How Basis Risk and Spatiotemporal Adverse Selection Influence Demand for Index Insurance:
Evidence from Northern Kenya

Nathaniel D. Jensen[*†], Andrew G. Mude[‡] and Christopher B. Barrett[†]
[†] *Cornell University, Ithaca, NY, USA;* [‡] *International Livestock Research Institute, Nairobi, Kenya*

October 2014

Abstract: Index insurance has emerged as a popular method for mitigating weather-related shocks among farmers and herders in low-income rural settings, but has had routinely exhibited low uptake rates. Evidence from several index insurance pilot projects in developing countries suggests that the supply and price of insurance influence demand, as do non-price factors such as liquidity, access to alternative risk pooling mechanisms, and product understanding. Basis risk – the remaining risk that an insured individual faces – is widely acknowledged as the Achilles Heel of index insurance, but to date there has been no direct study of its role in determining demand for index insurance. Further, insurers' practice of spatiotemporal aggregation in pricing index insurance products leaves open the possibility of a form of adverse selection that naturally attenuates demand. We use rich longitudinal household data from northern Kenya that enable us to control for individual-level basis risk and spatiotemporal variation in covariate loss rates to determine which factors affect demand for index based livestock insurance (IBLI) in that region. We find that both price and the non-price factors studied previously are indeed important, but that basis risk and spatiotemporal adverse selection play a major role in dampening demand for index insurance. The average marginal effect of our proxy for division average idiosyncratic risk is to reduce the likelihood of uptake by 29.6% while households that purchase coverage but expect good rangeland conditions purchase 11.2% less coverage than those that expect poor rangeland conditions. A Shapley's goodness-of-fit decomposition finds that the combined role of product characteristics and adverse selection are at least as important as household characteristics in determining demand for IBLI.

[*]Corresponding author: Nathaniel D. Jensen ndj6@cornell.edu.
This research uses data collected by a collaborative project of the International Livestock Research Institute, Cornell University, Syracuse University and the BASIS Research Program at the University of California at Davis. The authors wish to specifically thank Diba Galgallo, Munenobu Ikegami, Samuel Mburu, and Mohamed Shibia for their remarkable efforts to collect useful and accurate data. This project was made possible by the generous funding of 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 and the World Bank's Trust Fund for Environmentally and Socially Sustainable Development Grant No: 7156906. This paper represents the views of its authors alone and not the positions of any supporting organizations. Any remaining errors are our sole responsibility.

1. Introduction

Risk management interventions have become a priority for development agencies as the enormous cost of uninsured risk exposure, especially to the rural poor, has become increasingly widely appreciated. Improved risk management through innovative insurance products is hypothesized to crowd in credit access, induce investment, support informal social transfers, and generally stimulate growth and poverty reduction (Hess et al. 2005; Skees, Hartell & Hao 2006; Barrett et al. 2007; Barnett, Barrett & Skees 2008; Boucher, Carter & Guirkinger 2008; Skees & Collier 2008; Giné & Yang 2009; Hellmuth et al. 2009; Karlan et al. 2014). Although insurance products offer a proven means to manage risk through formal financial markets, the asymmetric information problems—adverse selection and moral hazard—and high fixed costs per unit insured effectively preclude conventional indemnity insurance for smallholder crop and livestock farmers in developing countries.

Index insurance products have flourished over the past decade as a promising approach to address these obstacles. Index insurance products use easily observed, exogenous signals to provide coverage for covariate risk. Anchoring indemnity payments to external indicators, not policyholder's realized losses, eliminates the need to verify claims, which is particularly costly in remote areas with poor infrastructure and clients with modest covered assets, and mitigates the familiar incentive challenges associated with moral hazard and adverse selection that plague traditional insurance. These gains do come at the cost, however, of "basis risk", defined as the residual risk born by insurees due to the imperfect association between experienced losses and indemnification based on index values. Furthermore, a form of adverse selection may remain if prospective purchasers have information about upcoming conditions that affect insured, covariate risk – such as climate forecasts – but that information is not incorporated into the index insurance product's pricing (Carriquiry & Osgood 2012).

