Basis Risk and the Welfare Gains from Index Insurance: Evidence from Northern Kenya

Jensen, Nathaniel D. and Barrett, Christopher B. and Mude, Andrew G.

Cornell University, International Livestock Research Institute

September 2014

Basis Risk and the Welfare Gains from Index Insurance: Evidence from Northern Kenya

Nathaniel D. Jensen*†, Christopher B. Barrett† and Andrew G. Mude‡

† *Cornell University, Ithaca, NY, USA;* ‡ *International Livestock Research Institute, Nairobi, Kenya*

September 2014

A revised version of this working paper has since been published as the following:
Jensen, N. D., Barrett, C. B., & Mude, A. G. (2016). Index Insurance Quality and Basis Risk: Evidence from Northern Kenya. *American Journal of Agricultural Economics*, 98 (5): 1450-1469, doi:10.1093/ajae/aaw046

Abstract: Index insurance products circumvent many of the transaction costs and asymmetric information problems that obstruct provision of low value conventional insurance policies in developing countries. Recent years have seen tremendous growth in index insurance pilots in developing countries, but there has been little progress in our understanding of the quality of those products. Basis risk, or remaining uninsured risk, is a widely recognized, but rarely measured drawback of index insurance that carries significant implications for the quality of any such product. This research uses a rich longitudinal household dataset to examine basis risk associated with an index based livestock insurance (IBLI) product available to pastoralists in northern Kenya since 2010. We find that IBLI coverage reduces downside risk for most households when purchased at actuarially fair premium rates and has net utility benefits for most even at commercial rates. Examining the components of basis risk, we find that IBLI reduces exposure to covariate risk due to high loss events by an average of 62.8%. The benefits of reduced covariate risk exposure are relatively small, however, due to high exposure to seemingly mostly random idiosyncratic risk, even in this population often thought to suffer largely from covariate shocks. Depending on covariate region, IBLI policy holders are left with an average of between 62.3% and 76.7% of their original risk due to high loss events. This research underscores the need for caution when promoting index insurance as a tool for reducing exposure to risk and the importance of *ex post* product evaluation.

*Corresponding author: Nathaniel D. Jensen ndj6@cornell.edu.

The survey and data described in this research was collected by a collaborative project of the International Livestock Research Institute, Cornell University, the BASIS Research Program at the University of California at Davis, and Syracuse University. Data collection was made possible, in part, by support provided by the generous funding of the UK Department for International Development, the Australian Department of Foreign Affairs and Trade and the Agriculture and Rural Development Sector of the European Union through DfID accountable grant agreement No: 202619-101, the UK Department for International Development through FSD Trust Grant SWD/Weather/43/2009, the United States Agency for International Development grant No: EDH-A-00-06-0003-00, the World Bank's Trust Fund for Environmentally and Socially Sustainable Development Grant No: 7156906, and the CGIAR Research Programs on Climate Change, Agriculture and Food Security (CCAFS) and Dryland Systems. All views and interpretations expressed in this document are those of the authors and not necessarily those of the supporting or cooperating institutions. The authors would like to thank the all persons who contributed to this survey and its various sub-processes, especially Andrew Mude, Chris Barrett, Michael Carter, John McPeak, Sommarat Chantarat, Mohamed Shibia, Oscar Naibei, Anne Gesare, Philemon Chelanga, Eddy Chebelyon, Diba Kone Galgallo, Brian Dillon, Sarah Janzen, Nishith Prakash, Samuel Mburu, Robert Ouma, Amy Kahn, the interviewees, enumerators, and supervisors. We also thank Michael Carter, George Jakubson, Sarah Janzen, Marc Rockmore, Vincent Smith, and Josh Woodard for their useful comments.

1. **Introduction**

Where insurance providers lack access to accurate, historic individual-level data, monitoring behavior is difficult, and/or claims validation is costly, conventional loss indemnity insurance is often cost prohibitive and thus inaccessible to consumers. These barriers exist and are exacerbated in the context of smallholder farmers and herders in developing countries that demand low value policies. One alternative to conventional policies—which provide indemnity payments based on verified individual losses—is to offer policies that provide indemnity payments based on an index related to average losses among groups of consumers. Such covariate insurance products eliminate the need to price polices at the individual level and reduce the costs of validating claims by using data on average, rather than individual, losses. In addition, as group size increases, the potential for moral hazard and the costs of monitoring for such behavior fall. In cases where an easily observed exogenous signal of covariate losses is available, index insurance policies can further reduce costs of estimating average losses and reduce the negative (positive) impacts of cross-sectional adverse selection and moral hazard on insurer profits (equilibrium premium rates).

Basis risk, or the risk to which an insured individual is still exposed, has been called "the most serious obstacle to the effectiveness of weather index insurance as a general agricultural risk management tool" (Miranda & Farrin 2012, p.48). Product design and remaining basis risk have been studied quite extensively within the agricultural finance and insurance literature in the context of index insurance (or weather derivatives) for crops in developed economies (e.g., Miranda 1991; Williams et al. 1993; Smith, Chouinard & Baquet 1994; Mahul 1999; Turvey 2001; Vedenov & Barnett 2004; Woodard & Garcia 2008). This research establishes that basis risk is often significant even for optimally designed index contracts offered to farmers in developed economies, but that even suboptimal index products can nonetheless offer a valuable tool for cost-effectively mitigating exposure to yield risk in such environments.

Recent years have seen a surge in the promotion and piloting of index insurance projects for agricultural households in developing countries. For example, in 2009 the International Finance Corporation and the World Bank jointly implemented the Global Index Insurance Facility (GIIF) to help grow and support index based insurance products in developing countries. By 2012, the GIIF was supporting projects insuring 228,000 clients for $USD 50.7 million in prospective indemnity payments (Global Index Insurance Facility, 2013). Unfortunately, most pilot projects have met with extremely low demand, even when premiums have been subsidized and extension efforts have been included. Basis risk is often cited as a likely cause of low demand (e.g., Hazell & Hess 2010; Miranda & Farrin 2012; Smith & Watts 2009) but the magnitude of this basis risk remains unknown.

To date, none of the studies associated with index insurance products in developing countries offer household—level estimates of basis risk. In fact, few studies explicitly include any measure of basis risk at all. The lack of empirical attention to basis risk is especially disturbing because without it, there is no guarantee that index insurance is risk reducing. In cases where an individual's idiosyncratic risk is high or if the index is inaccurate, index products can represent a risk increasing gamble rather than the risk reducing insurance they are advertised to offer. Discerning the magnitude and distribution of basis risk should be of utmost importance for organizations promoting index insurance products, lest they inadvertently peddle lottery tickets under an insurance label.

The dearth of empirical estimates of basis risk in developing countries has multiple reasons. First, longitudinal household data are required in order to identify the distribution of basis risk. Because administrative cost savings due to reduced data collection are a key selling point of index insurance, such data are commonly lacking. In addition, premiums are usually calculated using the expected indemnity rate of the index (rather than on observed household losses) and insurers face little risk from moral hazard, so there is little profit incentive to collect individual-level verification data.[1]

Second, because there are multiple measures of basis risk, it is not obvious which metric is most salient to potential consumers or to which aspects of basis risk insurance providers should pay most attention. For example, a high loss event with no indemnity payment is often cited as a worst case scenario for policy holders, but these "false negatives" may be due to idiosyncratic losses which contractually fall outside of the product's coverage. To complicate matters further, indemnity payments may improve the net expected outcome while increasing its variance by over-estimating (over-indemnifying) losses. Such events reduce the usefulness of mean-variance analysis, a method commonly used to examine risky choices. Finally, most index insurance policies use an index measured in units fundamentally different from the ultimate objective of insurance – stabilizing standards of living – as in the case of weather insurance contracts that aim to insure against crop loss, significantly complicating the estimation of basis risk.

Several authors have used clever approaches to approximate the impact of basis risk in the absence of direct basis risk estimates. Studying rainfall insurance in India, Mobarak and Rosenzweig (2012) exploit the likely reduction in the correlation of precipitation between two locations as the distance between the two increases, using perceived distance to rain gauge as a proxy for perceived basis risk. This perceived distance measure is intuitively sound, but requires that rain gauges correctly identify delayed monsoon onset (the insured risk) at their location and assumes that precision falls (basis risk increases) linearly and symmetrically with distance from the rain gauge. Giné, Townsend and Vickery (2008), also studying rainfall insurance for Indian crop farmers, use the proportion of a household's cultivated land that is planted with either castor or groundnut crops, the species used to generate the insurance policy parameters, as a proxy for basis risk. In theory, precipitation triggers set by the policy best reflect risks for groundnut and castor crops, while other crops' vulnerabilities are correlated with the index to a lesser degree. In that case, basis risk should increase as the proportion of planted castor and groundnuts

[1] The simplest method for determining actuarial premiums is to calculate the expected indemnity payment. This "burn rate" approach to pricing requires information on the average losses within the product area but not on cross-sectional heterogeneity within the area.

falls. Both studies find that basis risk has a statistically significant and negative impact on farmers' demand for insurance, but the use of proxies makes it difficult to assess the magnitude of basis risk.

Hill, Robles and Cebellos (2013), also studying weather insurance in India, find that price sensitivity of demand increases near the weather stations, where basis risk is presumed lower. Using that basis risk proxy as an indicator of product quality, the authors estimate that demand for higher quality products is more price sensitive than demand for low quality products. One interpretation of these findings is that high quality (i.e., low basis risk) index products are normal goods, and that low price elasticity of demand may be a signal of poor index quality due to high basis risk. This interesting finding begins to unpack the relationship between quality of the product and demand, an extremely relevant topic for insurance providers as they develop strategies aimed at increasing demand.

Other papers have used simulations, aggregate-level data, and/or experiments to examine basis risk (e.g., Breustedt, Bokusheva & Heidelback 2008; Clarke 2011; Dercon et al. 2014; Elabed et al. 2013; Leblois, Quirion & Sultan 2014; Norton, Turvey & Osgood 2012). Again, basis risk is consistently identified as a key factor in product quality and uptake, but little or nothing can be said about the relative magnitude or distribution of basis risk among households. Although basis risk is widely acknowledged as a potentially serious issue even as interest in index insurance has exploded globally, it remains remarkably under-researched.

The Index Based Livestock Insurance (IBLI) product was developed and commercially piloted among pastoralists in the Marsabit region of northern Kenya in 2010 (Chantarat et al. 2013). The IBLI index predicts livestock mortality rates using an innovative response function that was generated econometrically using historical data on household herd losses specifically with the objective of minimizing basis risk. If basis risk significantly limits the benefits from IBLI, one might naturally wonder whether other products, not designed to minimize basis risk, might suffer similar or worse shortcomings.

Because the IBLI index is measured in the same units as the insurable household losses, it (perhaps uniquely) allows for direct estimation of the magnitude and cross-sectional heterogeneity of basis risk. This paper uses a four-year household panel dataset, which includes eight distinct semi-annual seasons of index values and household-level loss data, in order to examine the magnitude and components of basis risk that pastoralists face with respect to IBLI. Using standard approaches that are often used to study index insurance in developed economies, we find that at unsubsidized, commercially loaded premium rates full IBLI coverage significantly increases variance in livestock survival rates by an average of 4.7% but improves skewness in survival rates by 45% (from -1.185 to -0.651). Restricting analysis to downside risk beyond the strike increases the ratio of households that benefit from IBLI and illustrates the vital role that premium rates play in determining the benefits of insurance. Utility analysis allows us to simulate the median willingness to pay rate, which is greater than the loaded unsubsidized premium rates in both contract regions.

We then extend the literature on basis risk by examining the components of basis risk and the factors that contribute to their heterogeneity. IBLI coverage reduces exposure of households to risk associated with large covariate shocks by an average of 62.8%. Although droughts, which represent insurable covariate risk, are the largest reported cause of livestock mortality there is considerable variation in livestock mortality in every season. After accounting for covariate losses, household continue to face 69.3% of their original downside risk. By design IBLI can do nothing about this remaining idiosyncratic risk. On average, IBLI coverage reduces exposure to downside risk by 30.9%, which reflects both under and over indemnification of covariate events as well as idiosyncratic losses.

Examining the ratio of covariate to total risk at various scales reveals considerable geographic heterogeneity. Covariate shocks represent only a small portion of households' risk portfolio in some locations, while in others the majority of livestock mortality is associated with covariate shocks. The degree of geographic heterogeneity in the relative importance of covariate shocks points towards regions where IBLI may be well suited and others where it may not offer an appropriate approach for reducing risk associated with livestock mortality. The idiosyncratic risk that index insured households continue to face is mostly the result of random, unobserved household characteristics and events, but is also positively associated with a higher household dependency ratio and income diversification away from livestock-related activities, both of which likely reflect reduced managerial attention to animal husbandry, as well as geographic location.

This paper links the established work on agricultural index insurance products in higher income economies with the emerging literature on index insurance in developing economies while also providing a benchmark for basis risk that is useful for all index products. More broadly, it underscores the dangers of assuming that cleverly designed financial instruments always perform as advertised. Given the considerable uninsured risk exposure faced by low-income rural households in the developing world, designing, implementing and evaluating risk management tools is a task of first order importance.

The rest of the paper is structured as follows. We begin with an examination of the components of basis risk in Section 2. Section 3 describes the context, the IBLI product, and data. Section 4 examines the impact that ILBI coverage has on the distribution of outcomes that households face. We then decompose basis risk into its various components in order to reveal which factors drive the product's imperfect performance and which are associated with idiosyncratic losses. We conclude in Section 5 with a discussion of the implications of our findings for IBLI and other index insurance products. Given the burgeoning interest in index insurance within the development, finance, and agricultural communities, and the glaring dearth of evidence on basis risk in these products, our findings offer a cautionary tale to researchers and practitioners alike.

2. **Basis Risk**

Insurance coverage that is priced to be actuarially fair necessarily has no impact on expected outcomes, thus the immediate welfare impacts of actuarially fair insurance are captured by changes to the higher order moments of the distribution of outcomes. Such a loss-indemnity insurance contract with no deductible weakly second-order stochastically dominates no insurance because it transfers resources from periods with good outcomes to periods with poor outcomes at no cost. A similar index insurance contract (i.e., actuarially fair with no deductible) with no basis risk does the same, intertemporally transferring at no cost to weakly stochastically dominate the no insurance alternative.

But such product designs are abstractions from the real world of commercially loaded (i.e., not actuarially fair) policies with deductibles (or, equivalently, non-zero strike levels) and basis risk. Of particular concern in this paper, most index insurance policy holders face some remaining basis risk. This has two potential sources: design risk due to differences between the index and the actual covariate risk it is meant to mimic, and idiosyncratic risk resulting from heterogeneity among individuals' losses within the same index region (Elabed et al. 2013).[2] Design risk results from imperfect index design. Idiosyncratic risk falls outside the scope of an index policy. It is an artifact of generalizing information and can only be changed by adjusting the index region.[3]

If we allow for either component of basis risk—heterogeneity between individuals or error in the index— there is no assurance that an index insurance product reduces risk exposure. Due to this positive probability of increases to risk, index insurance does not necessarily weakly second order stochastically dominate the no insurance alternative. That is, a risk averse individual may prefer no insurance over index insurance with the possibility of basis risk, even at actuarially fair premium rates.

Once overhead costs (loadings) are included, even loss indemnity insurance can be stochastically dominated by a no insurance state. In fact, the extremely high cost of monitoring and verification has made loss indemnity insurance loadings so high that it is nearly impossible to sustain commercially in many situations, such as to smallholder farmers or pastoralists in remote locations. It is specifically this dilemma that index insurance attempts to address by providing low cost insurance based on exogenous indicators of covariate shocks and indemnity payment schedules that require little (or no) verification.

