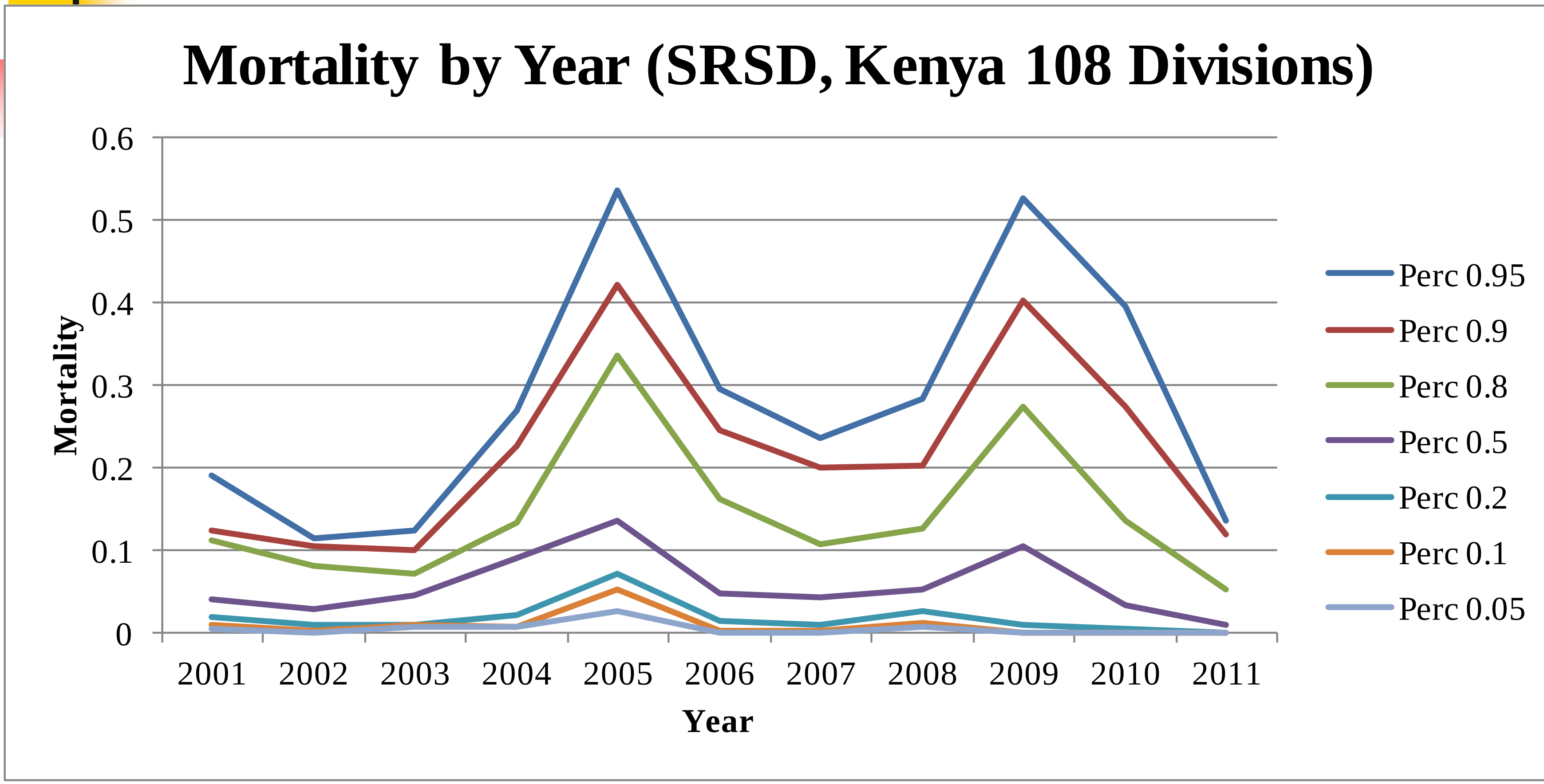
Mortality Index Construction

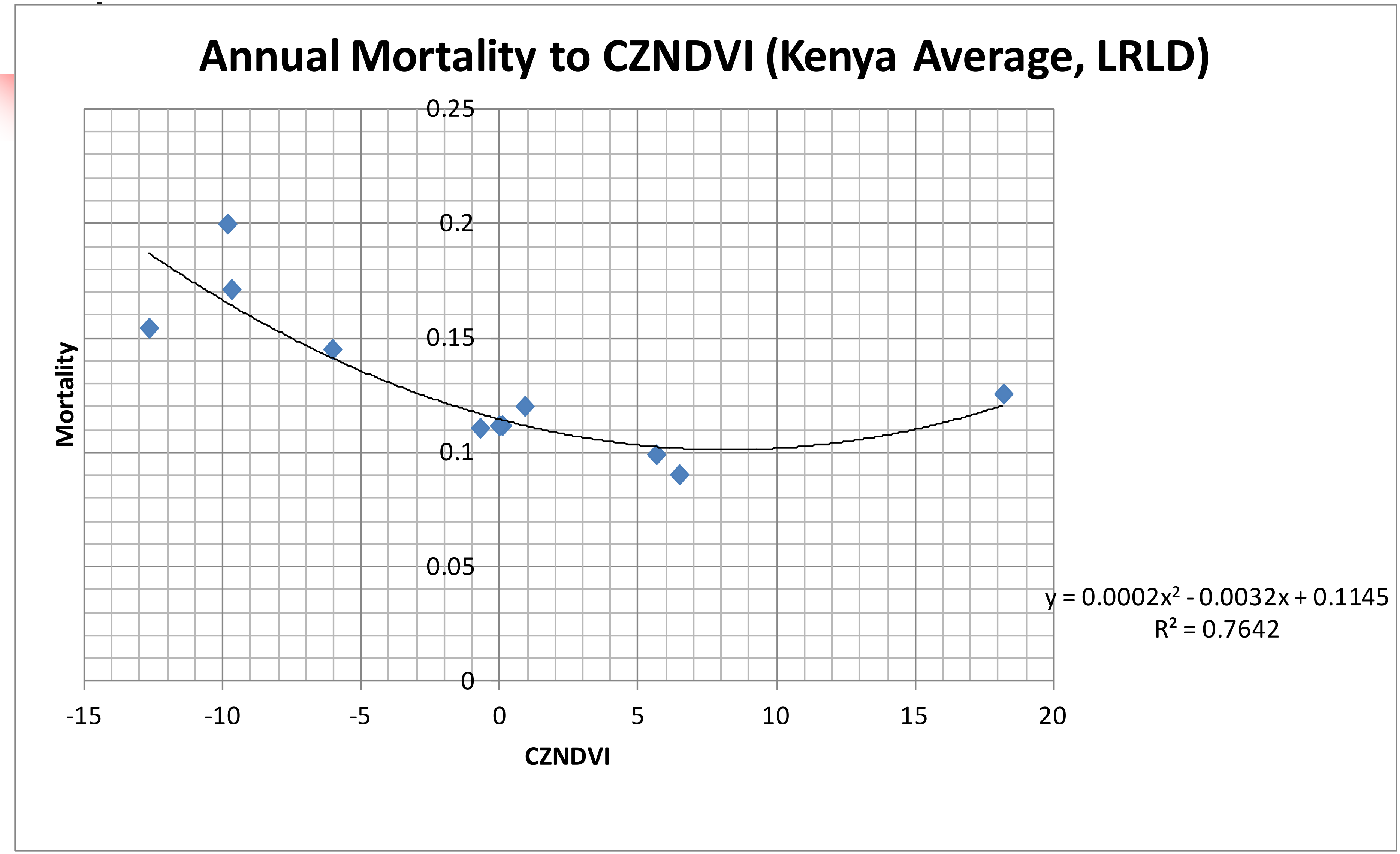
- Mortality is calculated for each season/year/division as the sum of herd mortality for all households across months in season, divided by the maximum herd size during the season.

$$M_{ls} \equiv \frac{\sum_{m \in s} \overline{H}_{mort,m}}{\operatorname*{Max}_{m \in s}\left(\overline{H}_{beg,m}\right)}$$