The explosion of interest in index insurance has resulted in a proliferation of pilot programs across the developing world. A burgeoning literature addresses various aspects of theoretical and applied concerns in the design, implementation, and assessment of index insurance products (e.g., Barnett & Mahul 2007; Barrett et al.2007; Binswanger-Mkhize 2012; Chantarat et al. 2007; Clarke 2011; Miranda & Farrin 2012). Despite the celebrated promise of index insurance, uptake in pilot programs around the globe has been far below anticipated levels, and there are as of yet no examples of clear success stories with demonstrable capacity for scalability or sustainability over the long run (Smith & Watts 2010; Hazell & Hess 2010; Leblois & Quiron 2010). As a result, most empirical research on index insurance in developing countries has focused on identifying the barriers to insurance uptake. Although demand appears to be price sensitive, as expected, studies find considerable variation in the price elasticity of demand, ranging from -0.44 to -1.16 (Mobarak & Rosenzweig 2012; Cole et al. 2013; Hill, Robles & Ceballos 2013). And, with the exception of the Ghanaian farmers studied by Karlan et al. (2014), uptake has been low even at heavily subsidized prices.[1] With evidence that price plays only a small part in determining demand,

[1] The high demand for rainfall insurance in Ghana is somewhat of a mystery. Karlan et al. (2014) point to the role that insurance grants and indemnity payments play, but those same processes have been observed elsewhere unaccompanied by similar high levels of demand.

researchers have turned to examining the role of non-price factors. Risk aversion, wealth, financial liquidity, understanding of the product, trust in the provider, and access to informal risk pooling commonly exhibit significant, although sometimes inconsistent, impacts on demand (Giné, Townsend & Vickery 2008; Chantarat et al. 2009; Pratt, Suarez & Hess 2010; Cai, de Janvry & Sadoulet 2011; Clarke 2011; Janzen, Carter & Ikegami 2012; Liu & Myers 2012; Mobarak & Rosenzweig 2012; Cole et al. 2013; McIntosh, Sarris, & Papadopoulos 2013; Dercon et al. 2014).

Although basis risk and the possibility of spatiotemporal adverse selection are widely understood as prospective weaknesses of index insurance, the empirical research has thus far not directly explored the role that either of these factors plays in influencing product uptake. But if the insurance index is imperfectly correlated with the stochastic welfare variable of interest (e.g., income, assets), then index insurance may offer limited risk management value; indeed it can increase, rather than decrease, purchasers' risk exposure (Jensen et al. 2014). Furthermore, prospective purchasers may perceive that an index insurance product is mispriced for their specific location or for the upcoming season, given information they have on covariate risk for the insured period and place.

Both of these problems exist generally in index insurance contracts and either should adversely affect uptake. A recent review of index insurance pilots concludes, "[g]iven the central role played by basis risk in determining benefits of and demand for index insurance, at least some modest efforts should be made to assess its magnitude" (Miranda & Farrin 2012, p.422). Yet the impact of these prospective weaknesses in index insurance products has not been carefully researched to date, although a few studies use coarse proxies for idiosyncratic risk (e.g., Karlan et al. 2014; Mobarak & Rosenzweig 2012). This lacuna arises primarily because the vast majority of products fielded to date remain unable to determine the level of basis risk inherent in their product design; the products were designed from data series on index variables (e.g., rainfall, crop growth model predictions), not from longitudinal household asset or income data from the target population to be insured.

This paper fills that gap, exploiting an unusually rich longitudinal dataset from northern Kenya and the randomization of inducements to purchase index-based livestock insurance (IBLI), a product designed from household data to minimize basis risk (Chantarat et al. 2013), in order to identify the impact of basis risk and spatiotemporal selection on index insurance uptake. We further distinguish between the two central components of basis risk, design error – associated with the imperfect match between the index and the covariate risk the index is meant to match – and idiosyncratic risk – individual variation around the covariate experience. Design error can be reduced by improving the accuracy of the index, while idiosyncratic risk inherently falls outside the scope of index insurance policies. We find that both spatiotemporal selection and basis risk are important and, in particular, that although correctable design error plays a role, the main demand effects arise due to idiosyncratic risk intrinsic to any index insurance product.