Since most consumers face loaded premium rates and design risk is practically inevitable, arguably even optimal given costly data collection, this makes the social value of index insurance an intrinsically empirical question because there exist many contracts with design risk that could reduce risk. And, because individuals do not face identical losses, products may be risk increasing for some individuals while for others they are risk reducing. Put differently, index insurance with basis risk might be a targetable product. The welfare effect of index insurance contracts and the distributional profile of those effects among heterogeneous agents are thus inherently empirical questions. The existing literature has not yet explored these issues.

[2] It is possible that there are multiple sources of covariate risk that may themselves be uncorrelated. In those cases, an index insurance policy may provide coverage for only one such covariate risk or an aggregate of the covariate risks. But, sources of agricultural risk are often correlated. For example, livestock mortality from disease is likely correlated with mortality due to drought. In the case of IBLI, the policy index is predicted livestock mortality rate, and was constructed using data that included all reported causes of death.

[3] There is likely a trade-off between scale and data requirements so that reducing design risk by increasing scale is likely to require more data lest the quality of the index suffer. Moreover, the finer the scale, the greater the chance for asymmetric information problems associated with adverse selection and moral hazard to reemerge as problems. There may therefore be an optimal scale-quality optimum that is a function of spatial correlation of insured losses and the cost of data collection.

The remainder of this section develops a framework for examining basis risk. We deviate from the commonly used model for index insurance developed by Miranda (1991) in order to more clearly separate basis risk into its idiosyncratic and design components.[4] Section 4 draws on this framework to empirically examine those components to learn about the factors that contribute to each.

As a stylized example, let individual *i* living in a spatially defined division *d* experience losses in period *t* at rate $L_{i,d,t}$.[5,6] Large scale events such as drought or floods can generate losses across many individuals in the same area. Such covariate losses are reflected in $\bar{L}_{d,t}$, the average or covariate losses in area *d* at time *t.* An individual's losses can then be divided into covariate losses and a remaining idiosyncratic component ($L_{i,d,t} - \bar{L}_{d,t}$).

The variance in loss rate that an individual faces over time ($Var_t[L_{i,d,t}]$) is one metric of risk.[7] Similar to loss rate, an individual's risk can be decomposed into a covariate component, an idiosyncratic component, and the covariance between idiosyncratic losses and covariate losses ($Var_t[L_{i,d,t}] = Var_t[L_{i,d,t} - \bar{L}_{d,t}] + Var_t[\bar{L}_{d,t}] + 2 * cov_t[L_{i,d,t} - \bar{L}_{d,t}, \bar{L}_{d,t}]$). Therefore, as $L_{i,d,t} \to \bar{L}_{d,t}, Var_t[L_{i,d,t}] \to Var_t[\bar{L}_{d,t}]$. Alternatively, households that typically experience much less risk than their neighbors face idiosyncratic risk that is larger than their total risk.[8] This points towards a potential population for whom a financial tool designed to indemnify covariate risk may be inappropriate because it would increase the variance of losses.

Let an insurance product be available that makes indemnity payments based on the values of an index generated in each division at every period ($Index_{d,t}$). The difference between experienced losses and the index ($L_{i,d,t} - Index_{d,t}$) is basis error. The variance of basis error, often called basis risk and shown in Equation (1), is the risk that an insured individual faces.

$$(1) \qquad Var_t[L_{i,d,t} - Index_{d,t}] = Var_t[L_{i,d,t}] + Var_t[Index_{d,t}] - 2 * Cov[L_{i,d,t}, Index_{d,t}]$$

So long as the variance introduced by the index is less than twice the covariance between the index and losses, an individual can reduce risk by purchasing the index insurance.

[4] A model in the style of Miranda (1991) and a discussion of the links between that model and the one presented here can be found in Appendix A.

[5] A division could be defined any number of ways. Defining index divisions spatially makes sense for products that hope to mitigate risk associated with weather-sensitive activities, such as agriculture, where losses are often spatially correlated.

[6] For consistency with IBLI and comparability with conventional insurance, where indemnity payments are based on individual losses, we assume an index that predicts loss rates. This discussion can easily be recast in terms of deviations from any value, such as precipitation below a benchmark or number of cooling days.

[7] Assume that variance is a suitable measurement of risk for the time being. We will extend this analysis to allow for asymmetric preferences by examining skewness and semi-variance after decomposing basis risk.

[8] Perhaps a more intuitive specification of the covariate risk faced by an individual is limited to that risk which positively co-varies with their division average and has a maximum value of the individual's total risk. In this case, idiosyncratic losses are limited to those individual losses that are greater than division average losses, and covariate risk is calculated using only that portion of division losses that are not greater than individual losses. The drawback to this alternative specification is that it does not capture variance associated with overestimation of losses such as those falling into the false positive region, as will soon be discussed.

An index that tracks average division level losses exactly maximizes total coverage and minimizes basis risk but is likely to be unachievable or at least generally not cost effective.[9] Differences between the division average and the index are called design errors. The variance in design error, design risk ($Var_t[\bar{L}_{d,t} - Index_{d,t}]$), is the remaining covariate risk that could theoretically be captured by a (better) division level index.

The risk that an insured individual faces can be described by the sum of design risk, idiosyncratic risk, and the covariance between design error and idiosyncratic error:

(2) $$Var_t[L_{i,d,t} - Index_{d,t}] = Var_t[\bar{L}_{d,t} - Index_{d,t}] + Var_t[L_{i,d,t} - \bar{L}_{d,t}] + 2Cov[L_{i,d,t} - \bar{L}_{d,t}, \bar{L}_{d,t} - Index_{d,t}]$$

In addition to the magnitude of basis risk, the sign and circumstances of basis error are also likely to be important to consumers. Figure 1 illustrates that point by displaying all of the possible loss-index combinations. The vertical and horizontal axis represent the range of time-specific individual losses ($L_{i,d,t}$) and index values ($Index_{d,t}$), respectively, where both index and losses refer to a loss rate ($L_{i,d,t}, Index_{d,t} \in [0,1]$). The 45° line represents the set of outcomes where the index and losses are identical and basis error is zero. Above the 45° line, losses are greater than those predicted by the index, while below the 45° line, the index predicts higher losses than experienced. The absolute difference between the index and experience increases as one moves away from the 45° line, so that basis errors are largest in the top left and lower right corners.

Contracts may map index values onto indemnity payments in a nonlinear fashion. For example, index insurance generally does not cover all losses. The strike (*S* in Figure 1) is the value that the index must exceed in order for there to be an indemnity payment, equivalent to a deductible in conventional indemnity insurance. If the index falls below the contractually specified strike level, no payments are made even if positive losses are predicted. If the index is greater than the strike, payments are made according to the conditions of the contract. The lower left shaded section represents those outcomes where both losses and the index are below the strike. Although events falling into this region may provide a signal as to the relationship between the index and events, it does not impact the contract's precision in providing accurate indemnity payments because the contract explicitly does not cover risk in this region.

Events during which high losses are suffered but the index remains below the strike level are termed false negatives. False negatives are likely to have a poor impact on the reputation of the product because households pay a premium and experience losses that exceed the strike, but none of those losses are indemnified. Analogously, a high index that initiates a payment while the individual losses are less than the strike falls in the false positive region. Although false positive indemnity payments are a

[9] If basis risk is the variance of the remainder generated by subtracting index values from individual losses ($\varepsilon_{i,d,t} = L_{i,d,t} - Index_{d,t}$), then basis risk is minimized by using an index perfectly identical to the division average losses. Or $arg \min_{Index_{d,t}} Var_{i,t}[L_{i,d,t} - Index_{d,t}] = arg \min_{Index_{d,t}} Var_{i,t}[\varepsilon_{i,d,t}] = Index_{d,t}{}^*$, such that $Index_{d,t}{}^* = \bar{L}_{d,t}$.

windfall for individuals, the payments are not necessarily risk reducing and may perversely transfer money from low to high income states through premiums to fund that windfall.

The basis risk faced by an individual would describe the distribution of $(L_{i,d,t}, Index_{d,t})$ realizations scattered in Figure 1. The Index Based Livestock Insurance (IBLI) product and IBLI household survey from northern Kenya provide a rare opportunity to examine this basis risk distribution using policy and household data, which are described in the next section.

3. **Background on Kenyan Pastoralists, IBLI and the Data**

Pastoralist households in northern Kenya depend on livestock for most of their income (mean =70% and median=100% in our data) as well as for a wide variety of financial and social services. Frequent droughts in the region play a large role in livestock mortality and household herd size. For example, in both 2009 and 2011, severe droughts hit the horn of Africa, causing mortality rates greater than 50% in some locations (OCHA 2011; USAID 2011; Zwaagstra et el. 2010). Indeed, drought is the single largest cause (47%) of livestock mortality in our survey data. For pastoralist households, herd loss represents a direct loss of wealth and productive assets on which both current and future incomes depend.

The Index-Based Livestock Insurance (IBLI) pilot started in the Marsabit district of northern Kenya in January 2010. IBLI is an index insurance product based on a remotely collected indicator: the normalized difference vegetation index (NDVI). NDVI is an indicator of the level of photosynthetic activity in observed vegetation and, being a good proxy of the available rangeland forage for animals, should be highly correlated with livestock mortality.[10] The NDVI data originally employed was sourced from NASA's Advanced Very High Resolution Radiometer (AVHRR) NDVI3g sensor.[11] The IBLI contract was designed by regressing historic livestock mortality rates on transformations of lagged NDVI data to estimate a seasonal livestock mortality rate response to NDVI observations (Chantarat et al. 2013).[12] The regression approach is appealing because minimizing the residual sum of squared errors is equivalent to minimizing the variance of the difference between the index and individual losses, or basis risk.

Division-specific indices are calculated for each of Marsabit's five administrative divisions. The five divisions were grouped into two contract divisions, upper and lower, each with its own response function. Figure 2 displays the five index (administrative) divisions. The legend shows how the index divisions are allocated into contract divisions. The IBLI strike and deductible are set at 15% so that indemnity rates are equal to max(index-0.15,0).

[10] Purchased feed is essentially non-existent in these populations.

[11] The current, updated IBLI product uses NASA's eMODIS dataset. See Vrieling et al. (2014) for analysis on selecting the best relevant NDVI source for IBLI and the intercalibration efforts to stitch the relatively new eMODIS dataset (available only from 2000 to the present) with the AVHRR data (available from 1981 to 2012).

[12] The IBLI contract was revised for scale-up and implemented in Marsabit as well as Isiolo and Wajir districts, in August 2013 (see Woodard, Shee & Mude 2014 for more information). As this paper focuses on the years 2009 – 2012 the analysis is based on the IBLI design as specified in Chantarat et al. (2013).

The Marsabit region experiences a bimodal rainfall pattern, which naturally produces two insurance seasons per year. The long rain/long dry contract season (LRLD) begins on March 1st and ends September 30th. The short rainy/short dry contract season (SRSD) begins October 1st and runs through end-February. Twelve month contracts are sold twice a year, during the two months preceding each insurance contract season (January-February and August-September), so that each twelve month contract covers two indemnity periods. See Chantarat et al. (2013) for more detailed information on the IBLI product.

Our analysis uses data from a longitudinal household survey collected annually for four years between 2009 and 2012. The first survey round took place three months before IBLI launched and subsequent rounds took place during the same October-November period each year thereafter. In order to maintain statistical power in an environment in which geographic clustering effects are likely large, the survey collected data from sublocations in four of the five IBLI pilot divisions: Central/Gadamoji, Laisamis, Loiyangalani, and Maikona. North Horr was omitted. The survey questionnaires were collected within 16 sublocations selected to provide a wide variety of market access, agro-ecological zones, ethnicity, and herd size. Within sublocations, households were randomly selected within herd size strata. The survey collects data on a wide variety of demographic, economic, and health characteristics but emphasizes livestock herd dynamics.

Because we are interested in comparing estimated sample variances, this analysis uses only those households that participated in all four rounds. Of the original 924 households in the survey, 832 were available for all four rounds. About 30 households (~3%) were replaced each round. Attrition, for the most part, was due either to the household moving to a distant location or unavailability of an appropriate household respondent. The first factor may be the result of shock or an indicator of household mobility, both of which are of interest in this study. Repeated visits were attempted to reduce the incidence of the second factor. Attrition analysis (reported in detail in Appendix B) finds that households that leave the survey tend to have fewer members, rely on livestock for a smaller portion of their income, and consume more per person.

We place the additional restriction that households have at least one animal in every round so that their livestock mortality rate is defined, reducing the sample further to 736 households. Those dropped due to periods with no livestock are similar to the exiting households, in regards to differences from the sample used, with the addition of having more education and smaller herd sizes.

Thus, our final sample is the product of attrition, due to households leaving the survey, and truncation as we study only those households that have livestock in every period. The result is a sample in which shocks are likely underrepresented. If attrition or reported zero livestock are due to livestock shocks, the sample selection process will bias shock related estimates (e.g., average livestock losses) downwards. Unfortunately, there is little that can be done to address this bias except to control for those variables known to be related to attrition or zero livestock, which we do.

Consistent with the IBLI contracts, the months are grouped into two insurance seasons: LRLD and SRSD.[13] Household-season livestock mortality rates are estimated using current herd size and recall data of births, deaths, slaughter, sale, and purchases. There are also a number of relevant annual premium rates for IBLI policies: the subsidized rate at which policies were sold during the survey period, the within-sample actuarially fair premium rate, and the loaded and unsubsidized commercial rate. Although IBLI coverage was only available for the last five of the eight insurance seasons captured in these data, all eight seasons are used in our analysis in order to better estimate the basis risk distributions that households face. A detailed description of the livestock mortality rate estimation process, the various premium rates, and index values are found in Appendix C.

In the following sections we examine the impact of IBLI coverage on risk and estimate a number of idiosyncratic and design risk metrics in order to provide a clear picture of IBLI's performance. We focus on full insurance rather than optimal coverage because we are specifically interested in learning about the distribution of basis risk and factors that determine where a household falls in the distribution. Fully insured households provide us with the opportunity to examine the factors that are associated with both positive and negative outcomes, be that from poor index design or high idiosyncratic risk.[14] A convenient byproduct of fully insured herds is that net outcomes are in units of livestock mortality/survival rate.[15]

4. **Results**

Welfare Effects

We begin this section by examining the effects of purchasing full IBLI coverage on the distribution of survival rates and utility. This provides a vantage point by which to better understand the magnitude and heterogeneity in coverage, and thus basis risk, provided by the IBLI product. It is also, to the knowledge of the authors, the first look at coverage provided by an index product in a low-income country that draws on household level data. In the interest of brevity, and because the methods used are standard, the full analysis and discussion has been placed in Appendix D.

Comparing the survival rates of those without IBLI coverage to the net survival rate of those with IBLI coverage—calculated as the survival rate less commercial premium rate plus indemnities—reveal that IBLI coverage changes the distribution of outcomes dramatically (Figure 3). Most apparent is a significant mass of households that experience a greater than one net outcome with insurance, when

---

[13] Each SRSD season runs from October through February, crossing into a new calendar year. They are dated by the year of the month that the season begins (October) rather than the year that the season ends (February). LRLD begins at the beginning of March and ends at the end of September so this is not an issue.

[14] Because coverage cannot be negative, an analysis of optimal coverage would only include those households for whom IBLI improves outcomes.

[15] At full insurance all calculations can be performed as a ratio of the full herd. The net survival rate on an insured herd is estimated by subtracting seasonal loss and premium rates from one, and adding indemnity rates when payments trigger.

(by construction) there are no households with greater than one livestock biological survival rate.[16] Households with greater than one net survival rate received indemnity payments exceeding the sum of their losses plus the premium.