Kenya Average Herd Mortality Rates



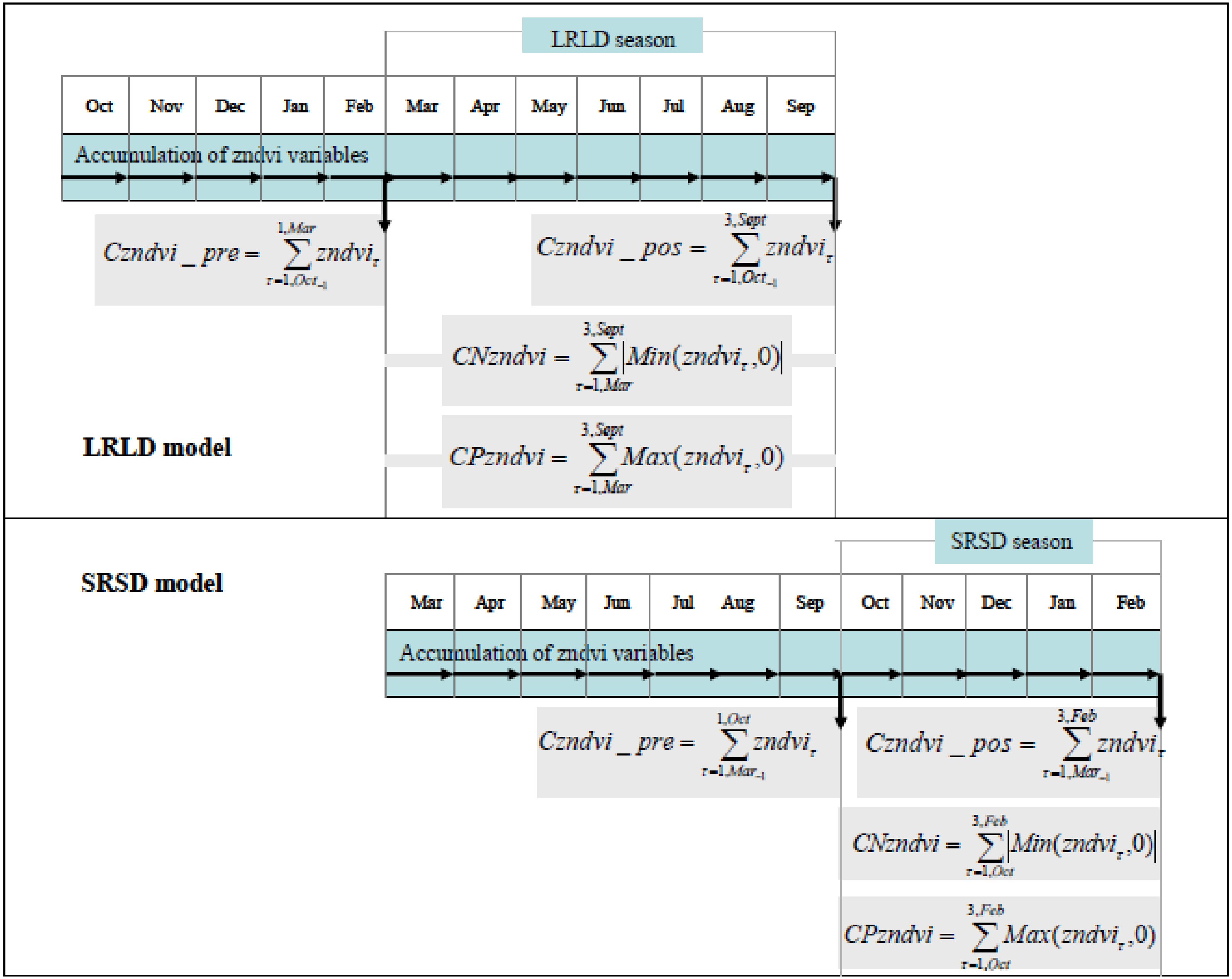

Mortality by Year (SRSD, Kenya 108 Divisions)
0.6
0.5
0.4
0.3
0.2
0.1
0
Mortality
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year
Perc 0.95
Perc 0.9
Perc 0.8
Perc 0.5
Perc 0.2
Perc 0.1
Perc 0.05


Annual Mortality to CZNDVI (Kenya Average, LRLD)
Mortality
0.25
0.2
0.15
0.1
0.05
0
-15
-10
-5
0
5
10
15
20
CZNDVI
y = 0.0002x² - 0.0032x + 0.1145
R² = 0.7642

NDVI Processing

- In general, NDVI is not a standardized measure for comparing across locations (due to differences in hydrology, elevation, etc.)
- Employ cumulative season z-indexes (CZNDVI) to measure deviations from “normal” for each division/year/season
- Procedure:
  - Pixels are spatially aggregated to division, and then for each 10 day period of the year, a mean and standard deviation are estimated
  - Each division/period is then assigned a z-value to measure its deviation from normal for that period of the year.
  - Division/period z-values are then summed together to arrive at a cumulative CZNDVI value for the year/season/division
  - Also employ smoothing procedures for outliers, cloud cover, etc.