Using longitudinal seasonal and annual household survey data collected over four years, we follow demand over multiple seasons as households learn about the product and the basis risk components that they face. Notably, uptake was healthy during the first sales window (27.8% of the sample purchased), but has dropped off rather dramatically in the following sales periods. Echoing the prior literature, we find that price, liquidity, and social connectedness affect demand in the expected ways. In addition, we find that basis risk and spatial adverse selection associated with division average basis risk dampen demand for IBLI. Households in divisions with greater average idiosyncratic risk are much less likely to purchase any insurance than those in divisions with relativity more covariate risk. Design error also plays a role in demand, reducing uptake and increasing price sensitivity among those who purchase coverage. But, between the two components of basis risk, design risk plays a much smaller role, reducing uptake by an average of less than one percent (average marginal effect [AME] =-0.0073, Std. Err.= 0.0025) while the division average covariance between individual and covariate losses effects uptake by nearly 30% on average (AME = 0.2964, Std. Err.=0.1617). Consequently the basis risk problem is not easily overcome through improved product design. There is also strong evidence of intertemporal adverse selection as households purchase less coverage, conditional on purchasing, before seasons for which they expect good conditions (AME=-0.2709, Std. Err.=0.0946). This impact represents an 11.2% reduction in average demand among those purchasing.

The remainder of the paper is organized as follows. Section 2 discusses risk among pastoralists and the motivation for and design of the IBLI product offered to them. Section 3 develops a stylized model of livestock ownership and the role of insurance so as to understand the structural determinants of demand. Section 4 presents the research design and data followed by an explanation and summary of key variables in Section 5. Section 6 describes the econometric strategy used to analyze demand for IBLI. The results are discussed in Section 7.

2. Drought-Related Livestock Mortality & Index Insurance in Northern Kenya

A first order concern in the design of an optimal insurance index is that it significantly reduces risk borne by the target population and that the index covaries strongly with observed losses. The IBLI Marsabit product expressly covers predicted area average livestock mortality due that arise due to severe forage shortages associated with drought precisely because drought-related livestock mortality has consistently emerged as the greatest risk faced by pastoralists in the arid and semi-arid lands (ASAL) of the Horn of Africa (McPeak & Barrett 2001; McPeak, Little & Doss 2012).

Livestock not only represent the principal source of income across most ASAL households (mean=69% and median=95% in our data) but also constitute the highest value productive asset they own. Livestock face considerable mortality risk, rendering ASAL households particularly vulnerable to herd mortality shocks. Among these, drought is by far the greatest cause of mortality, and drought-related deaths largely occur in times of severe forage shortages. For example, between June 2000 and June 2002, surveyed pastoralists reported that drought-related factors accounted for 53% of the livestock deaths that they experienced, and disease, which is often associated with droughts, caused an additional 30% mortality during that period (McPeak, Little & Doss 2012). Drought is the cause of 62% of the reported livestock mortality in our 2009-12 sample from northern Kenya. Droughts represent a covariate risk that may be especially difficult for existing social risk pooling schemes to handle because losses can impact all members of the risk pool. At the same time, the seemingly largely covariate risk profile pastoralists face seems well-suited for coverage by an index product.

Launched in January 2010 in the arid and semi-arid Marsabit District to address the challenge of drought-related livestock mortality, the index based livestock insurance (IBLI) product is derived from the Normalized Difference Vegetation Index (NDVI), an indicator of photosynthetic activity in observed vegetation as reflected in spectral data remotely sensed from satellite platforms at high spatiotemporal resolution (Chantarat et al. 2013). These NDVI data are reliably and cheaply accessible in near real-time, and with sufficiently long historical record to allow for accurate pricing of the IBLI product (Chantarat et al. 2013). The statistical relationship between NDVI and livestock mortality was estimated using historic household level livestock mortality rates and NDVI values from January 2000 through January 2008 and then tested out-of-sample against a different set of seasonal household panel data collected 2000-2 in the same region.[2] The resulting response function generates estimates of division average livestock mortality rate on the basis of NDVI inputs.[3] IBLI appears to be the only index insurance product currently on the market that was developed using longitudinal household data so as to minimize the design component of basis risk.[4]

A commercial underwriter has offered IBLI contracts written on this predicted livestock mortality rate index (see Chantarat et al. 2013 for more details on data and product design). The index is calculated separately for each of the five administrative divisions in Marsabit, allowing for variation between divisions. The commercial underwriter set a single strike level—the index level at which indemnity payments are initiated—at 15% livestock mortality and aggregated the five index divisions into two premium regions. Notably, the aggregation of index divisions into premium regions results in variation in loadings/subsidies between index divisions, opening the door for spatial adverse selection.[5] A detailed summary of the contract parameters (e.g., geographical segmentation of coverage, temporal coverage of the contract, conditions for contract activation, indemnification schedule, pricing structure) is presented in Appendix A.