Notice as well that a small number of observations have moved to the left of zero livestock survival in the insured case. A less than zero net outcome is due to households paying premiums and suffering extremely high losses but receiving very little or no indemnity payment.[17] Thus, at the population level, there is a small but real chance that an insured household may face a net outcome of less than zero. Because of there is no similar possibility for uninsured households, the insured distribution of outcomes fails to stochastically dominate the uninsured distribution at any degree of analysis.

Fully insuring at the commercial premium rate statistically significantly worsens average expected outcomes and variance in outcomes by 1.3% (t-stat=19.91) and 4.7% (t-stat=3.31), respectively (Table 1). But, as is suggested by the histograms of outcomes (Figure 4), IBLI indemnity payments significantly improve the skewness of the survival rate distribution by shifting it towards the right by 45.1% (t-stat=10.13, Table 1). These outcomes are consistent with any loaded and unsubsidized insurance product: a reduction in expected outcome and improved skewness.

The welfare impact of an increase to variance is not clear in this case, as it is the result of both under- and over-indemnification of losses. Downside risk, or the risk a household faces losses beyond the strike, is unencumbered by such ambiguity. During such high-loss seasons, households with insurance see a statistically significant improvement to mean (t-statistic=1.98) and semi-variance (t-statistic=2.70) of outcomes when the premiums are at the actuarially fair rate (Table 2). At the commercial rates, the average expected outcome is worse with insurance than without (t-statistic=-3.68) and the change in semi-variance is not statistically different from zero (t-statistic=-1.50).

Although the semi-variance analysis allows us to focus on the set of high-loss events that are likely to be most important to the household, we continue to be faced with households that we cannot order because they see benefits by one metric at the expense of losses by another. Utility analysis can be used to order the outcomes for all the households but requires greater assumptions about households' preferences. Following other work on the utility gains from insurance (e.g., Woodard et al. 2012), we assume constant relative risk aversion (CRRA). Herd size for the fully insured scenario is simulated using observed household-specific herd growth and mortality rates and by assuming all premiums (indemnity payments) are made via herd offtake (intake). At all three levels of risk aversion examined, most

[16] That is not to say that there are not observations of net seasonal growth to herd size. Herd size increased between seasons in about 32% of the observations. Here, we are examining only the insured risk, which is livestock mortality, not changes to herd size.

[17] In 16 of the 5,888 observations, households experienced less than zero net livestock survival rate due to premium rates being added to an already high livestock mortality rates. The minimum net outcome is -0.0212.

households in both the lower and upper IBLI contract regions see gains to utility at premium rates above the commercial rates (Table 3).[18]

This section examined the impact that IBLI coverage has on household welfare outcomes. We find that the benefits of IBLI coverage outweigh its cost for most of the population and that the majority of households would continue to benefit from IBLI purchases even at higher premium rates. But, the net benefits are far from universal in magnitude or even sign. The next section examines basis risk at the household level to determine which factors contribute to the net benefits of IBLI.

Decomposing Basis Risk

Although many, and by some measures most, households benefit from IBLI, there are clear signs that policy holders continue to shoulder significant basis risk. This section examines household-level basis risk to determine which contract and household level characteristics are associated with greater basis risk. In order to focus on index design shortfalls we make two changes to our procedure. First, outcomes and net outcome are now measured in terms of livestock mortality and net mortality rates rather than survival rates, as in the previous section. Survival rates can be recovered by subtracting the mortality rate outcome from one. Second, we do not include a premium in this analysis so that our estimates are an examination of the relationship between the index and household data rather than the policy's premium parameters.

Table 4 summarizes the downside risk without insurance and the downside basis risk associated with index shortfalls during high covariate loss events. Downside risk is estimated as the semi-variance of livestock mortality rate beyond the strike. Downside basis risk is estimated as the semi-variance of the difference between livestock mortality rates and the indemnity rate, conditional on high livestock mortality rates (>0.15) and a shortfall in indemnity rates.[19] This focuses our analysis on those periods when households suffer severe losses and on IBLI's performance in reducing risk caused by such losses. The overall average reduction to squared deviations from the strike during high loss events due to IBLI coverage is about 30.7%.

Design Risk

Design risk arises due to differences between the index and the covariate losses. The level of design risk is necessarily shared among all policy holders in the same index division (administrative districts in this case). Figure 2 shows a map of the Marsabit region and the five index divisions; a different index value is calculated for each.

[18] Utility is of the form $U(x_{idt}) = \frac{(x_{idt})^{(1-R_i)}}{(1-R_i)}$, where x_{idt} is herd size held by household *i* in IBLI contract division *d* at period *t,* and with Arrow-Pratt absolute risk aversion R_i. We simulate utility using three level of risk aversion, R=[0,1,2].

[19] For greater detail on the computations used to estimate downside risk see Appendix D.

Figure 4 plots the 32 index-covariate loss observations on the domain described by Figure 1. Fitted lines above and below the strike are also included, along with confidence intervals. There is clearly large variation across the sample in how well the index performs. Below the strike, the fitted line lies above the 45 degree line indicating that index is likely to underestimate division level mortality rates when those rates are below 0.15. Above the strike, the index generally overestimates the covariate losses. In total, there are eight (25%) observed false positives and four (12.5%) false negatives. The high rate of discrete error observed on an index designed explicitly to minimize basis risk and tested out-of-sample using a data set other than the design data (Chantarat et al. 2013) serves as a strong caution against overconfidence in the quality of index insurance products.

To examine the accuracy of the index we focus on those events when covariate losses were greater than the strike (above the horizontal red line in Figure 4). Table 5 provides summary statistics of the covariate and design risk associated with those events. The covariate risk is estimated using the target semi-variance in order to examine the risk associated with severe events and represents only 20% of the average downside risk estimated in Table 4.[20] Design risk is then calculated as the semi-variance of the shortfall of the index during those events. Notice that the average conditional design risk represents less than 10% of the average conditional basis risk presented in Table 4, foreshadowing the large role that idiosyncratic risk plays. The precision is an estimate of the portion of conditional covariate risk that the index covered. On average, the index reduces covariate risk by about 62.8% but there is significant heterogeneity in covariate risk and index precision between divisions.

Regressing covariate losses on the index shows that there are systematic differences between the index and covariate losses (Table 6). The index consistently under predicts both covariate losses and conditional covariate losses; the estimated index coefficient is significantly less than one in both the restricted and unrestricted case. The R^2 statistics provide an indication of the amount of covariate risk that the index is able to account for. Once again, the index performs much better when the sample is restricted to high loss events.

The fact that the index under-predicts high covariate losses is evident by the much greater number of points above the strike but to the left of the 45 degree line than to the right of it in Figure 4. This structural relationship between the index and covariate losses means that a shift and rotation of the index according to the parameter estimates in Table 6 could reduce design risk and thereby increase the accuracy of the insurance product during these severe covariate events within this sample period.

A second potential approach to reducing design risk is to adjust the strike. Calculating design error conditional on covariate losses greater than the strike where the strike falls in the interval $[0,0.25]$, we examine how well the index predicts covariate losses above the strike at various strike levels. We find that varying the strike rate has no significant impact on the accuracy of the index; there is a great deal of variation in design error at all strike levels (Figure 5).

[20] Notice that IBLI coverage is reducing exposure to risk from extreme events at an average rate (30.7%) that is greater than the average share of risk that is associated with large covariate shocks (20%). That is because the IBLI index predicts catastrophic losses in a number of periods during which covariate losses are below 0.15 but there is a subsample with high losses.

Design errors are a significant component of basis risk. These design errors arise due to covariate losses that could be indemnified by the IBLI policy but are not captured by the index as presently designed even though it was explicitly designed to minimize basis risk. Our estimates of the relationship between the index and covariate losses point towards a systematic error that could be addressed by shifting and rotating the index to increase predicted livestock mortality rate during poor seasons. The strike level is a second parameter that could be readily and easily changed if there were gains in precision to be had. But there is no evidence to support one strike level over another. Because the expected absolute value of design error does not change significantly as strike levels change, they might be left open as a contract parameter chosen according to consumer or provider preferences.

Idiosyncratic Risk

A second and far larger portion of downside basis risk arises due to idiosyncratic losses, or mortality not reflected in the division average, or covariate, losses. Although much idiosyncratic loss is likely associated with random events, it may also have a systematic relationship with household or geographic characteristics. If such patterns are known to prospective purchasers, a form of adverse selection subtly returns even though index insurance is pitched in part as an approach to obviate adverse selection problems in conventional insurance.[21] We now examine factors associated with idiosyncratic risk.

The size of the covariate region may affect the level of covariate (and thus remaining idiosyncratic) risk. In theory, index products capture a greater portion of risk as the size of the index region shrinks. The entire IBLI study region covers about 66,700 km^2 (about the size of West Virginia) and is composed of four divisions. Each division consists of sublocations (administrative subunits within divisions), 16 of which are captured by the survey.

Figure 6 shows the ratio of covariate risk to average total risk at various geographic scales of aggregation.[22] This ratio captures the risk faced by households that could be covered by an index product at each covariate scale in this setting. The average ratio of covariate to total risk more than doubles as the covariate area shrinks from a large aggregate region composed of a single IBLI division, to separate divisions defined by sublocation. There is also a great deal of variance between sublocations. Covariate risk within sublocations is less than 15% of total risk in five survey sublocations, while it is greater than 40% in four. In those locations with very low covariate risk, even a local and extremely accurate (i.e., zero design risk) index product could not cover much of the risk that households face. On the other hand, households in many survey sublocations face a great deal of covariate risk, making them prime candidates for index insurance.[23]

---

[21] Note that this sort of adverse selection does not affect equilibrium pricing of the insurance since it does not affect insurer indemnification rates. It merely induces selection effects among prospective clients.

[22] The numerator, covariate risk, is the variance of covariate losses within each covariate region ($CR_d = Var_t[\frac{1}{N_d}\sum_i L_{idt}]$). The denominator is the within region average household variance in losses or average risk ($\overline{Risk}_d = \frac{1}{N_d}\sum_i Var_t[L_{idt}]$).

[23] It is possible that the differences in average covariate risk share are related to variation in the shapes and sizes of the sublocations. But regressing the sublocation average ratio of covariate risk to risk on sublocation area and the ratio of area to perimeter yields no statistical evidence of such a relationship. Results of that analysis are available upon request.

At the division level, there is clearly the potential for geographic patterns to the benefits of IBLI. Risk averse households in sublocations most similar to their division averages are likely to benefit more from insurance covering covariate losses than would households in sublocations with little covariate risk. In addition, three of the index divisions were, until recently, aggregated into a single contract division (the Lower IBLI contract included sublocations Dakabaricha through South Horr from left to right on the x-axis of Figure 6) with a single premium rate, but maintain separate indices. If the division level index correctly predicts covariate losses within each of the three divisions, the Central/Gadamoji division (sublocations Dakabaricha, Dirib Gombo, Sagante, and Karare) will have much higher expected indemnity payments then the remaining households in the contract region even though they pay the same premium rate. In that case, policy holders in the other two divisions are inadvertently subsidizing the premium rate in the Central/Gadamoji division and spatially defined opportunistic behavior (i.e., spatial adverse selection) should emerge; and it does (Jensen, Mude and Barrett 2014).

There is also variation among households and even within households over time. In this final analysis of basis risk patterns, we explore which factors are associated with deviations of households from the average losses experienced within their index division. A number of easily observed characteristics could reasonably impact livestock loss rates. For example, Lybbert et al. (2004), studying a very similar system in neighboring southern Ethiopia, find a strong positive association between herd size and livestock mortality rate, which would translate into a similar relationship with respect to idiosyncratic losses. Access to labor, herd size and composition, cash liquidity, informal insurance network participation and level of risk aversion all might impact how well a household's herd fares compared to the household's division's average losses. A description of the household characteristics considered here and their summary statistics are found in Appendix E. Idiosyncratic losses and the semi-variance of idiosyncratic losses are regressed on household characteristics in order to determine which are associated with idiosyncratic risk. The semi-variance is used rather than variance in order to isolate variance associated with household losses that are greater than covariate losses.

Spatial correlation of idiosyncratic risk could arise due to local environmental shocks or spatially correlated household characteristics. Although we cannot fully disentangle the two here, we can examine household characteristics for explanatory value with and without sublocation fixed effects, in order to reveal when factors are important due to between-sublocation variation and within-sublocation variation. Sublocation fixed effects alone are able to account for a fairly large portion of variation in downside risk (idiosyncratic semi-variance) between households (R^2=0.125, column 4, Table 7) but very little of the variation in idiosyncratic losses (R^2=0.026, column 1, Table 7). Indeed, household characteristics do no better in explaining idiosyncratic losses or downside risk than do sub-location fixed effects as revealed by comparing columns 1 with 2, and 4 with 5 in Table 7. Including both controls for sublocation fixed effects and household characteristics provides the best fit, the R^2 is nearly the sum of those from the considering location and household characteristics separately indicating that the two processes are fairly distinct (columns 1 and 2 vs. 3, and columns 4 and 5 vs. 6).

The ratio of income generated from livestock is the only livestock-related characteristic that is consistently (negatively) associated with idiosyncratic risk, even when we control for sublocation fixed effects. There does seem to be a weak relationship between herd size and exposure to idiosyncratic risk, the average marginal effect of herd size is negative and statistically significant in the analysis presented in Table 7 columns 3 and 6, and the third order polynomial coefficients estimates are jointly statistically significant Table 7 (analysis not included). Households with relatively more dependents also have greater idiosyncratic risk.

What is perhaps the most striking finding of this analysis is how little idiosyncratic risk is associated with household characteristics or can be captured by sublocation fixed effects. Idiosyncratic losses cannot be very well explained by sublocation average losses nor by a host of household characteristics that could reasonably be associated with livestock mortality rates. Idiosyncratic losses seem to be almost entirely random while variance in losses is much more predictable, but still more than 75% of the variation in semi-variance is unexplained by readily observable household characteristics and sub-location fixed effects, as might be practical for targeting purposes.

As a robustness check, we estimate a fixed effects model to determine if unobserved time-invariant household characteristics drive our findings. Only column (2) from Table 7 can be estimated in this way because the within-household variation in sublocation is nearly zero and semi-variance of idiosyncratic losses has no within household variance. In addition, risk aversion, age, and gender variables are dropped due to lack of within household variation. The fixed effects model reported in Appendix E also captures very little of the rate of idiosyncratic losses and there is little indication that those losses are anything but random.

In addition, we test to make sure that our approach and findings in this section are compatible with the benefits found in the earlier welfare analysis. To do so, we re-estimate columns 4-6 in Table 7, replacing the dependent variable with the reduction to risk due to insurance coverage as a proxy for net benefits.[24] As expected, variation in design risk between divisions results in a high degree of correlation between benefits and geographic division (R^2=0.30) while the randomness in idiosyncratic risk allows household characteristics alone to explain very little of the between-household variation in benefits from IBLI (R^2=0.17). See Appendix G for results.

In summary, households that depend on livestock for only a small amount of their income but have relatively large herds and have many dependents will likely suffer from high idiosyncratic losses even after accounting for community fixed effects. The sublocation effects seem to be mostly in addition to household characteristics indicating that they capture factors associated with local environmental conditions. While there is some geographic targeting capacity when the index regions are made sufficiently small in size, none of these observable variables explain much idiosyncratic loss, which is both large in magnitude and mainly random.

5. **Discussion**

[24] The reduction risk is estimated by $\frac{1}{T-1}\{\sum_{t=1}^{T}(M_{idt})^2 - (\max[M_{idt} + Premium_{idt} - Indemnity_{idt}, 0])^2\}$, which places greater value on indemnities during seasons with higher losses and assigns no additional benefits to indemnity payments above the losses that a household experiences.