Temporal/Seasonal Structure of NDVI Regressors



Source: Chantarrat, Barrett, and Mude (2012, JRI)

Spatial Econometric Models

- Standard Cross-Sectional Spatial Lag Model:

$$\mathbf{y}_{N\times 1} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\mathbf{y} - \rho\mathbf{W}\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$(I - \rho\mathbf{W})\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\mathbf{y} = (I - \rho\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$$

Standard Spatial Lag Model

- Note that The term $\mathbf{W}X$ is a spatially weighted average
- Essentially, in a spatial lag model, every observation is a function of itself and its neighbors.
- Noting that $(I - \rho \mathbf{W})^{-1}$ is non-sparse, every observation is also a function of the neighbors of its neighbors (with influence decaying with distance), and the magnitude of the effect is mediated by the spatial lag coefficient.
- Advantages: Allows for some degree of modeling spatial feedback effects (net) and spatial autocorrelation.
- Model is endogenous, so OLS is biased and inefficient.
- Several methods for estimating model, including MLE, GMM, and Bayesian models

Spatial Panel Models

Spatial Panel Model:

$$\mathbf{y}_{TN\times 1} = \rho(\mathbf{I_T} \otimes \mathbf{W_N})\mathbf{y} + \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

where $\mathbf{y}_{TN\times 1}$ and $\mathbf{X}$ are stacked observations (stacked by time).

Similar to the standard model, we then have:

$$[\mathbf{I_T} \otimes \mathbf{I_N} - \rho(\mathbf{I_T} \otimes \mathbf{W_N})]\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

$$\mathbf{y} = (\mathbf{I_T} \otimes \mathbf{I_N} - \rho(\mathbf{I_T} \otimes \mathbf{W_N}))^{-1}(\mathbf{X\beta} + \boldsymbol{\varepsilon})$$

Aggregation Level of the Mortality Index Response Function

- Recall, in general we seek to estimate a function of livestock mortality on various measures of weather or drought (in our case NDVI), $\mathbf{X}$:

$$M = f(g(W), \theta)$$

$$\mathbf{M}_{TN\times 1} = \rho(\mathbf{I_T} \otimes \mathbf{W_N})\mathbf{M} + \mathbf{X\beta} + \boldsymbol{\varepsilon}$$

- Suppose we have 1 weather variable. Should each location get its own parameter? Should each region get its own parameter? Or should we just impose the same parameter for the whole dataset?

Aggregation Level of the Mortality Index Response Function

- In our case, we have a country, which is made up of districts, which are made up of divisions (N = # divisions).
- Country/System Level Design matrix:

$$\mathbf{X}^{Country} = \begin{bmatrix} x_{1,1} \\ x_{2,1} \\ . \\ . \\ x_{N,T} \end{bmatrix}$$

Aggregation Level of the Mortality Index Response Function

- Suppose we have 2 districts in the country and three divisions, and divisions 1 and 2 belong to District 1, and Division 3 belongs to District 2.
- District Level Design matrix:

$$\mathbf{X}^{Dist} = \begin{bmatrix} x_{1,1} & 0 \\ x_{2,1} & 0 \\ 0 & x_{3,1} \\ x_{1,2} & 0 \\ x_{2,2} & 0 \\ 0 & x_{3,2} \\ . & . \\ . & . \\ x_{1,T} & 0 \\ x_{2,T} & 0 \\ 0 & x_{3,T} \end{bmatrix}$$

Aggregation Level of the Mortality Index Response Function

- Bias vs. Efficiency tradeoff…
- If underlying mortality processes are identical across locations, then the country level design matrix is preferred, but if the mortality functions (i.e., the parameters themselves) differ substantially, then may less biased using division or district level fixed effects in the design matrix.
- Similarly, if somewhat similar, then estimating individual response functions for each location will be inefficient.
- Alternatives to using regressor division/district fixed effects:
  - Spatial expansion (interaction of regressors with x-y coordinates)
  - Geographically Weighted Regression (GWR)
  - Others potentially, but all have same problem

Aggregation Level of the Mortality Index Response Function

- One could add a mixing parameter to each model, and then estimate jointly, but this increases the number of parameters substantially and may be very inefficient
  - Lower aggregation model will usually outperform in any case if component models estimated independently.
  - Number of parameters will be too large if estimated jointly in short panels if we have a large number of regressors.
- Also, may be difficult or impossible to solve if embedded in the more complicated model like a spatial panel model
- Alternative solution: estimate models independently, and then use cross validation to solve for the mixing weights.

In-Sample and Out-of-Sample Likelihoods

- In-Sample: $$L_k^{In} = \prod_{i=1}^{N} f_{k,i}(y_i, \theta_k^*(Y))$$
- Out-of-Sample: $$L_k^{Out} = \prod_{i=1}^{N} f_{k,i}(y_i, \theta_{k,i}^*(y_{-i}))$$

where we have data $Y = \{y_1, y_2, \ldots y_N\}$ and $y_{-i} = \{y_j \in Y : y_i \notin y_j\}$ and *K* candidate models

- This method of dropping one observation is known as "leave-one-out" cross validation