During the first sales season in January 2010, the IBLI product sold 1,974 policies to cover the long rain/long dry season of 2010 (LRLD10) and following short rain/short dry season (SRSD10), from March 1, 2010-Ferburary 28, 2011. The intention was to have a sales window during the two-month period before the onset of each bimodal rainy season. Due to logistical and contractual complications, IBLI was

---

[2] Monthly household-level livestock mortality data were collected by the Arid Lands Resource Management Project (ALRMP, http://www.aridland.go.ke/). The seasonal household panel data used for out-of-sample evaluation come from the Pastoral Risk Management project (http://dyson.cornell.edu/special_programs/AFSNRM/Parima/projectdata.htm).

[3] "Divisions" are existing administrative units in Kenya that IBLI has used to define the geographic boundaries of a contract. Division boundaries are suitable because they are large enough to reduce moral hazard to a negligible level, small enough to capture a large portion of covariate risk, and were well known by pastoralists.

[4] An index based livestock insurance program in Mongolia, which protects pastoralists from the risk of severe winters known as dzud, seems to have been designed off area average herd mortality rates (see Mahul & Skees 2007 for a full description of the IBLI Mongolia project). As of writing, the Mongolian program has yet to make its findings public so we are unable to use the similarities between programs to inform this research.

[5] The aggregation of index divisions into premium regions had been dropped in the newer IBLI products.

not available for purchase during the August/September 2010 or January/February 2012 periods. In total, there have been four sales windows and six seasons of coverage during the timeframe considered in this paper. Table 1 presents summary statistics for IBLI sales over the four rounds that fall within our sample period.

There was a consistent fall in IBLI uptake over the 2010-2012 period. Although inconsistency of sales windows, a change in the commercial insurance provider, and variation in extension and sales protocols are likely to have depressed sales, heterogeneity in demand suggests that other factors also influenced purchases. Tracking household purchase patterns across seasons shows considerable variation in when households make their first purchase, if they continue to purchase, or if they allow their contract to lapse (Table 2). Such behavior suggests dynamic factors play a significant role in insurance demand. In the next section, we offer a simple model of index insurance demand and examine the role that basis risk and spatiotemporal adverse selection could play in determining demand.

3. Demand for Index Based Livestock Insurance

This section sets up a simple model of household demand for insurance that offers a set of empirically testable hypothesis concerning basis risk and spatiotemporal adverse selection. This is meant merely to motivate the empirical exploration that is this paper's primary contribution. So we simplify this as a static problem under uncertainty and ignore dynamic considerations in the interests of brevity.

Let households maximize their expected utility, which is an increasing and concave von Neumann-Morgenstern function that satisfies U'>0, U''<0. Utility has wealth, measured as end-of-period herd size in tropical livestock units (TLU), as its argument.[6] Households have an initial livestock endowment, TLU_0, but the herd is subject to stochastic losses (L). Households have the option of purchasing livestock insurance at the rate of p per animal insured ($\widetilde{tlu}$) where $\widetilde{tlu}$ is in units of livestock and $p \in [0,1]$.[7] The insurance makes indemnity payments according to an index, which is the predicted rate of division average livestock losses ($I \in [0,1]$).[8] The utility maximization problem and budget constraint can be described as follows, where E is the expectation operator.

$$\max_{\widetilde{tlu}} E[U(TLU)], \tag{1}$$

$$subject\ to: TLU = TLU_0 - L - \widetilde{tlu} * p + \widetilde{tlu} * I$$

[6] Tropical livestock units (TLUs) are a conversion rate used to aggregate livestock. The IBLI contracts use the conversion rate of 1 TLU = 0.7 camels = 1 cattle = 10 sheep or goats as suggested by the FAO Livestock and Environment Toolbox (1999).