Index insurance provides a promising means for overcoming many of the barriers that have impeded insurance delivery in poor rural regions of the world. But index insurance has its own weaknesses, chief among which is basis risk. As a result, index insurance may only prove appropriate in certain risk environments and at certain scales. Knowing both the idiosyncratic and design components of basis risk is important in determining the value proposition of index insurance. Regrettably, in practice neither the consumer nor the provider has perfect information. Providers can only learn about the relative magnitude of covariate risk and the accuracy of their index by collecting longitudinal consumer-level information to determine covariate risk, a rare practice. In a similar fashion, consumers can only begin to estimate the design risk once they have observed a number of periods of product coverage.

The result is that although basis risk is widely recognized as the Achilles heel of index insurance, it has to date gone unmeasured and unstudied in index insurance products developed for smallholder farmers and herders in the low-income world. This study provides the first detailed study of basis risk related to index insurance products in developing countries. It examines an insurance contract that is best-in-class in at least two important ways. First, there is a great deal of evidence that large covariate droughts are the largest cause of livestock mortality in the population for whom IBLI is available (e.g., Barrett et al. 2006; Lybbert et al. 2004; McPeak, Little, & Doss 2012; Santos & Barrett 2006). Second, IBLI polices are based on an index that was generated using a long panel of household data and regression methods expressly to minimize basis risk (Chantarat et al. 2013). Other index products fielded in the developing world typically lack similar foundations. These features should make this product something close to a best case scenario for assessing basis risk in index insurance products for farmers and herders in the developing world.

The results are only mildly encouraging and offer a cautionary tale about the prospective benefits of index insurance. Tests for stochastic dominance underscore that index insurance with a positive probability of large false negatives cannot stochastically dominate remaining uninsured. Mean-variance and utility analysis show that IBLI coverage likely improves the outcomes faced by most – but far from all – households in Marsabit, but only modestly. Most importantly, fully insuring with IBLI still leaves households bearing a significant amount of uninsured risk. Some of this basis risk is due to correctable design risk as the index proved an imperfect predictor of covariate livestock mortality rates, underscoring the need for careful *ex post* evaluation and adjustment of index products even when designed *ex ante* to minimize basis risk.

A second, much larger, portion of basis risk is due to idiosyncratic risk. Although the study population is plagued by severe droughts during which nearly all households experience higher than normal livestock mortality, households also experience a tremendous degree of nearly random idiosyncratic variation in every season, even in high covariate loss seasons. These findings echo earlier research showing a dramatic increase in the variation of herders' expectations of their own herd's dynamics when those herders expect poor rainfall conditions rather than good or normal conditions (Santos & Barrett 2006; Barrett & Santos 2014). In addition, the degree of covariate risk is closely tied to how covariate losses are defined spatially and temporally. The ratio of covariate risk to total risk varies by a factor of 6 across the 16 sublocations included in this study.

This research illustrates the complexity of providing index insurance, even in an environment that in some respects seems ideal. It emphasizes the spatial sensitivity of covariate risk to the covariate region and the resulting prospect for spatial adverse selection in demand patterns. It reveals that basis risk, especially idiosyncratic risk, is substantial, pointing towards the continued importance of informal risk sharing agreements and other complementary risk management mechanisms even when index insurance is available. An optimally designed index insurance product yields risk-reducing welfare gains for many prospective purchasers but offers far-from-full coverage. Caution seems warranted in the promotion of index insurance as a risk management instrument for low-income populations underserved by conventional insurance markets.

References

Barrett, C. B., Belle mare, M. F., & Osterloh, S. M. (2006). Household-level livestock marketing behavior among northern Kenyan and southern Ethiopian pastoralists. In McPeak, J., & Little, P. (Eds.), *Pastoral Livestock Marketing in Eastern Africa: Research and Policy Challenges*, (15-38). London: Intermediate Technology Development Group Publishing.

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J., Rasambainarivo, J., & Wangila, J. (2006). Welfare dynamics in rural Kenya and Madagascar. *Journal of Development Studies*, *42*(2), 248-277.

Barrett, C.B., & Santos, P. (2014). The Impact of Changing Rainfall Variability on Resource-Dependent Wealth Dynamics, *Ecological Economics,* 105(1), 48-54.

Breustedt, G., Bokusheva, R., & Heidelbach, O. (2008). Evaluating the potential of index insurance schemes to reduce crop yield risk in an arid region. *Journal of Agricultural Economics*, *59*(2), 312-328.

Chantarat, S., Mude, A. G., Barrett, C. B. and Carter, M. R. (2013). Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya. *Journal of Risk and Insurance*, 80, 205–237.

Clarke, D. J. (2011). *A theory of rational demand for index insurance*. Department of Economics D. Phil. dissertation, University of Oxford.

Dercon, S., Hill, R. V., Clarke, D., Outes-Leon, I., & Seyoum Taffesse, A. (2014). Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia. *Journal of Development Economics*, 106, 132-143.

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.

Giné, X., Townsend, R., & Vickery, J. (2008). Patterns of rainfall insurance participation in rural India. *World Bank Economic Review*, 22(3), 539-566.

Global Index Insurance Facility. (2013). Global Index Insurance Facility Factsheet. Updated April 2013.

Hazell, P. B., & Hess, U. (2010). Drought insurance for agricultural development and food security in dryland areas. *Food Security*, 2(4), 395-405.

Hill, R. V., Robles, M., & Ceballos, F. (2013). *Demand for Weather Hedges in India: An Empirical Exploration of Theoretical Predictions*. Food Policy Research Institute (IFPRI) Discussion Paper 1280.

Jensen, N., Mude, A., & Barrett, C.B. (2014). How Basis Risk and Spatiotemporal Adverse Selection Influence Demand for Index Insurance: Evidence from Northern Kenya. Working paper, Cornell University.

Kenya National Bureau of Statistics. (2013). Consumer Price Indices. Retrieved May 28, 2014, from http://www.knbs.or.ke/

Leblois, A., Quirion, P., & Sultan, B. (2014). Price vs. weather shock hedging for cash crops: Ex ante evaluation for cotton producers in Cameroon. *Ecological Economics*, 101, 67-80

Lybbert, T. J., Barrett, C. B., Desta, S., & Layne Coppock, D. (2004). Stochastic wealth dynamics and risk management among a poor population. E*conomic Journal, 114(*498), 750-777.

Mahul, O. (1999). Optimum area yield crop insurance. *American Journal of Agricultural Economics*, 81(1), 75-82.

McPeak, J. G., Little, P. D., & Doss, C. R. (2012). *Risk and social change in an African rural economy: livelihoods in pastoralist communities*. New York: Rutledge.

Miranda, M. J. (1991). Area-yield crop insurance reconsidered. *American Journal of Agricultural Economics*, 73(2), 233-242.

Miranda, M. J., & Farrin, K. (2012). Index insurance for developing countries. *Applied Economic Perspectives and Policy*, 34(3), 391-427.

Mobarak, A., & Rosenzweig, M. (2012). Selling formal insurance to the informally insured. Yale University Economic Growth Center Discussion Paper 1007.

Norton, M. T., Turvey, C., & Osgood, D. (2012). Quantifying spatial basis risk for weather index insurance. *Journal of Risk Finance, 14*(1), 20-34.

OCHA. 2011. *Eastern Africa: Drought - Humanitarian Snapshot*. http://www.fews.net/docs/Publications/Horn_of_africa_drought_2011_06.pdf.

Santos, P. & Barrett, C. (2006). Heterogeneous Wealth Dynamics: On the Roles of Risk and Ability, Working paper, Cornell University.

Smith, V. H., Chouinard, H. H., & Baquet, A. E. (1994). Almost ideal area yield crop insurance contracts. *Agricultural and Resource Economics Review*, 23(1).

Smith, V., & Watts, M. (2009). Index based agricultural insurance in developing countries: Feasibility, scalability and sustainability. Report to the Gates Foundation.

Turvey, C. G. (2001). Weather derivatives for specific event risks in agriculture. *Review of Agricultural Economics*, *23* (2), 333-351.

USAID/OFDA. 2011. *Horn of Africa-Drought,* Fact Sheet #1, Fiscal Year 2011.

Vedenov, D. V., & Barnett, B. J. (2004). Efficiency of weather derivatives as primary crop insurance instruments. *Journal of Agricultural and Resource Economics*, 387-403.

Vrieling A, Meroni M, Shee A, Mude A, Woodard J, & de Bie Rambold, F. (2014). Historical Extension of Aggregated NDVI from Operational Products for Livestock Insurance in Kenya. *International Journal of Applied Earth Observation and Geoinformation*, 28 (238-251).

Williams, J. R., Carriker, G. L., Barnaby, G. A., & Harper, J. K. (1993). Crop insurance and disaster assistance designs for wheat and grain sorghum. *American Journal of Agricultural Economics*, 75(2), 435-447.

Woodard, J. D., & Garcia, P. (2008). Basis risk and weather hedging effectiveness. *Agricultural Finance Review*, *68* (1), 99-117.

Woodard, J. D., Pavlista, A. D., Schnitkey, G. D., Burgener, P. A., & Ward, K. A. (2012). Government Insurance Program Design, Incentive Effects, and Technology Adoption: The Case of Skip-Row Crop Insurance. *American Journal of Agricultural Economics*, *94*(4), 823-837.

Woodard J, Shee A, & Mude A. (2014). A Spatial Econometric Approach to Scalable Index Insurance against Drought Related Livestock Mortality in Kenya. Under review*, Journal of Risk and Insurance.*

Zwaagstra, L., Sharif, Z., Wambile, A., Leeuw, J. D., Said, M. Y., Johnson, N., Njuki, J., Said, M., Ericksen, P., & Herrero, M. (2010). *An assessment of the response to the 2008 2009 drought in Kenya*. International Livestock Research Institute (ILRI).

Figures

Figure 1. The domain of basis error

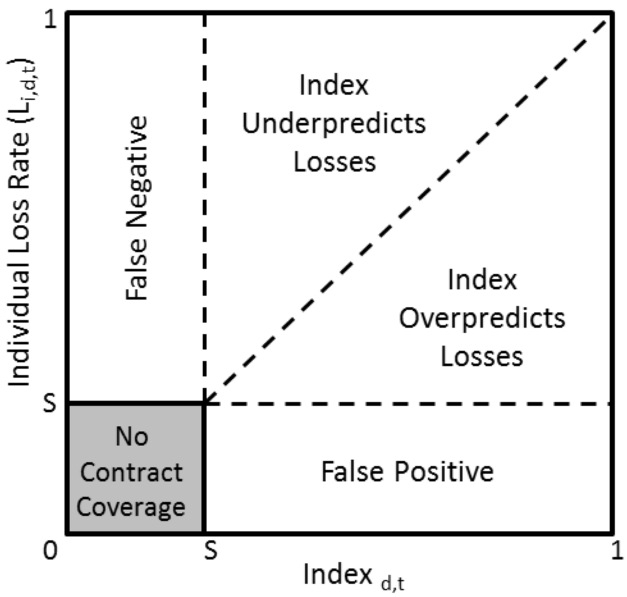


Figure 2. IBLI Geographical Coverage and Index Divisions

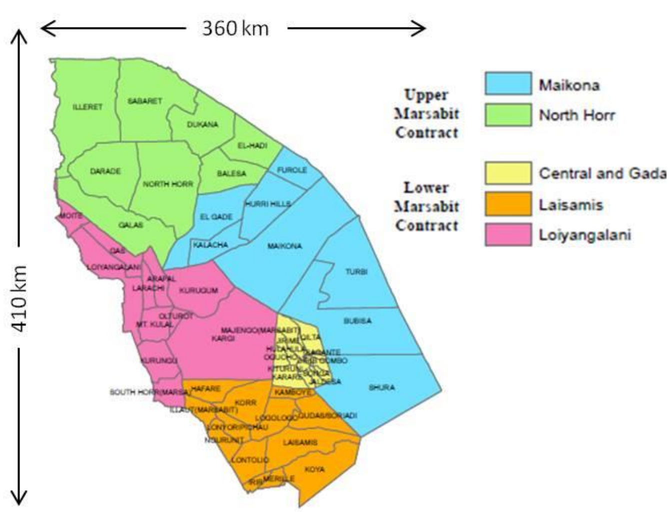
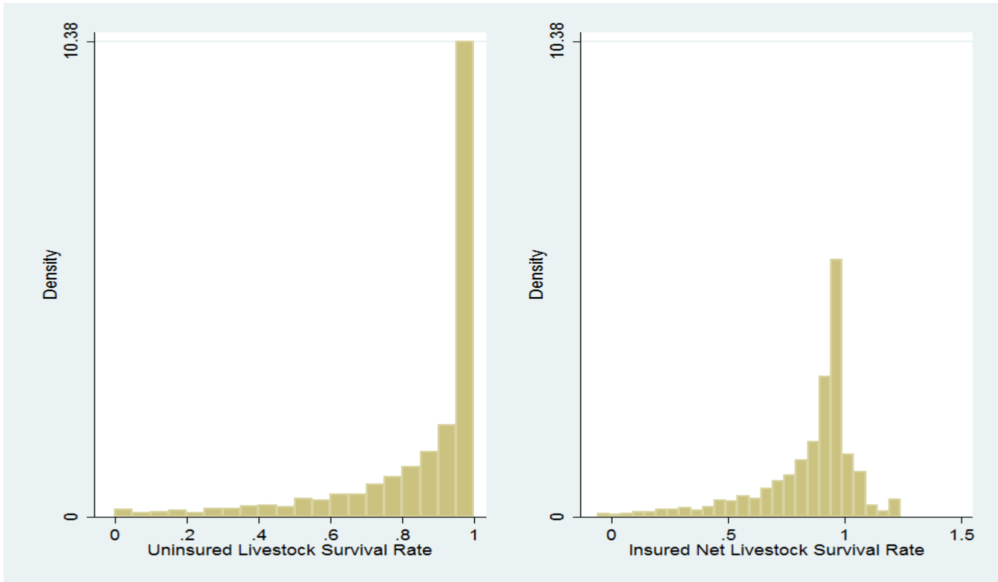
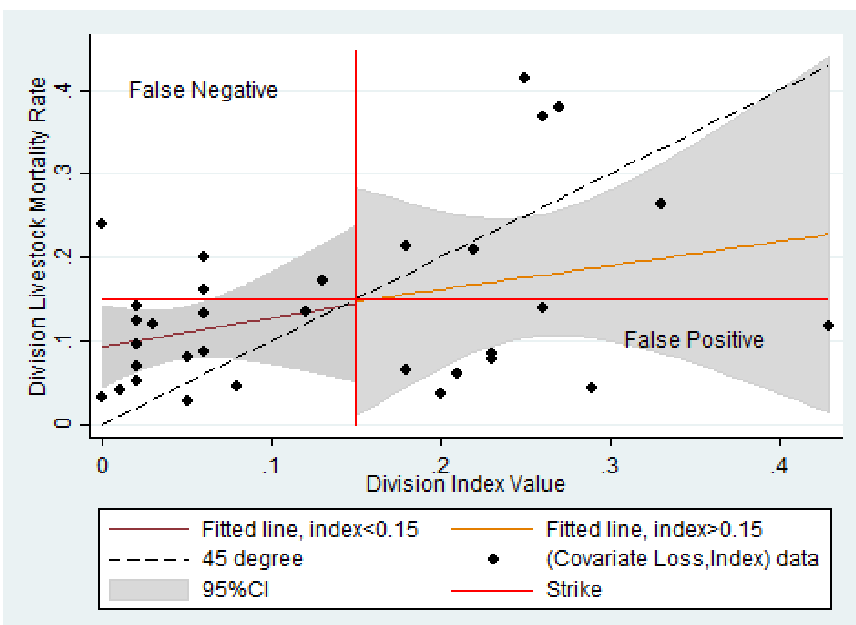


Figure 3. Histograms of livestock survival rate and net livestock survival rate with full insurance



Note: Net livestock survival rate with full insurance equals the household's seasonal survival rate less the commercial premium (loaded and unsubsidized) plus indemnity payments.