CV Model Mixing Procedure

- OOSL of Mixture Model:

$$L_{MIX}^{Out}(W,Y)=\prod_{i=1}^{N}\left[\sum_{k\in K}\left(w_k\cdot f_{k,i}(y_i,\theta_{k,i}^{*})\right)\right]$$

- Mixing Weights: $W=\left\{w_1,w_2,\ldots,w_S:\sum_{i\in K}w_i=1;w_i\geq 0\ \forall i\right\}$

- Optimal Weights: $W^{*}=\arg\max_{W}\left(L_{MIX}^{Out}(W,Y)\right)$

Some Observations

- Note that component model parameters are estimated independently of mixing weight
- Method is different than maximizing an in-sample likelihood to find a mixing weight and also different from EM
- Could also use different penalty functions in lieu of likelihood (e.g., MSE, LPM, etc.) within framework

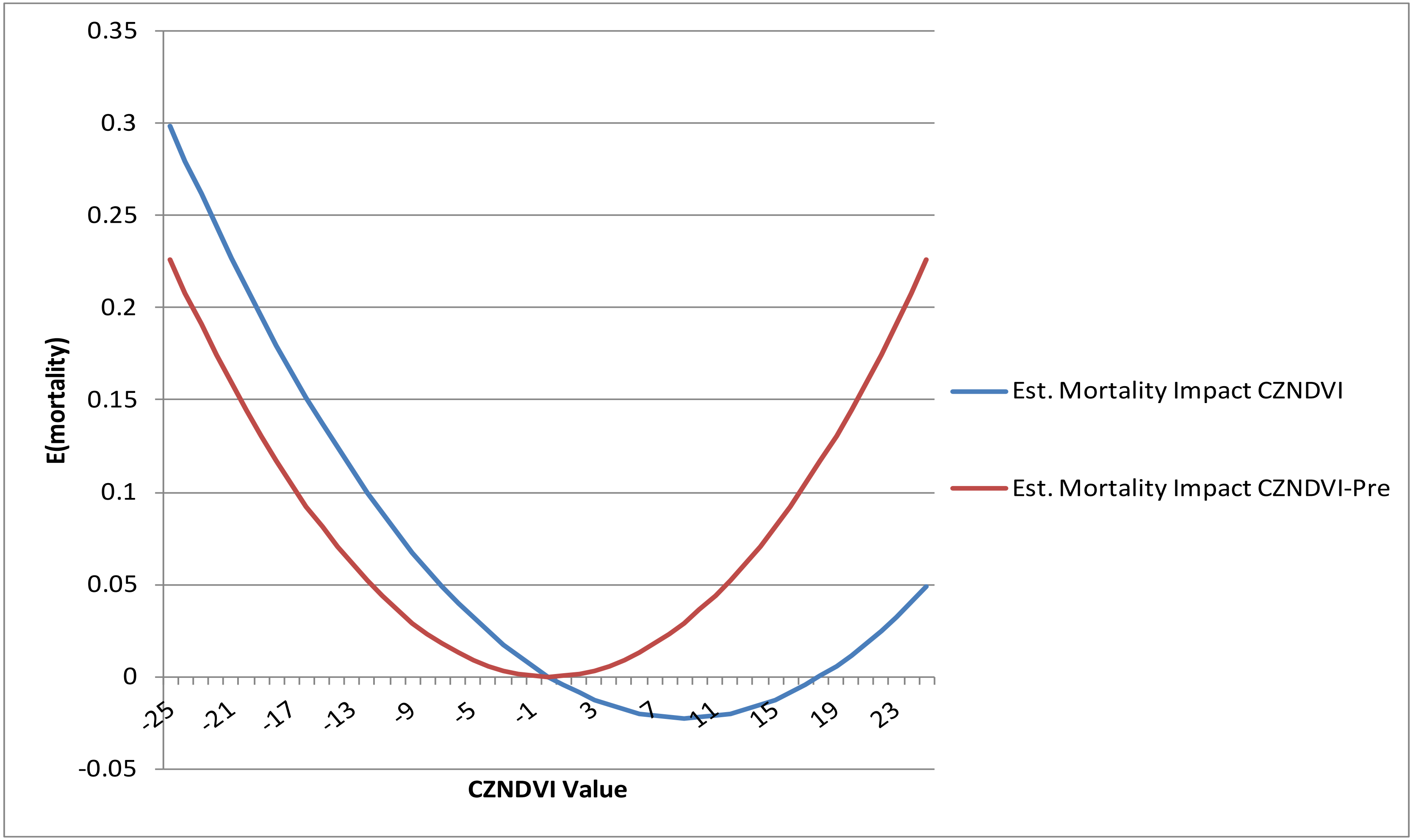
Results-Spatial Panel Model for Alternative Response Function Parameter Aggregations