[7] The premium and index are defined as ratio of the value insured to avoid the need to place a monetary value on livestock. This specification is appropriate in the context of livestock insurance in northern Kenya because households often sell off a small animal in order to purchase insurance on remaining animals. If the cost of insuring one animal was equivalent to the value of the animal, p=1.

[8] The division refers to the geographic region defined by the insurance product.

Normalize the variables $TLU, TLU_0, L, \widetilde{tlu}$ by TLU_0 so that they are now all expressed as proportions of the household's initial herd endowment. Substituting the budget constraint into the utility function and using a second order Taylor expansion allows us to approximate the expected utility maximization problem as a function of original livestock endowment and deviations from the endowment associated with losses, premium payments and indemnity payments.[9] The necessary first order condition becomes

(2) $$E\left[U'(TLU_0)(-p+I) + U''(TLU_0)\left[Lp - L*I + \widetilde{tlu}*p^2 - 2p*I*\widetilde{tlu} + \widetilde{tlu}*I^2\right]\right] = 0$$

The first order condition can be solved for optimal insurance purchases. We use the representations $E[x] = \bar{x}$, $Cov(x,y)$= the covariance of x and y, and $Var(x)$ = variance of x, where x and y are representative variables. In addition, we use U=U(TLU_0) to reduce simplify notation. With some algebra, the optimal number of animals to ensure can be written as equation (3).

(3) $$\widetilde{tlu}^* = \frac{U''[\bar{L}(\bar{I}-p) + Cov(I,L)] - U'(\bar{I}-p\,)}{U''\left((\bar{I}-p)^2 + Var(I)\right)}$$

If premiums are actuarially fairly priced, then the premium rate is equal to the expected index value ($\bar{I} = p$). In that case, optimal coverage is $\widetilde{tlu}^* = \frac{Cov(I,L)}{Var(I)}$, which is greater than zero as long as the covariance between the index and losses is positive. If the insurer adds loadings to the policy premium so that $\bar{I} < p$, then optimal insurance purchase volumes can be zero even when the index is positively correlated with household losses.

Basis risk

If there is no basis risk ($cov(I,L) = Var(I)$) and the premiums remain actually fair, then the index and losses are identical and $\widetilde{tlu}^* = 1$, i.e., full insurance is optimal. As the covariance between the index and individual losses falls, however, so does optimal coverage ($\frac{d\,\widetilde{tlu}^*}{dCov(I,L)} = \frac{1}{Var(I)} > 0$).

To more closely examine the role that basis risk plays, let the index equal individual losses multiplied by a coefficient, a constant, and a random error term ($I = \beta_0 + \beta_1 L + \varepsilon$). The expected difference between the index and losses (expected basis error) is captured by the relationship $\beta_0 + \beta_1 L$, in particular deviations from the null $\beta_0 = 0$ and $\beta_1 = 1$, while $Var(\varepsilon)$ is the variance in basis error.

Because the covariance between the error term and losses is zero by construction, optimal coverage for actuarially fairly priced index insurance with basis risk is $\widetilde{tlu}^* = \frac{\beta_1 Var(L)}{\beta_1 Var(L) + Var(\varepsilon)}$. Clearly, as the variance

[9] $\max_{\widetilde{tlu}} E\left[U(TLU_0) + U'(TLU_0)(-L - \widetilde{tlu}*p + \widetilde{tlu}*I) + \frac{1}{2}U''(TLU_0)(-L - \widetilde{tlu}*p + \widetilde{tlu}*I)^2\right]$

in basis error increases, demand falls. Alternatively, as β_1 increases so does demand as long as there is some variance in basis error ($Var(\varepsilon) \neq 0$).[10] At actuarial fair premium rates with no variance in basis error, households can adjust their purchase levels to account for expected basis error at no change to expected net costs, and full coverage continues to be optimal.

Relaxing the premium constraint, let premiums be set so that $p + \delta = E[I]$, where δ represents the net loading on the policy, reflecting any subsidy to purchasers less than loading above actuarially fair rates (i.e., expected indemnity payments) by the underwriter. Thus, if there is a net subsidy, $\delta > 0$, while if the premiums are loaded beyond the subsidy, $\delta < 0$. Optimal coverage is not monotonic in premium rates because changes to premium rates not only effect the opportunity cost of premium payments but also have wealth effects that are ambiguous in the impact on demand ($\frac{\partial \widetilde{tlu}^*}{\partial \delta} = \frac{\{U''\bar{L} - U'\}}{D} - \frac{2\delta U''\{U''[\bar{L}(\delta)+\beta_1 Var(L)] - U'(\delta)\}}{D^2}$). Clarke (2011), discusses a similar outcome.