Figure 4. Design error above and below the strike (0.15)



Notes: Covariate loss-index observations are seasonal division average mortality paired with the index value for that division-season. Fitted lines and confidence intervals are generated by regressing livestock mortality rates on the index.

Figure 5. Index accuracy at various covariate loss strike levels

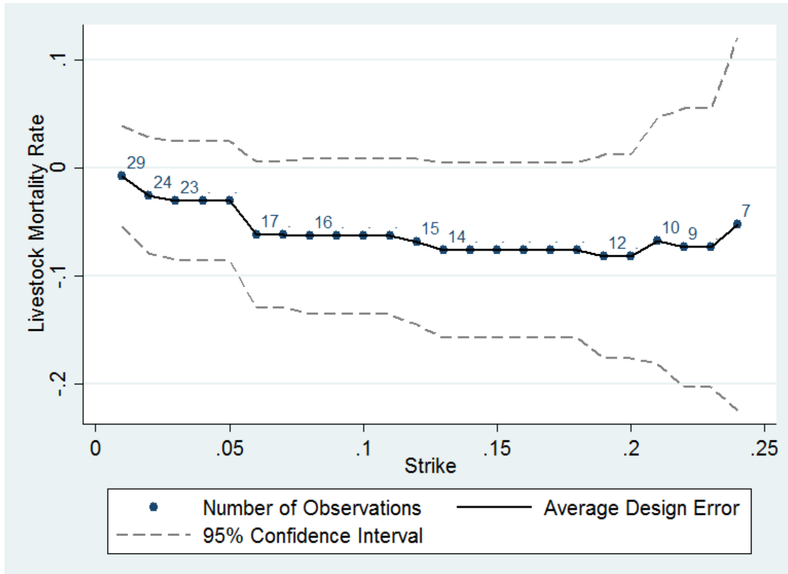
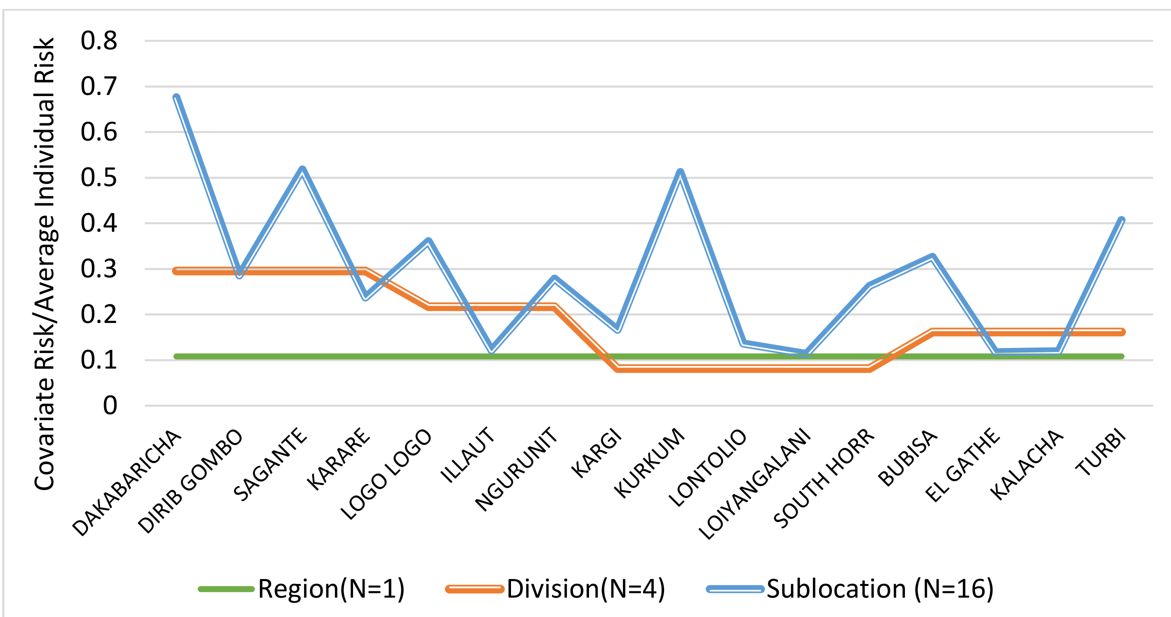


Figure 6. Ratio of average covariate risk to average individual risk at different geographic scales



	Average Risk (R)	Average Covariate Risk (CR)		
		Entire Sample	Division	Sub-Location
Variance	0.0431	0.0047	0.0076	0.0110
Ratio (CR/R)	*1*	*0.108*	*0.156*	*0.222*

Tables

Table 1. The impact of IBLI on average within-household mean, variance, and skewness of survival rate

Statistic	Uninsured	Insured	Difference	Standard Error	t-statistic
Mean	0.864	0.854	0.011	0.001	19.91***
Variance	0.043	0.045	-0.002	0.001	-3.31***
Skewness	-1.185	-0.651	-0.535	0.053	-10.31***

Notes. Analysis uses the commercial premium rate.

Table 2. Impact of IBLI on downside risk in mortality rate during severe events (mortality rate > 0.15)

Premium	Statistic	Uninsured	Insured	Difference	t-statistic
Commercial[1]	Expected Losses >.15	0.080	0.083	-0.002	-3.68***
	Semi-Variance	0.035	0.036	-0.001	-1.50
Actuarially Fair[2]	Expected Losses >.15	0.080	0.079	0.001	1.98**
	Semi-Variance	0.035	0.034	0.001	2.70***
Subsidized[3]	Expected Losses >.15	0.080	0.072	0.008	13.37***
	Semi-Variance	0.035	0.031	0.005	10.17***

Notes. [1]The commercial annual premium rates are Central/Gadamoji=10.6%, Laisamis=11.3%, Loiyangalani=9.2%, and Maikona=10.7%. [2]The within-sample actuarially fair annual premium rates are Central/Gadamoji=9.25%, Laisamis=7.5%, Loiyangalani=7.0%, and Maikona=12.25%. [3]At that time that this data was collected, the subsidized annual rates were Central/Gadamoji=3.325%, Laisamis=3.325%, Loiyangalani=3.325%, and Maikona= 5.5%.

Table 3. Proportion of households that are better off with IBLI than without

	Subsidized Rate[1]		**Actuarially Fair**[2]		**Loaded & Unsubsidized**[3]	
Coefficient of Risk Aversion	**Lower**	**Upper**	**Lower**	**Upper**	**Lower**	**Upper**
R=0 *(risk neutral)*	0.999	0.952	0.931	0.167	0.539	0.611
R=1	0.998	0.939	0.933	0.254	0.560	0.605
R=2 *(risk averse)*	0.997	0.921	0.935	0.303	0.544	0.595

Notes. [1]At that time that this data was collected, the annual subsidized rates were Central/Gadamoji=3.325%, Laisamis=3.325%, Loiyangalani=3.325%, and Maikona= 5.5%. [2]The within-sample actuarially fair annual premium rates are Central/Gadamoji=9.25%, Laisamis=7.5%, Loiyangalani=7.0%, and Maikona=12.25%. [3]The commercial annual premium rates are Central/Gadamoji=10.6%, Laisamis=11.3%, Loiyangalani=9.2%, and Maikona=10.7%.

Table 4. Average risk without insurance, basis risk with insurance and coverage rate during periods with high covariate livestock mortality rates

	Central/Gadamoji	Laisamis	Loiyangalani	Maikona	Overall
Conditional Losses (rate)[1]	0.505	0.406	0.358	0.381	0.394
Conditional Index (rate) [1]	0.203	0.163	0.125	0.156	0.150
Conditional Risk (X100) [2]	6.082	4.002	2.985	2.966	3.636
Conditional Basis Risk (X100)[2]	4.410	3.047	2.753	2.035	2.896
Coverage[3]	0.360	0.345	0.233	0.377	0.307
Observations	157	109	247	223	736

Notes. [1] Conditional losses and index values are the average index and livestock mortality rates from the pool of household observations with greater than 0.15 livestock mortality rate. [2] Conditional (downside) risk is estimated using semi-variance which is calculated by $100 * \frac{1}{T-1}\sum_{t=1}^{T}(O_{idt} - \dot{O}_{dt})^2 I(Z_{idt}))$ where O_{dt} is the outcome experienced in division d, in time period t, $\dot{O}_{dt}$ is the target, and $I(Z_{idt})$ is an indicator function that is equal to one if $O_{idt} > \dot{O}_t$ and equal to zero otherwise. Conditional risk is calculated using $O_{idt} = L_{i,d,t}$ and $\dot{O}_{dt} = 0.15$. Conditional basis risk is calculated using $O_{idt} = \max(L_{i,d,t} - 0.15, 0)$, $\dot{O}_t = \max(index_{dt} - 0.15, 0)$ to measure risk associated with shortfall in the index during high losses events. [3] Coverage is the average reduction in risk.

Table 5. Mean covariate risk and design risk during seasons when covariate losses were above the strike

	Central/Gadamoji	Laisamis	Loiyangalani	Maikona	Overall
Conditional Covariate Losses (rate)[1]	0.392	0.228	0.207	0.255	0.262
Conditional Index (rate)[1]	0.255	0.170	0.120	0.165	0.176
Conditional Covariate Risk (X100)[2]	1.699	0.781	0.094	0.323	0.725
Conditional Design Risk (X100)[2]	0.572	0.167	0.053	0.119	0.228
Precision[3]	0.663	0.786	0.431	0.631	0.628
Seasons w/ mean loss>0.15	2	4	2	2	10

Notes. [1] Division averages for seasons during which the covariate losses are greater than 0.15. [2]Conditonal (downside) risk is estimated using semi-variance which is calculated by $100 * \frac{1}{T-1}\sum_{t=1}^{T}(O_{dt} - \dot{O}_{dt})^2 I(O_{dt}))$ where O_{dt} is the outcome experienced in division d, in time period t, $\dot{O}_{dt}$ is the target, and $I(O_{dt})$ is an indicator function that is equal to one if $O_{dt} > \dot{O}_t$ and equal to zero otherwise. Conditional covariate risk is calculated using $O_{dt} =$ covaritate losses (CL_{dt}) and $\dot{O}_t = 0.15$. Conditional design risk is calculated using $O_{dt} = \max(CL_{dt} - 0.15,0)$ and $\dot{O}_t = \max(index - 0.15,0)$. [3] Precision is the average ratio of conditional covariate risk captured by the index.

Table 6. The relationship between covariate losses and the index

	Covariate losses	Conditional covariate losses[1]
Index	0.318**	0.527*
	(0.151)	(0.237)
Constant	0.095***	0.168***
	(0.027)	(0.048)
F-stat: H_0: $\alpha_d = 0\ and\ \delta_d = 1$	10.26***	8.01**
F-stat: H_0: $\delta_d = 1$	20.50***	3.99*
Observations	32	10
R-squared	0.129	0.383

Notes.[1] Conditional covariate losses are covariate losses during season's when covariate losses were greater than the strike (0.15). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7. Factors that contribute to idiosyncratic livestock mortality rate and downside risk

	Idiosyncratic Loss Rate (IL)			Semi-Variance(IL)*		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Age (/100)		-0.1102	-0.1401		-0.0496	-0.0756
		(0.1283)	(0.1173)		(0.0735)	(0.0682)
Age^2 ($age^2/100^2$)		0.0873	0.1341		0.0308	0.0638
		(0.1092)	(0.1010)		(0.0650)	(0.0614)
Gender (=1 if male)		-0.0181**	-0.0076		-0.0062	-0.0038
		(0.0085)	(0.0085)		(0.0043)	(0.0040)
Household size (count/100)		-0.0917	-0.1143		-0.0320	-0.0412
		(0.1837)	(0.1768)		(0.1064)	(0.1003)
Dependency ratio		0.0699***	0.0528**		0.0469***	0.0367***
		(0.0253)	(0.0242)		(0.0140)	(0.0138)
Asset index#		-0.0198	-0.0874		0.0473	0.0541
		(0.0786)	(0.1013)		(0.0574)	(0.0710)
Asset index squared#		-0.0283	0.2301		-0.2896	-0.3186
		(0.3662)	(0.4092)		(0.2211)	(0.2548)
HSNP participant		0.0051	-0.0059		-0.0038	-0.0087
		(0.0089)	(0.0097)		(0.0046)	(0.0054)
Ratio herd camels&		-0.0036	0.0090		-0.0032	0.0166
		(0.0226)	(0.0242)		(0.0177)	(0.0201)
Ratio herd cattle&		-0.0053	0.0032		0.0021	0.0093
		(0.0186)	(0.0180)		(0.0207)	(0.0258)
Herd size (TLU/100)&		0.0191	-0.1142*		-0.0289	-0.1389**
		(0.0534)	(0.0629)		(0.0565)	(0.0603)
Herd $size^2$ ($TLU^2/100^2$)&		-0.0330	0.0763		0.0929	0.2643**
		(0.0665)	(0.0746)		(0.1087)	(0.1097)
Herd $size^3$ ($TLU^3/100^3$)&		0.0030	-0.0209		-0.0455	-0.1201**
		(0.0171)	(0.0194)		(0.0511)	(0.0503)
Ratio income from livestock#		-0.0235**	-0.0226*		-0.0372***	-0.0275*
		(0.0110)	(0.0126)		(0.0102)	(0.0148)
Log (1+Savings)#		0.0011	0.0016		-0.0015**	-0.0011
		(0.0014)	(0.0016)		(0.0007)	(0.0011)
Social groups (count)#		-0.0046	-0.0032		-0.0056	-0.0043
		(0.0060)	(0.0065)		(0.0035)	(0.0039)
Moderately risk averse		-0.0109	-0.0070		-0.0022	0.0001
		(0.0086)	(0.0070)		(0.0050)	(0.0041)
Extremely risk averse		-0.0049	-0.0078		0.0015	-0.0005
		(0.0095)	(0.0079)		(0.0051)	(0.0044)
Sublocation Fixed Effects (16)	Yes	No	Yes	Yes	No	Yes
F-stat testing: All sublocation Fixed Effects=0	6.74***		3.82***	4.39***		3.81***
Observations	5,888	5,120	5,120	736	736	736
R-squared	0.0260	0.0129	0.0326	0.1254	0.1196	0.2103

Notes. Regression also included an intercept term. * Semi-variance of idiosyncratic losses is calculated as the within household sum of squares of $\widetilde{IL_{\iota,t}}$ where $\widetilde{IL_{\iota,t}} = \max(IL_{i,t}, 0)$ #Variable is lagged by one period in the idiosyncratic losses estimation in order to reduce potential endogeneity. &Variable uses seasonal average monthly herd size. Household clustered-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Basis Risk and the Welfare Gains from Index Insurance: Evidence from Northern Kenya

—Online Appendix—

September 2014

Nathaniel D. Jensen
Cornell University

Christopher B. Barrett
Cornell University

Andrew G. Mude
International Livestock Research Institute

Contents:

Corresponding author: Nathaniel Jensen, Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14850; ndj6@cornell.edu

Appendix A: Regression Analysis of Basis Risk

Regression analysis provides one method for examining basis error patterns across index levels. Examining only design risk for a moment, Equation (A1) expresses division level average losses as a function of the index.