Season = LRLD	System Level		District Level		Division Level	
ρ	0.530	***	0.455	***	0.598	***
pre-CZNDVI	-0.00292	***	-0.00412	†	-0.00557	†
preCZNDVI-sq	0.00037	***	0.00044	†	0.00060	†
CZNDVI	-0.00280	***	-0.00281	†	-0.00388	†
CZNDVI-Sq	0.00020	***	0.00012	†	0.00020	†
$pre\text{-}CZNDVI \times CZNDVI^{-}$	0.00076	***	0.00152	†	0.00140	†
$CZNDVI^{+} \times CZNDVI^{-}$	0.00007		-0.00192	†	-0.00024	†
R-squared	0.5013		0.5391		0.8497	
sigma-sq	0.0076		0.0071		0.0023	
log-likelihood	1158.8292		1221.2485		1857.5948	

*** Denotes significance at the 1% level.

† Denotes the average parameter values for the variable across all districts/divisions.

Estimated Average Mortality Response Rate to CZNDVI Measures

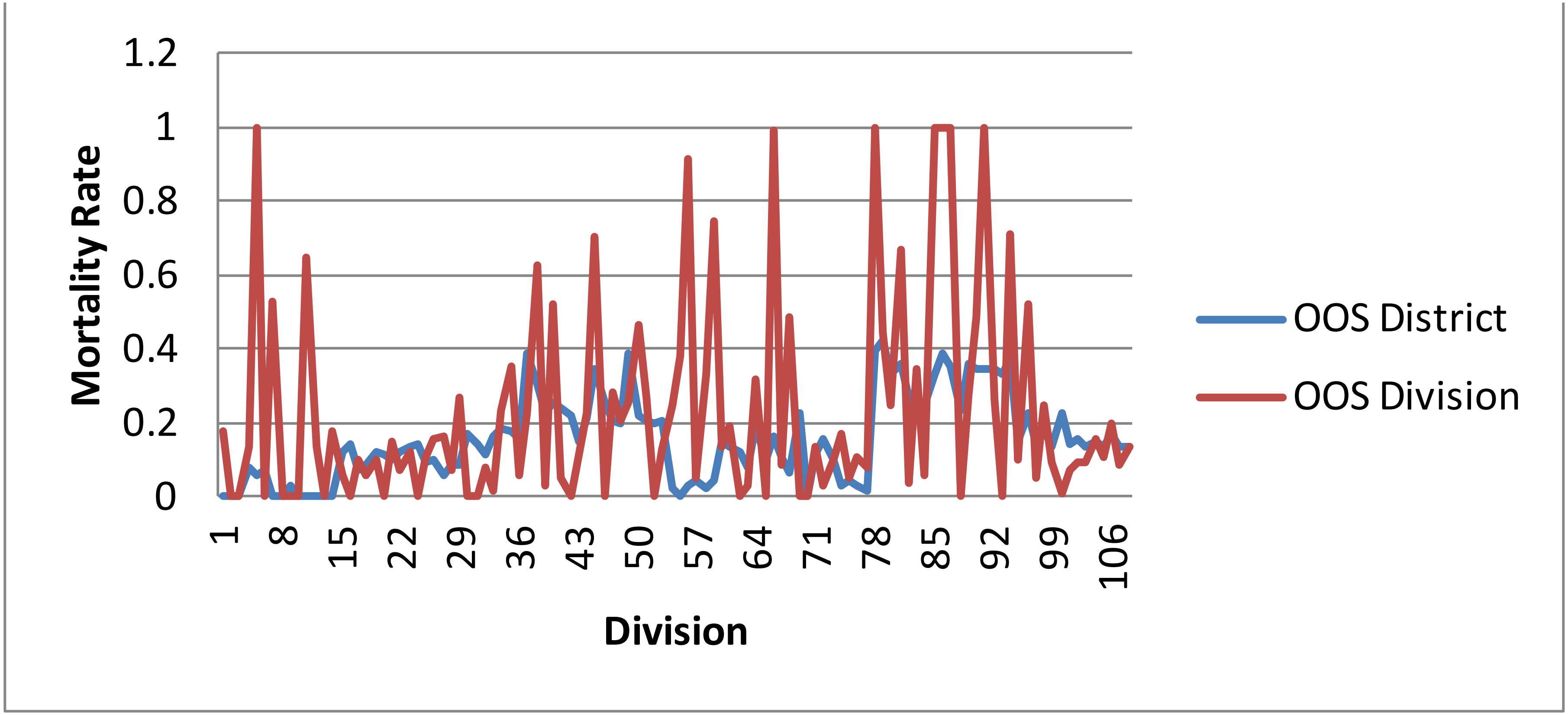


Optimal OOS Mixing Weights (Average-All Divisions)