(A1)
$$\bar{L}_{d,t} = \alpha_d^0 + \alpha_d^1 Index_{d,t} + \mu_{d,t}$$

$$E[\mu_{d,t}] = 0$$

Here α_d^0 is the intercept, α_d^1 is the expected change to covariate losses for a unit change in the index, and $\mu_{d,t}$ is mean zero error.[25] Together α_d^0, α_d^1 tell us the expected design error for any index value. The variance of the error informs on the uncertainty around the expectation. Notice that the variance of the error term ($Var_t[\bar{L}_{d,t} - \alpha_d^1 Index_{d,t}]$) is similar but not equivalent to our earlier definition of design risk ($Var_t[\bar{L}_{d,t} - Index_{d,t}]$) unless$\alpha_d^1 = 1$.

An index without design error will have coefficients $\alpha_d^0 = 0, \alpha_d^1 = 1$ and $Var[\mu_{d,t}] = 0$. Deviations from this zero design risk ideal can manifest in a number of ways. For example, if $\alpha_d^0 < 0$ and $\alpha_d^1 \approx 1 - \frac{\alpha_d}{S}$ where S is the strike, then the index generally over predicts losses below the strike and under predicts losses above the strike. Perhaps most usefully, if Equation A1 is estimated and $\widehat{\alpha_d^0} \neq 0$ or $\widehat{\alpha_d^1} \neq 1$ basis risk could be reduced by transforming the existing index ($Index_{d,t}$) to a new index $Index_{d,t}{}^*$, where $Index_{d,t}{}^* = \alpha_d^0 + \alpha_d^1 Index_{d,t}$.

Individual losses can be similarly expressed as a function of division level covariate losses and idiosyncratic losses. Equation A2 expresses that relationship as the sum of an intercept ($\rho_{i,d}^0$), a parameter ($\rho_{i,d}^1$) times covariate losses, and an idiosyncratic component ($\vartheta_{i,d,t}$).

(A2)
$$L_{i,d,t} = \rho_{i,d}^0 + \rho_{i,d}^1 \bar{L}_{d,t} + \vartheta_{i,d,t}$$

$$E[\vartheta_{i,d,t}] = 0$$

The risk described by $Var_t[\bar{L}_{d,t}(\rho_{i,d}^1 - 1) + \vartheta_{i,d,t}]$ is individual level losses that could be insured by a loss indemnity insurance product but is not, by design, covered by an index product.[26] It represents the minimum possible risk that an individual could be exposed to after purchasing an index insurance product based on $\bar{L}_{d,t}$. Index insurance is ideal for individuals with $\rho_{i,d}^0$ near zero, $\rho_{i,d}^1$ near one, and a low level of idiosyncratic losses.

[25] We assume stationarity of losses and the relationship between individuals, divisions and the index. Nonstationarity could be modeled by allowing coefficients α_d^0 and α_d^1 to vary with time but adds little to this discussion.

[26] Idiosyncratic risk can be calculated by $Var_t[L_{i,d,t} - \bar{L}_{d,t}] = Var_t[\bar{L}_{d,t}(\rho_{i,d}^1 - 1) + \vartheta_{i,d,t}]$. If $\rho_{i,d}^1 = 1, V[\vartheta_{i,d,t}]$ is equal to idiosyncratic risk.

The relationships illustrated in Equations (A1) and (A2) can be combined to express the relationship between individual losses and the index (Equation A3). Notice that the basis error parameters can be divided into division level components (design error) and individual level components (idiosyncratic error).

(A3)
$$L_{i,d,t} = \beta^0_{i,d} + \beta^1_{i,d} Index_{d,t} + \varepsilon_{i,d,t}$$

Where:

$$\begin{aligned}
\beta^0_{i,d} &= \alpha^0_d + \{\rho^0_{i,d} + \alpha^0_d(\rho^1_{i,d} - 1)\} \\
\beta^1_{i,d} &= \rho^1_{i,d}\alpha^1_d \\
\varepsilon_{i,d,t} &= \mu_{d,t} + \{\vartheta_{i,d,t} + \mu_{d,t}(\rho^1_{i,d} - 1)\}
\end{aligned}$$

Equation A3 can be estimated using historic data to examine the components of basis risk at different index levels, in different divisions, and for individual households. Holding the parameters constant within divisions estimates the expected performance of the index product within each division. If the index perfectly predicts a household's experience, then $\beta^0_{i,d} = 0$, $\beta^1_{i,d} = 1$ and $\varepsilon_{i,d,t} = 0\ \forall t$.

Appendix B: Attrition and Selection Analysis

The level of sample attrition is less than 4% per year; 37 households between first and second rounds, 30 between second and third rounds, and 25 between third and fourth rounds. There are significant differences between the survey households that exit and those that remain in the survey (Table B1). Households that leave the survey are larger, consume less per person, and generate a greater portion of income from livestock related activities. About 12% of the remaining households are dropped because they have periods with zero reported livestock so that their livestock mortality rate is undefined. The dropped households are similar to the exit households but also have significantly lower education, greater herd size and income than the control households.

Table B1. Balancing Table (2009 data): Attrition, dropped, and full data households (2009-2012)

Variable	Maintained [1] (N=736)	Exit, or Dropped	Difference	T-statistic	
Exit households (N=92[2])					
Max education[3]	4.31	4.74	-0.43	-0.88	
Household members (count)	5.76	4.89	0.87	3.38	***
Dependency ratio[4]	0.62	0.59	0.02	0.96	
Consumption per capita (KShs)	1,377	1,736	-360	-2.76	***
TLU owned[5]	19.71	16.14	3.57	1.28	
Income (KShs)	5,259	3,504	1,755	1.01	
Ratio of income form livestock	0.56	0.30	0.26	3.54	***
Risk category[6]	2.49	2.65	-0.16	-0.84	
Savings (KShs)	6,893	13,795	-6,901	-0.98	
Households with zero livestock holdings in at least one period (N=96)					
Max education[3]	4.31	5.28	-0.98	-2.03	**
Household members (Count)	5.76	4.79	0.97	3.87	***
Dependency Ratio[4]	0.62	0.62	-0.01	-0.23	
Consumption per capita (KShs)	1,377	1,989	-612	-4.68	***
TLU owned[5]	19.71	4.30	15.41	5.78	***
Income (KShs)	5,259	5,258	1.29	0.00	
Ratio of income form livestock	0.56	0.08	0.48	9.33	***
Risk category[6]	2.49	2.73	-0.23	-1.28	
Savings (KShs)	6,893	5,217	1,676	0.26	

Notes. [1] Households that are in all four survey rounds and never have zero livestock for an entire IBLI season (March-September or October-February). [2] N=92 is composed of 88 households that left the survey and were replaced, and 4 that miss one survey round but did not leave the survey. [3] Maximum level of education achieved by any household member where 1-8 are standards, 9-12 are forms 1-4, 15 is a diploma, 16 a degree and 17 a postgraduate degree. [4] Ratio of household members aged less that 15 or older than 54 years to the total household size. [5] Tropical Livestock Units (TLU) are calculated as follows: Camel=1TLU, Cattle=0.7 TLU, Sheep & goats=0.1 TLU. [6] Risk categories are discrete values ranging from 0 (most risk averse) to 5 (most risk taking) elicited using a real lottery with variation in expected winnings and variance of outcomes similar to that described by Binswanger (1980). *** (p<0.01), ** (p<0.05) and * (p<0.1).

Appendix C: Livestock Mortality Rate, IBLI Premium Rates and Index Values

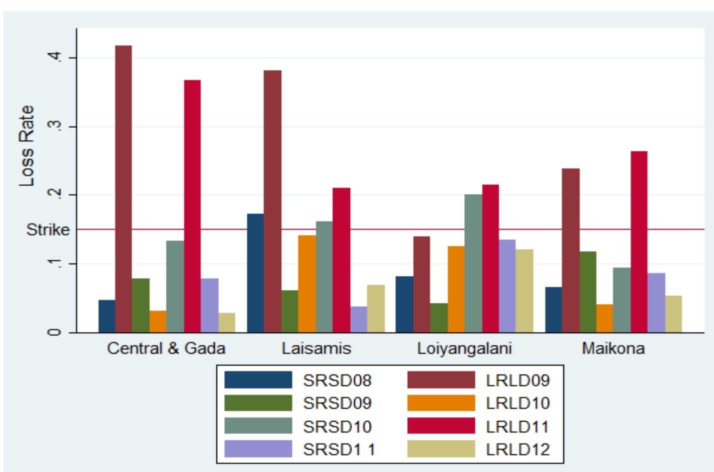
The ideal estimate of seasonal livestock mortality rate is the ratio of animals entering a season that die during the season. But the data do not allow for tracking specific animals through the season so we construct an alternative estimate of seasonal livestock mortality rate. The numerator of this alternative estimate is the sum of monthly losses ($M_{i,d,m}$) for individual *i* in division *d* during month *m* for all months that fall into season *s*. The denominator is composed of the sum of the herd size at the beginning of the season ($H_{i,d,start}$) and all monthly additions to the herd over the following season ($\sum_{m\in s} A_{i,d,m}$).[27] Thus, seasonal livestock mortality rates ($L_{i,d,s}$) are estimated by dividing the season's cumulative livestock mortality by the total herd owned by household that season (Equation C1).[28]

(C1)
$$L_{i,d,s} = \frac{\sum_{m\in s} M_{i,d,m}}{H_{i,d,start} + \sum_{m\in s} A_{i,d,m}}$$

Where:
$$s = \begin{Bmatrix} LRLD & if\ m = [March, \dots, Sept] \\ SRSD & if\ m = [Oct, \dots, Feb] \end{Bmatrix}$$

Average mortality rates vary widely between the four study divisions and across seasons (Figure C1). More important for this analysis, there is clear evidence of large covariate losses within divisions, as is revealed by seasons with high average mortality rates. IBLI can only be an effective risk mitigation tool if individual level catastrophic losses are correlated. An ideal IBLI product would indemnify those (average) losses that are above the strike (0.15) in Figure C1.

Figure C1. Division level average livestock mortality rate across seasons



Notes. The index strike value is 0.15. SRSD is short rain/short dry insurance season. LRLD is the long rain/long dry insurance season.

[27] $H_{i,d,start}$ is calculated using reported herd sizes at the time of the survey and iterating backwards, adjusting for monthly birth, death, purchase, sale, and slaughter. Herd size is constrained by $0 \leq H_{i,d,m}\ \forall\ i, d, m$ to address errors in recall that occasionally lead to erroneous negative livestock herd size estimates.

[28] We rely on estimates of livestock mortality rate because the data does not track individual livestock through each season. The qualitative results presented in this paper are robust to using an alternative method for calculating livestock mortality rate, which is described and used in Chantarat et al (2013).

There are three important premium rates to consider for IBLI (Table C1). The subsidized rates that were made available to pastoralists during the periods covered by this analysis offer insight into the conditions that the survey households actually faced in these periods. The within-sample actuarially fair premium rates provide the best estimates, however, if the intent is to focus on the intertemporal smoothing effect of insurance. Finally, the unsubsidized loaded annual premium rates calculated by the insurance providers in 2014 provide information on outcomes associated with commercially sustainable, unsubsidized premium rates. These final rates reflect a reevaluation of the expected indemnity payments in 2014 in response to severe conditions between 2009 and 2013. Notice that the premium rates are no longer common in the upper and lower contract divisions as of 2014.

Table C1. Annual Premium Rates in Percent of Insured Value

	Subsidized Rates[1]	Within-Sample Actuarially Fair Rates	Unsubsidized & Loaded Commercial Rates
Central/Gadamoji	3.325%	9.25%	10.60%
Laisamis	3.325%	7.50%	11.30%
Loiyangalani	3.325%	7.00%	9.20%
Maikona	5.50%	12.25%	10.70%

Notes. [1] The subsidized rates were available to pastoralists form January 2010-January 2012.

This research includes analysis of basis risk before IBLI was available for sale. In those non-sale periods, there are no publically available index values. In the seasons before LRLD 2010, index values were collected from internal documents: "IBLI Pricing 2010" (SRSD 2008 LRLD 2009 and SRSD 2009) and "IBLI Marsabit Pricing June 2012" (LRLD 2010). The remainder (SRSD 2010 though LRLD 2012) were collected from the publically available IBLI index archive available at http://livestockinsurance.wordpress.com/ibli-kenya/mortality-index-update/index-archive//.The indemnity payments represent a percentage of the value of the insured asset and are calculated according to the IBLI contracts (max (index-0.15,0)).

Table C2. IBLI Index Values and Imputed Indemnity Payments

	Central & Gadamoji		Laisamis		Loiyangalani		Maikona	
Seasons	Index	Indemnity	Index	Indemnity	Index	Indemnity	Index	Indemnity
SRSD 2008[1]	0.08	0.00	0.13	0.00	0.05	0.00	0.18	0.03
LRLD 2009[1]	0.25	0.10	0.27	0.13	0.26	0.11	0.00	0.00
SRSD 2009[1]	0.23	0.08	0.21	0.06	0.29	0.14	0.42	0.27
LRLD 2010	0.00	0.00	0.02	0.00	0.02	0.00	0.01	0.00
SRSD 2010[1]	0.06	0.00	0.06	0.00	0.06	0.00	0.02	0.00
LRLD 2011	0.26	0.11	0.22	0.07	0.18	0.03	0.33	0.18
SRSD 2011	0.23	0.08	0.20	0.05	0.12	0.00	0.06	0.00
LRLD 2012[1]	0.05	0.00	0.02	0.00	0.03	0.00	0.02	0.00

Notes. [1]IBLI was not sold during these seasons.

Appendix D: Welfare Effects of IBLI Coverage

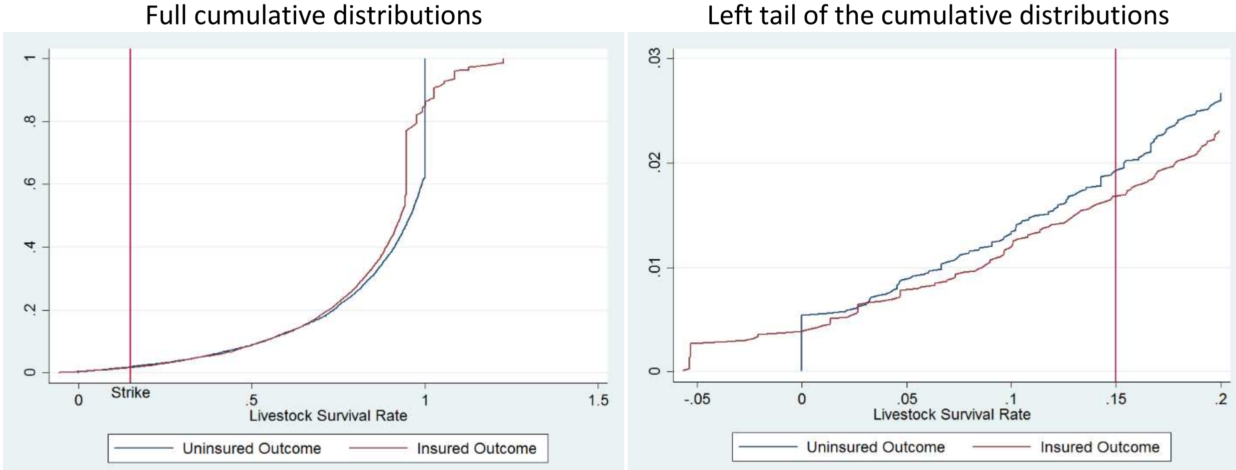
Stochastic Dominance

Testing for stochastic dominance is one approach for ordering risky choices in a manner consistent with expected utility theory. The main advantage of the stochastic dominance approach is that is allows for ordering with few assumptions about the utility function. Unfortunately, with only eight seasonal observations per household, our data do not allow for powerful tests of stochastic dominance at the household level. Rather, we test for stochastic dominance at the population level.

Let *f(x)* describe the distribution of observed livestock survival rates and *g(x)* describe the net outcome of fully insuring (i.e., net of premium and indemnity payments). If the insured survival rate distribution first order stochastically dominates (FSD) the uninsured distribution, $F(x) \equiv \int_{-\infty}^{x} f(x)dx \gg G(x) \equiv \int_{-\infty}^{x} g(x)dx$, then the expected outcome with insurance is better than without insurance. Figure D1 shows that the insured distribution does not FSD the uninsured state. In particular, as shown in the right panel of Figure D1, which focuses on just the left tail of the distribution depicted in the left panel, no insurance dominates insurance when households experience extremely high losses and do not receive indemnity payments greater than the premium. Indeed, the insured distribution necessarily fails to stochastically dominate the uninsured case at any degree of stochastic dominance because of the positive probability of negative net survival rates under insurance due to catastrophic losses with little or no indemnity payment.

Figure D1. Cumulative distribution of livestock survival rate and net outcome:

Full cumulative distributions — Left tail of the cumulative distributions



Mean-Variance Metrics

The mean-variance method for analyzing choices under risk is common in the insurance literature. For example, Miranda (1991) defines the change to yield risk due to insurance as the variance in yield without insurance less the variance of the net yield, which includes premiums and indemnity payments. This approach is intuitive and requires the estimation of very few parameters, allowing for more powerful household level analysis than does testing for stochastic dominance, and is consistent with expected utility as long as mean and variance are sufficient for describing differences in outcomes (Meyer 1987). But insurance may lead to changes beyond those that are captured by mean and variance, so that mean—variance analysis is inconsistent with important classes of preferences. For example, risk averse individuals may distinguish asymmetrically between deviations from the mean due to extremely good outcomes and extremely poor outcomes (Alderfer & Bierman 1970). Agricultural insurance products specifically target those negative outcome events rather than all variation (Turvey 1992). Higher moments (beyond mean and variance) can be calculated to examine changes to distributions that are not symmetrical while semi-variance analysis examines changes to downside risk.

Loaded and unsubsidized insurance in unlikely to be mean preserving or improving, since it is priced above the actuarially fair level. Comparing the expected net outcome of being insured with the uninsured case shows that the loading indeed results in a net decrease in survival rates from about 86.4% to 85.4% for a difference of about 1.1% per season (Table D1), which is very near the estimated loading rate.[29]

Table D1. The impact of IBLI on average within-household mean, variance, and skewness of survival rate

Statistic	Uninsured	Insured	Difference	Standard Error	t-statistic
Mean	0.864	0.854	0.011	0.001	19.91***
Variance	0.043	0.045	-0.002	0.001	-3.31***
Skewness	-1.185	-0.651	-0.535	0.053	-10.31***

Notes. Table D1 is also included in the main body of the paper as Table 1.

But the primary motivation for purchasing insurance is presumably not to increase expected outcomes but to reduce the risk of extremely poor outcomes. In this case, the average variance with insurance is slightly greater (4%) than without. This is not surprising as the domain of potential outcomes has increased for insured households and we expect over-indemnification to also contribute to outcome variance. The histograms of outcomes (Figure 4) suggest that IBLI impacts the downside risk that households face via indemnity payments that shift outcomes to the right. Analysis of skewness supports that hypothesis. Distributions are negatively skewed in both the uninsured and insured cases, but

[29] The loading rate is about 40% of the actuarial premium rate. One can back out the provider's estimated average actuarially fair premium rate by dividing the average seasonal premium (5.37%) by 1.4 which comes to about 3.84%. Thus the average loading is the difference between the two or about 1.52%

insurance significantly reduces the skewness magnitude, by 45.1% (t-stat=10.13, Table D1). The skewness values indicate that the impact of IBLI is not a symmetric contraction of the variance. Rather, IBLI reduces the likelihood of large shocks at a small cost to expected outcomes, as is to be expected from a loaded insurance product.

We now focus our attention on downside risk. By examining only risk associated with shocks producing greater than 15% livestock mortality, we reveal how IBLI performs in the domain that falls within the coverage parameters of the IBLI policy. To do so, we use an approach similar to that described in Turvey (1992). Downside risk is calculated by $\frac{1}{T-1}\sum_{t=1}^{T}(O_{it}-\dot{O}_t)^n I(Z_{it})$ where O_{it} is the outcome experienced by individual *i* in time period t, $T=1,2,\ldots,8$, $\dot{O}_t$ is the target, n is the weight given to deviations from the target, and $I(Z_{it})$ is an indicator function that is equal to one if a condition is met and equal to zero otherwise.

In this case, the outcome under examination is livestock mortality rate and the indicator function is used to identify severe events defined by those seasons in which the household experienced at least 15% livestock mortality.[30] The target is used to reference the magnitude of the shock, which we set to the strike in order to capture the risk beyond the strike, associated with those extreme losses. The outcome set of measures are the average sum of the distance between outcome and strike with distance weighted by n. Because the distance measure is not in relation to the mean, as it is with variance, the addition of a constant premium rate affects this measure of downside risk. This is important as risk coverage is often discussed quite separately from premium levels. To explore the effects of premium levels on downside risk we include estimates of downside risk for the subsidized, within-sample actuarially fair, and commercial, unsubsidized rates.

Setting $n=1$ provides an estimate of the expected losses beyond the strike. The expectation of the outcome will rest on the level of loading or subsidy applied to the premium and the timing of the indemnity payments. If indemnity payments are *perfectly* made during high loss events, households with insurance could experience an improvement to expected conditional losses even at the commercially loaded premium rate. Conversely, if the product is not making payments during the high loss events we could see an increases in expected net losses even at subsidized rates. The estimates indicate the index is performing somewhere between the two boundary outcomes described above, triggering indemnity payments during seasons with high losses enough of the time to statistically significantly improve expected outcomes at the subsidized and actuarially fair premium rates, but not enough to overcome the additional 40% loading of the commercial rates (Table D2).

[30] The equation used to estimate downside risk includes a degree of freedom correction (T-1) because it is a transformation of variance, which can be consistently estimated by setting $\dot{O}_t$ to the mean of O_{it}, n to 2, and the indicator function to one.

Table D2. Impact of IBLI on downside risk in mortality rate during severe events (mortality rate > 0.15)

Premium	Statistic	Uninsured	Insured	Difference	t-statistic
Commercial[1]	Expected Losses >.15	0.080	0.083	-0.002	-3.68***
	Semi-Variance	0.035	0.036	-0.001	-1.50
Actuarially Fair[2]	Expected Losses >.15	0.080	0.079	0.001	1.98**
	Semi-Variance	0.035	0.034	0.001	2.70***
Subsidized[3]	Expected Losses >.15	0.080	0.072	0.008	13.37***
	Semi-Variance	0.035	0.031	0.005	10.17***

Notes. Table D2 is also included in the main body of the paper as Table 2. [1]The commercial annual premium rates are Central/Gadamoji=10.6%, Laisamis=11.3%, Loiyangalani=9.2%, and Maikona=10.7%. [2]The within-sample actuarially fair annual premium rates are Central/Gadamoji=9.25%, Laisamis=7.5%, Loiyangalani=7.0%, and Maikona=12.25%. [3]At that time that this data was collected, the subsidized annual rates were Central/Gadamoji=3.325%, Laisamis=3.325%, Loiyangalani=3.325%, and Maikona= 5.5%.

Semi-variance around the target is estimated by setting $n = 2$. As with the conditional expected losses, the estimates indicate that the benefits associated with reductions to semi-variance during severe events are very sensitive to the premium levels (Table D2). At the commercially loaded rate, the average household's semi-variance remains about the same with or without IBLI, but at the actuarially fair rate or subsidized rate households are, on average, better off with IBLI coverage. It worth noting that perfect loss-indemnifying insurance above 15% would drive both the expected losses above 15% and the semi-variance above the strike to zero. But, perfect index insurance would not cover all losses above the strike unless all individuals within the covariate region sufferer from identical losses at all times. For example, 49.2% of the non-zero observations used in Table D2 (experiencing livestock mortality rate>0.15) occurred during periods when covariate losses were below 0.15, and thus fall outside the parameters of the IBLI contract. We examine the index design and idiosyncratic contributions towards this remaining basis risk, represented by the semi-variance here, in Section 4 of the main text.

On average, IBLI sold at the commercially loaded premium rate significantly reduces expected net survival rate but also adjusts the distribution to one more favorable to the household as indicated by a significant reduction in skewness. Restricting our analysis to those periods when households experience greater that 15% livestock mortality reveals that the benefits of IBLI coverage on downside risk are highly sensitive to the premium rates and are positive at the actuarially fair rates. Yet, the impact of IBLI is likely to be heterogeneous across loss rates, premium levels, and households, so that while many households do benefit from IBLI, many others do not.

Table D3 reports the share of households for whom IBLI improves survival rates, variance, skewness or semi-variance in order to begin unpacking the distribution of benefits/cost. At the division level, the commercially loaded, unsubsidized premium rates exceed the expected IBLI indemnity payments in all divisions but Maikona where the expected seasonal indemnity payment rate=6.1% and premium rate=5.6%. Using only the mean standard, those 226 households in Maikona are better off with IBLI than without. At the subsidized rate, all households realize improved expected outcomes with IBLI.

Table D3. Proportion of households for whom IBLI improves their position with respect to each statistic

Statistic	Proportion	
	Loaded & Unsubsidized	Subsidized
Mean	0.232	1.000
Variance	0.359	0.359
Skewness	0.817	0.817
Semi-Variance	0.374	0.609

Assuming that the motive for purchasing IBLI is to reduce (downside) risk, higher moments are more important metrics than is the mean. Variance is reduced for 35.9% of the population, skewness is reduced for 81.7%,[31] and downside risk is improved for 37.4% and 60.9% at the unsubsidized and subsidized rates, respectively. At either premium level, many of these households experience net benefits as measured by one metric and net costs by another. The mean variance framework cannot order outcomes for these households.

Utility Analysis

Utility analysis allows us to order the outcomes for all the households but requires specific assumptions about households' preferences. Following other work on the utility gains from insurance (e.g., Woodard et al. 2012), we assume constant relative risk aversion (CRRA) such that utility is of form $U(x_{idt}) = \frac{(x_{idt})^{(1-R_i)}}{(1-R_i)}$, where x_{idt} is herd size held by household *i* in IBLI contract division *d* at period *t,* and with Arrow-Pratt absolute risk aversion R_i.[32] CRRA utility has been shown to provide a good fit to the risk preferences of individuals so long as the scale of risk (payoffs) does not change drastically (Holt & Laury 2002). In this case, the scale of risk is defined by changes to herd size due to losses, which are rarely more than a factor of ten.[33]

To simulate being fully insured by IBLI, we begin by calculating the observed net herd growth (loss) rate between sales seasons (g_{idt}). In the uninsured case, livestock at the end of each period is calculated by multiplying the herd size during the sales season (TLU_{idt}) by the growth rate.[34] Insurance is simulated by adjusting herd size by the net of premiums paid to fully insure and indemnity payments, if received. The sequence of events is as follows: the household enters time period *t* with livestock TLU_{idt}, the herd grows by g_{idt} (the rate calculated above) and indemnity payments In_{idt}, then is reduced by the amount

[31] Note that the variance and skewness are unaffected by the difference in premium rates, which is constant and thus only affects the mean and semi-variance.
[32] When $R_i = 1$, utility takes the form $U(x_{idt}) = \ln(x_{idt})$.
[33] In only 77 of 5,704 observations do households lose more than 90% of their herds in a single season.
[34] The outcome of this calculation is, of course, the original herd size observations.

required to fully insure the updated herd. The seasonal premium rate is Pr_d per TLU.[35] The herd size at the end of period *t* is TLU_{idt+1}, which is the beginning herd size for period *t+1*.

Indemnity payments in time period *t* are made according to livestock insurance purchased at the end of period *t-1,* the index value from the end of period *t* ($index_{dt}$), and the parameters of the IBLI contract (strike=0.15).[36] To simplify the model, we assume that IBLI contracts last for a single season and that the premium rate for a single season contract is exactly one half the premium rate of an annual IBLI contract. Thus, the household makes a fresh insurance purchase in each period and contracts do not overlap. They model can be described as follows:

(D1)
$$U_t\left((TLU_{idt} * g_{idt} + In_{idt}) * (1 - \frac{Pr_d}{1 + Pr_d})\right) = U_t(TLU_{idt+1})$$

$$In_{idt} = TLU_{idt} * \max(index_{dt} - 0.15, 0).$$

This process for simulating the impact of IBLI coverage assumes that premiums are paid through a reduction in herd size and indemnity payments are reinvested directly into livestock. Households in this region have access to livestock markets (Barrett, Bellemare & Osterloh 2004) and can reinvest indemnity payments into livestock if that is optimal for them. The impacts of this assumption are ambiguous *a priori.*

The expected utility with and without insurance is estimated as the average of the utility estimates over all seven periods.[37] We then determine the number of households who have greater expected utility with IBLI coverage than without it at various levels of risk aversion and over the three premium levels. For consistency, we continue to group the index divisions into upper and lower groups.

As expected, fewer households benefit from IBLI as the premium rates increase. At the subsidized rates, the rates at which policies were sold during this period, nearly every household is better off with insurance.

Perhaps the most interesting case is at the within-sample actuarially fair price. These findings are driven purely by the intertemporal reallocation of livestock, drawing herds down to pay premiums during good years and receiving indemnity payments in poor years. The reallocation provided by the IBLI schedule improves the expected utility for about 90% of households in the lower division, but does not perform as well in the upper division. This is because the index greatly over-predicted covariate losses in the upper division during SRSD09 and LRLD11, driving the within-sample actuarially fair premium rate quite high. Here, design risk has resulted in overestimation of the within sample actuarially fair premium rate.

---

[35] The household fully insures by solving $TLU'_{idt} - Pr_d\widetilde{TLU}_{idt} = TLU'_{idt} - Pr_d(TLU'_{idt} - Pr_d\widetilde{TLU}_{idt})$, where $TLU'_{idt} = TLU_{idt} * g_{idt} + In_{idt}$ and $\widetilde{TLU}_{idt}$ is the number of livestock insured. The solution is $\widetilde{TLU}_{idt} = TLU'_{idt}/(1 + Pr_d)$ which costs $Pr_d * TLU'_{idt}/(1 + Pr_d)$ in reduced herd size.

[36] We assume that full coverage is purchased on all the livestock owned in each period.

[37] Because growth rate in period *t* includes information from *t+1* we cannot simulate into the final (eighth) period. In addition, we assume that households enter the first period with no insurance coverage.

The loaded unsubsidized premium rate provides another interesting scenario, those that would benefit if the policies were sold at their commercially sustainable rate. The ratio of households that benefit from IBLI in the upper region is greater with these rates than with the within-sample actuarially fair premiums because the premium continues to fall below the within-sample expected indemnity payment. The loaded premiums are estimated using a much longer dataset than only these eight seasons, which have had an abnormally high number of severe droughts. Notice that the loaded provider rates used here are the same rates as those used in the mean-variance section (Table D3). This utility analysis is able to order all the households and allows for insurance coverage in one period to affect herd size in following periods. With this complete ordering and accounting for dynamic effects due to endogenous herd growth, the proportion of risk averse households that benefit from IBLI increases dramatically from 35% in the mean variance analysis (column 1,Table D3) to more than 54% (Table D4).