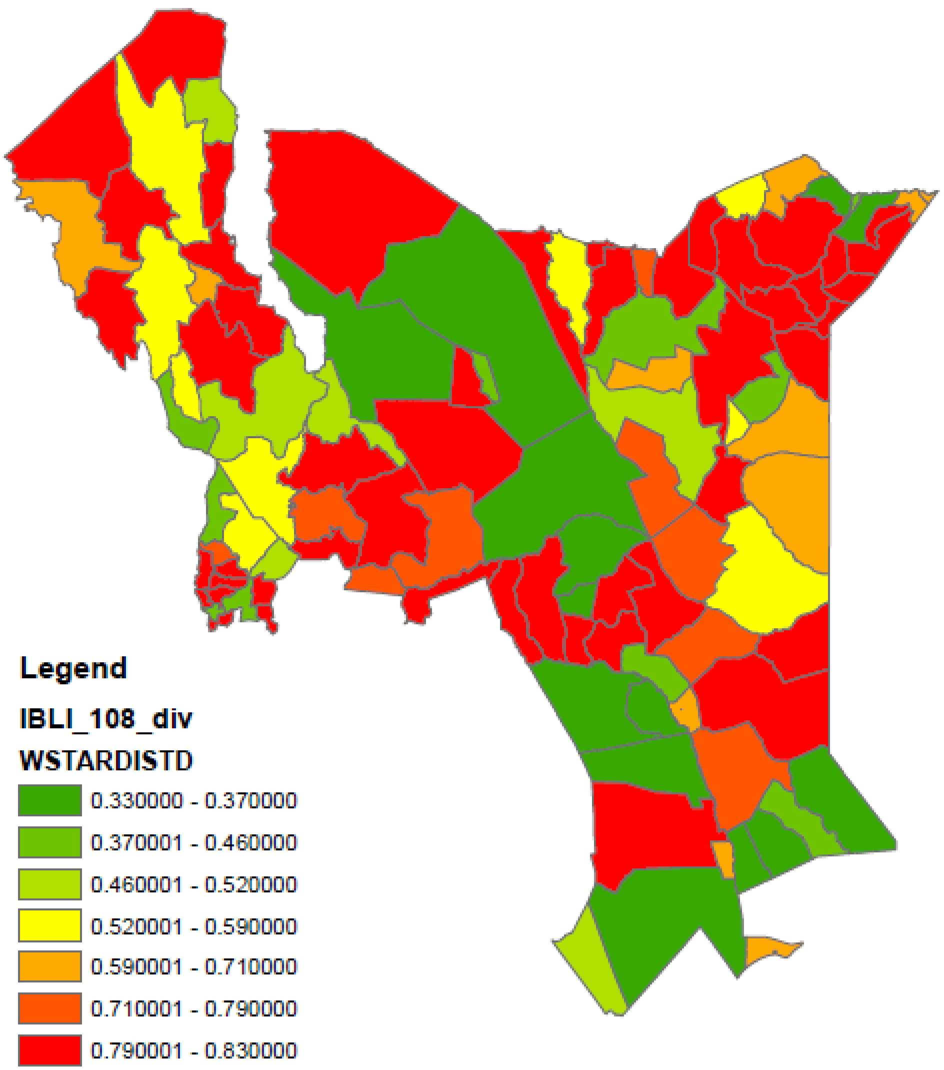
	System Model	District Model	Division Model
System Model		0.32	0.49
District Model	***0.68***		***0.67***
Division Model	0.51	0.33	

Note: Value reported is optimal weight on model in row.