Table D4. Proportion of households that are better off with IBLI than without

	Subsidized Rate[1]		**Actuarially Fair**[2]		**Loaded & Unsubsidized**[3]	
Coefficient of Risk Aversion	**Lower**	**Upper**	**Lower**	**Upper**	**Lower**	**Upper**
R=0 *(risk neutral)*	0.999	0.952	0.931	0.167	0.539	0.611
R=1	0.998	0.939	0.933	0.254	0.560	0.605
R=2 *(risk averse)*	0.997	0.921	0.935	0.303	0.544	0.595

Notes. Table D4 is also included in the main body of the paper as Table 3. [1]At that time that this data was collected, the annual subsidized rates were Central/Gadamoji=3.325%, Laisamis=3.325%, Loiyangalani=3.325%, and Maikona= 5.5%. [2]The within-sample actuarially fair annual premium rates are Central/Gadamoji=9.25%, Laisamis=7.5%, Loiyangalani=7.0%, and Maikona=12.25%. [3]The commercial annual premium rates are Central/Gadamoji=10.6%, Laisamis=11.3%, Loiyangalani=9.2%, and Maikona=10.7%.

Note that there is little evidence of a U-shaped relationship between risk aversion and benefits from IBLI, as predicted by Clarke (2011). But in this case, the (non)linearity of the relationship between risk aversion and benefits from insuring is only part of the story. For some households, IBLI increases risk but has an expected payoff that is greater than the premium rate even when prices are actuarially fair while for others the opposite is true.[38] The utility function coordinates exchange rates between the moments.

Willingness to pay for seasonal contracts is estimated as the highest premium value at which at least 50% of households are better off with insurance than without it. We also examine how willingness to pay changes across a range of risk preferences. We find that the upper contract region has greater willingness to pay than the lower region (Table D5). Consistent with our findings in Table D4, willingness to pay is always greater than the commercial rate. In the lower contract division, at least half of the simulated risk averse households would continue to be better off purchasing IBLI at a rate 42% higher than the expected indemnity rate (3.9%) while in the upper contract division risk averse households

[38] Because herd sizes are not constant, households can net profits from always fully insuring even when premiums are equal to the expected index. For example, a household that has a small herd coming into a good season and then a large herd coming into a season that triggers an indemnity payment would net a profit by fully insuring in both seasons even if the premium rate is equal to the two—season average index value.

have a willingness to pay that is below the expected indemnity rate as the result of a series of high design error periods.

Table D5. Average willingness to pay for seasonal coverage as a share of insured livestock value

	Mean willingness to pay	
Coefficient of Risk Aversion	**Lower**	**Upper**
R=0 *(risk neutral)*	5.48%	5.54%
R=1	5.54%	5.56%
R=2 *(risk averse)*	5.60%	5.58%

Appendix E: Household Variables and Summary Statistics

Table E1. Description of Household Characteristics Used to Examine Idiosyncratic Risk

Variable	**Description**
Idiosyncratic Losses	Seasonal difference between household loss rate and division average loss rate.
Semi-Variance	Within household sum of squares of the difference between losses and covariate losses, conditional on individual losses greater than covariate losses.
Age	Age of head of household to capture lifecycle and herding experience effects.
Household Size	Number of individuals in the household as a control for access to labor.
Dependency Ratio	The ratio of persons under 15, over 65, chronically ill, and disabled to total household members.
Asset Index	An index constructed by factor analysis of a large list of household construction materials and assets. The asset index is discussed in more detail in below.
HSNP	A dummy variable indicating that the household is a participant in the Hunger Safety Net Program (HSNP), an unconditional cash transfer program.
% Camels	Ratio of herd that are camels
% Cattle	Ratio of herd that are cattle
Herd Size	Total herd size in tropical livestock units (TLU) where Camel=1TLU, Cattle=0.7 TLU, Sheep or goats=0.1 TLU
Income	Total cash and in-kind income in real (February 2009) Kenya shillings adjusted for changes to the consumer price index found at Kenya National Bureau of Statistics.
Ratio Income Livestock	Share of income generated from livestock and their byproducts.
Savings	Total savings in real (February 2009) Kenya shillings adjusted for changes to the consumer price index found at Kenya National Bureau of Statistics. Log(1+savings) is used in the regressions.
Social Groups	A count of the number of the following groups that the household participates in: self-help group, women's group, youth group, group related to a water point, group related to pasture, group related to livestock business, merry-go-round savings and lending group, and other.
Risk Aversion	Risk aversion is elicited by offering a one-time lottery similar to the process described in Binswanger (1980). Each household choose one lottery, a coin was flipped, and the household received payment accordingly. The households were given the following set of gambles to choose from: A: Heads- 50 KShs , Tails – 50KShs B: Heads- 45 KShs , Tails – 95KShs C: Heads- 40 KShs , Tails – 120KShs D: Heads- 30 KShs , Tails – 150KShs E: Heads- 10 KShs , Tails – 160KShs F: Heads- 0 KShs , Tails – 200KShs In this analysis, household's level of risk aversion is categorized according to their lottery choice by the following: A or B are considered extremely risk averse, C or D are moderately risk averse, E or F are extremely risk averse.

Table E2. Summary Statistics of Idiosyncratic Risk and Household Characteristics

	Mean	Std. Dev.	Min	Max
Idiosyncratic Losses	0.00	0.20	-0.42	0.95
Semi-Variance of Idiosyncratic Losses	0.035	0.036	0.001	0.230
Age of household head	48.07	18.13	18	99
Number of household members	5.66	2.17	1	19
Dependency ratio	0.60	0.20	0	1
HSNP participant (1=yes)	0.25		0	1
Ratio of herd: camels	0.29	0.30	0	1
Ratio of herd: cattle	0.33	0.31	0	1
Herd size (TLU)	15.3	20.1	0	344.1
Income (Ksh/month)	7,276	11,990	0	236,000
Ratio income from livestock	0.69	0.42	0	1
Savings (Ksh)	3,810	35,100	0	1,515,000
Asset Index	-0.19	0.79	-0.94	5.69
Number of social groups	0.59	0.82	0	6
Extremely Risk Averse[1]	0.25		0	1
Moderately Risk Averse[1]	0.47		0	1
Risk Neutral/Low Risk Aversion[1]	0.29		0	1

Notes. [1]Dummy variable =1 if true.

The asset index is constructed using factor analysis of a list of important household assets and characteristics in the spirit of Sahn and Stifel (2000). Included are counts of assets that fall into very small, small, medium, and large assets. Small, medium, and large categories are also each divided into two categories according to use (e.g., productive vs. other). There are also indicators of water source, household construction, lavatory facilities, fuel sources, education, cash on hand, land holdings, poultry, and donkeys. Cattle, camels, goats, and sheep are not included in the index as they are captured directly in herd size. The factor loadings are found in Table E3.

Table E3. Factor loadings estimated by factor analysis and used to generate an asset index

Variables	Factor Loading
Improved Wall	0.1324
Improved Floor	0.1302
Improved Toilet	0.1285
Improved Light	0.1178
Improved cooking appliance	0.0766
Improved Fuel	0.0643
Improved furniture	0.1650
Water Source: Open	0.0039
Water Source: Protected	0.0042
Water Source: Borehole	-0.0082
Water source: Tap	0.0398
Water Source: Rainwater catchment	0.0792
Water Source: Tanker	0.0214
Education	0.1214
Total cash savings	0.0851
Land	0.0511
Irrigation	0.0331
Poultry	0.0814
Donkeys	0.0188
Very small	0.0397
Small tools	0.1263
Small other	0.0531
Medium tools	0.1636
Medium other	0.1351
Large	0.0373
Large with motor	0.0891

Notes. Division-period dummies included in factor analysis.

Appendix F: First Differences Robustness Check

If there are time-invariant household level fixed effects, the estimates found in Table 5 may be biased. Taking advantage of the panel characteristic of the survey data, we re-estimate with a fixed effects estimator. This method requires within-household variation in all the variables of interest, so re-analysis is necessarily restricted to that found in column 2, Table 5, excluding time invariant characteristics. The results are qualitatively the same as the pooled results. There are very few statistically significant relationships between the household characteristics examined here and idiosyncratic losses and those characteristics explain very little of the variation in idiosyncratic losses.

Table F1. Fixed effects regression of factors that contribute to idiosyncratic livestock mortality losses

VARIABLES	**Idiosyncratic Losses**
Household size	1.0130
	(0.9695)
Dependency ratio	0.0253
	(0.0556)
Asset index[#]	-0.0892
	(0.2293)
Asset index squared[#]	0.8592
	(0.6611)
HSNP participant	-0.0052
	(0.0129)
% herd camels[&]	0.0304
	(0.0365)
% herd cattle[&]	0.0245
	(0.0387)
Herd size (TLU/100)[&]	-0.0862
	(0.0949)
Herd size^2 (TLU^2/100^2)[&]	0.0145
	(0.1068)
Herd size^3 (TLU^3/100^3)[&]	-0.0121
	(0.0270)
Ratio income from livestock[#]	-0.0245*
	(0.0137)
Savings (KShs/1,000)[#]	0.0042**
	(0.0020)
Social groups count[#]	0.0021
	(0.0090)
Constant	-0.0748
	(0.0700)
Observations	5,124
Number of households	736
R-squared	0.011

Notes. [#]Variable is lagged by one period in order to reduce potential endogeneity. [&]Variable uses seasonal average monthly herd size. Cluster-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, *p<0.1

Appendix G: Factors that contribute to net benefits due to IBLI coverage

Although the focus of this paper is to decompose basis risk in order to examine the factors that drive each of its components, policy makers may be more interested in how those factors net out to drive net benefits. A thorough examination of the impacts of IBLI on household welfare falls well beyond the scope of this paper (see Jensen, Mude & Barrett 2014 for such an analysis). But we can use the processes used in this paper to examine idiosyncratic risk to also examine the factors that contribute to net improvement to outcomes. Here, as above, we define household period outcomes without insurance as the livestock mortality rate and the outcome with insurance as the net of livestock mortality rates, premiums paid, and indemnity payments received. The loaded unsubsidized seasonal premium rate is used here.

Our variable of interest is then the reduction to risk due to insurance coverage. If we assume that households are risk averse, they place more weight on marginal differences to large shocks than to small shocks, which we approximate by squaring the outcome variable. Because squaring suppresses the distinction between overpayments and underpayments, the period specific net outcome is restricted to be greater than or equal to zero.[39] The average of the difference of the eight period sum of squares (Equation G1), is then our approximation of benefits.

$$(G1) \qquad Benefit_{id} = \frac{1}{T-1}\left\{\sum_{t=1}^{T}(M_{idt})^2 - (\max[M_{idt} + Premium_{idt} - Indemnity_{idt}, 0])^2\right\}$$

To determine the factors that contribute to net benefits, we perform an analysis similar to that found in columns (4)-(6) in Table 9 but also include an analysis only controlling for index division fixed effects. We expect that the factors that are associated with improved net benefits are correlated with either idiosyncratic risk or design risk. Because design risk is determined at the division level, any household characteristic that contributes significantly to net benefits but not to idiosyncratic risk is likely to be a reflection of regional differences between households, such as ethnicity and related herding practices or access to environmental services.

The estimates can be found in Table G1. Division level fixed effects capture 29% of the variation in benefits between households. According to the earlier analysis of design risk, we expect significant and large differences between the accuracy of the index across divisions. It is not surprising that variation in index accuracy manifests as variation in average benefits. In both the design risk analysis and in this benefit analysis, IBLI performs much better in Central/Gadamoji and Maikona divisions than in Laisamis or Loiyangalani.

[39] This definition is the same as semi-variance described in the body of the text, placing the target at zero and the indicator function is equal to one if the period specific net outcome is greater than zero and zero otherwise.

Table G1. Factors associated with net benefits (reduction to semi-variance) due to IBLI coverage

	$Benefit_{id}$			
VARIABLES	(1)	(2)	(3)	(4)
Age (/100)			0.0447***	0.0075
			(0.0167)	(0.0128)
Age^2 ($age^2/100^2$)			-0.0392***	-0.0069
			(0.0142)	(0.0111)
Gender (=1 if male)			0.0019	0.0001
			(0.0013)	(0.0011)
Household size (count/100)			-0.0259	0.0161
			(0.0310)	(0.0254)
Dependency ratio			0.0011	-0.0012
			(0.0031)	(0.0022)
Asset index (/10)			0.0325**	-0.0207
			(0.0156)	(0.0169)
Asset $index^2$ ($/10^2$)			-0.1192**	-0.0406
			(0.0596)	(0.0521)
HSNP participant			-0.0019*	0.0010
			(0.0011)	(0.0013)
Ratio herd camels			0.0018	-0.0049
			(0.0033)	(0.0032)
Ratio herd cattle			0.0052	-0.0046
			(0.0057)	(0.0035)
Herd size (TLU/100)			0.0197	0.0374***
			(0.0136)	(0.0121)
Herd $size^2$ ($TLU^2/100^2$)			-0.0397	-0.0653***
			(0.0263)	(0.0239)
Herd $size^3$ ($TLU^3/100^3$)			0.0220*	0.0333***
			(0.0123)	(0.0115)
Ratio income from livestock			-0.0142***	-0.0038
			(0.0026)	(0.0027)
Log (1+Savings)			-0.0002	0.0002
			(0.0003)	(0.0003)
Social groups (count)			0.0002	-0.0000
			(0.0010)	(0.0010)
Moderately risk averse			-0.0014	0.0001
			(0.0014)	(0.0011)
Extremely risk averse			0.0013	0.0013
			(0.0015)	(0.0012)
Division Fixed Effects (4)	Yes	No	No	No
Sublocation Fixed Effects (16)	No	Yes	No	Yes
F-stat testing: All location fixed effects=0	77.37***	24.71***		13.10***
Observations	736	736	736	736
R-squared	0.304	0.357	0.165	0.386

Notes. Regression also included an intercept term. Household clustered-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Moving the fixed effects to the sublocation level accounts for an additional 5.5% of the variation in benefits, indicating that there is some within-division variation in benefits. But most of the geographically defined variation takes place between index divisions.

As expected from the idiosyncratic risk analysis, household characteristics contribute very little to net benefits. Placing these findings in the context of our decomposition of basis risk, the majority of the basis risk or net benefits associated with IBLI coverage cannot be easily attributed to any of the household characteristics that we observe. Rather, it seems that the portion of net benefits (or basis risk) that we can account for are due to significant design risk in the product and losses that are correlated at a smaller geographic scale than the index scale. The remaining idiosyncratic component seems to be mostly random. Thus geographic targeting of a covariate loss product like IBLI seems the best strategy for maximizing net benefits.

References

Alderfer, C. P., & Bierman, H. (1970). Choices with risk: Beyond the mean and variance. *Journal of Business*, *43* (3), 341-353.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, *62*(3), 395-407.

Chantarat, S., Mude, A. G., Barrett, C. B. and Carter, M. R. (2013). Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya. *Journal of Risk and Insurance*, 80, 205–237.

Clarke, D. J. (2011). *A theory of rational demand for index insurance*. Department of Economics D. Phil. dissertation, University of Oxford.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, *92*(5), 1644-1655.

Jensen, N., Mude, A., & Barrett, C. (2014). Index Insurance and Cash Transfers: A Comparative Analysis from Northern Kenya. Unpublished.

Meyer, J. (1987). Two-moment decision models and expected utility maximization. *American Economic Review*, 421-430.

Miranda, M. J., & Farrin, K. (2012). Index insurance for developing countries. *Applied Economic Perspectives and Policy*, 34(3), 391-427.

Sahn, D. E., & Stifel, D. C. (2000). Poverty comparisons over time and across countries in Africa. *World development*, *28*(12), 2123-2155.

Turvey, C. G. (1992). An economic analysis of alternative farm revenue insurance policies. *Canadian Journal of Agricultural Economics*, *40* (3), 403-426.

Woodard, J. D., Pavlista, A. D., Schnitkey, G. D., Burgener, P. A., & Ward, K. A. (2012). Government Insurance Program Design, Incentive Effects, and Technology Adoption: The Case of Skip-Row Crop Insurance. *American Journal of Agricultural Economics*, *94*(4), 823-837.