Out of Sample Forecasts



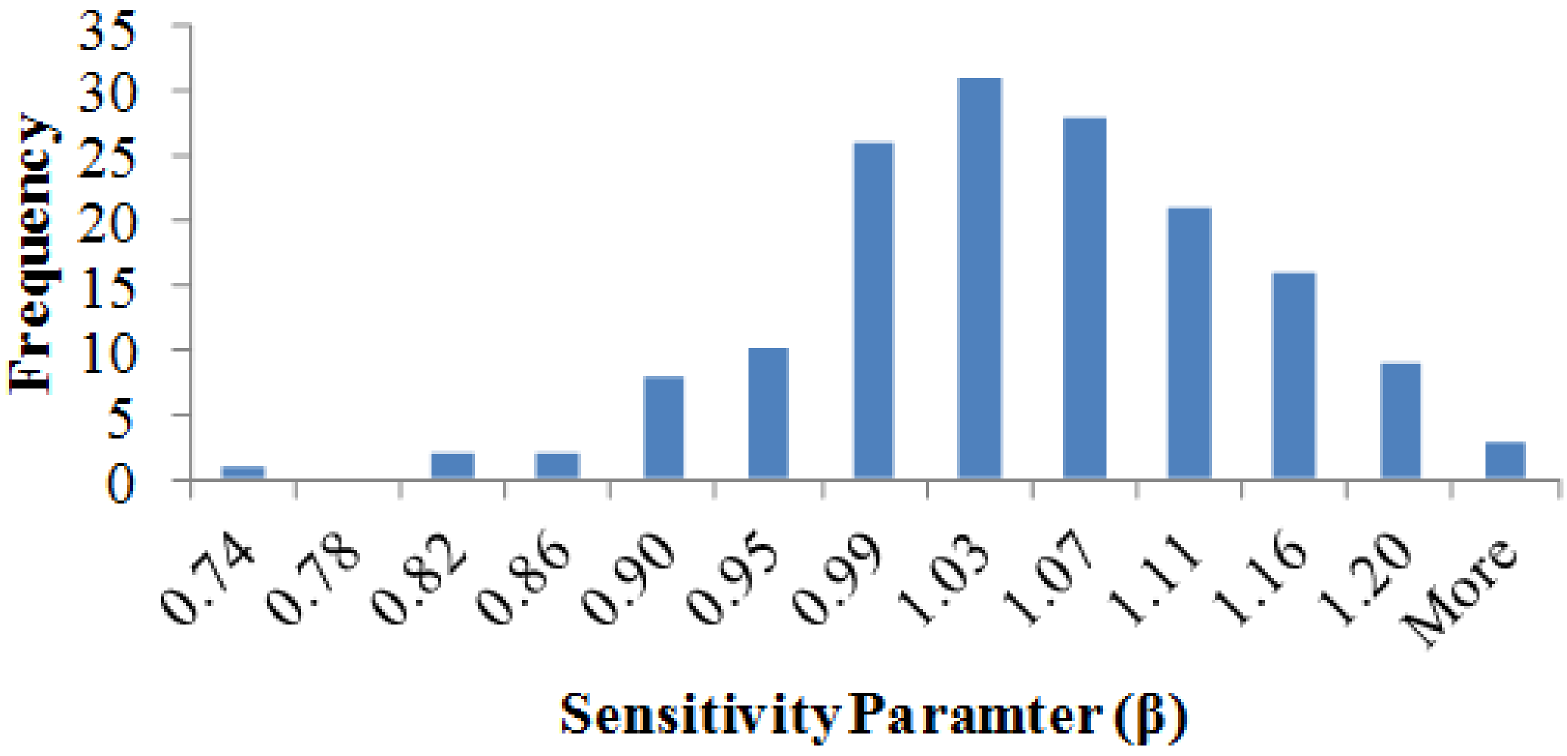
Optimal District Model Mixing Weight (District-Division Mixture)



Basis Risk Analysis Results and Decomposition of Idiosyncratic vs. Systemic Risk, Central Marsabit

	Estimated Index (RC Model)	*Ideal Index (1 Unit)*
α *(mean)*	-0.073	0
β *(mean)*	1.029	1
α *(st. dev.)*	0.041	-
β *(st. dev.)*	0.089	-
SST	21.77	21.77
SSE Total	12.48	14.71
SSE Overpay	5.19	5.50
SSE Underpay	7.29	9.21
FUV (B) Total	57.34%	67.57%
FUV (B) Underpayments	23.85%	25.27%
FUV (B) Overpayments	33.49%	42.30%
R^2	42.66%	32.43%

Histogram of Estimated Household Sensitivity Parameters-Panel RC Model Validation



Results (cont.)

- Results support proposed aggregation effect.
  - That is, better hedging opportunities may exist at higher levels of spatial aggregation.
  - Indicates that risk is relatively more systemic at higher levels of aggregation.
- Strong non-linear effects
- Out of sample performance much better for models aggregated at a higher level, even when the exposure itself being modeled is disaggregated

Practical Implications

- Perhaps more feasible to sell and educate one centralized institution (such as a county government or aid organization) on the use and functioning of contract, than it is to go out and deliver to individuals
- Perhaps better to leverage existing aid delivery networks, rather than try to replace them (at least for starters)

Conclusion

- Establishes the basis for the spatial aggregation effect in index insurance applications
- Suggest weather derivatives and index insurance will typically have lower basis risk at higher levels of aggregation and may be more useful for aggregators of risk
- Lower transaction/delivery costs, lower risk of trust issues

Questions?