Adjusting the earlier model with basis risk to allow for variation in premium rates, optimal coverage is now $\widetilde{tlu}^* = \frac{U''[\bar{L}(\delta)+\beta_1 Var(L)] - U'(\delta)}{[U''((\delta)^2+\beta_1 Var(L)+Var(\varepsilon))]}$ and demand still falls with increased variance in basis error.[11] The importance of basis risk might also change with prices. Analytically, we find that demand response to basis risk changes with premium rates but is also subject to the ambiguous wealth effects (equation 4).

(4)
$$\frac{\partial^2 \widetilde{tlu}^*}{\partial p \partial Var(\varepsilon)} = \frac{U''(U' - U''\bar{L})}{D^2} + \frac{4{U''}^2 \delta N}{D^3} \leq 0$$

Where $D = U''[\delta^2 + \beta_1 Var(L) + Var(\varepsilon)]$ and $N = U''[L\delta + \beta_1 Var(L)] - U'\delta$. This leads to

Hypothesis 1: As basis risk grows, demand falls, and that response changes with premium levels.

We also expect that the impact of the premium changes with basis risk in the same direction as $\frac{\partial^2 \widetilde{tlu}^*}{\partial p \partial Var(\varepsilon)}$ due to symmetry of cross partials in the Hessian matrix. This is consistent with Karlan et al.'s (2014) finding that households were less responsive to price incentives in regions with low product quality (high design error).

In some cases it may be that households do not understand the insurance product well. For example, a household might think that the insurance product indemnifies all losses or that indemnity payments are always made at the end of every season. In either of these cases, basis risk should play no role in the purchase decision, although it could have a large impact on the eventual welfare outcomes of the purchase decision. Between those two extremes, there may be households that partially understand the

---

10 $\frac{d\widetilde{tlu}}{d\beta} = \frac{var(L)*Var(\varepsilon)}{(\beta_1 Var(L)+Var(\varepsilon))^2} \geq 0$. There is a discontinuity in demand where $\beta_1 = -\frac{Var(\varepsilon)}{Var(L)}$ but demand is increasing with β_1 on either side of the discontinuity.

11 $\frac{\partial \widetilde{tlu}^*}{\partial var(\varepsilon)} = -\frac{U''\{U''[\bar{L}(\delta)+\beta_1 Var(L)] - U'(\delta)\}}{D^2} \leq 0$

insurance contract but have some misconceptions. Let an individual's understanding of the product be summarized by the term ($I_i = I + z_i$) where I continues to the index that determines indemnity payments, $z_i$reflects the individual's misinformation and I_i is the index required to produce the indemnity payment that the individual expects to receive. Assuming actuarially fair premium rates, the optimal purchase is $\widetilde{tlu}^* = \frac{Cov(I,L)+Cov(z,L)}{Var(I)+Var(z)+2*Cov(I,z)}$. If the misconceptions are negatively and highly correlated with the index, the consumer's optimal purchases could increase with increased basis risk.[12] Otherwise, households with misconceptions reduce optimal purchases with increased basis risk but that response is mitigated by basis risk.[13] This relationship leads to our next hypothesis:

Hypothesis 2: Poor understanding of the product moderates the negative demand response to increases in basis risk. At the most extreme levels of misinterpretation of the contracts, households may not respond at all to basis risk or might increase demand with basis risk.

Spatiotemporal Adverse Selection

Indemnifying covariate losses, rather than individual losses, eliminates the prospective impact on insurer profits of within index-division cross-sectional adverse selection by decoupling indemnity payments from individual losses.[14] But group-level adverse selection can reemerge if households have information on the likelihood of an indemnity payment in the coming season that is not reflected in the premium. For example, ecological conditions during the sales window may have predictive power as to the likelihood of an upcoming drought. In this case, the consumer has a signal (observed ecological conditions) that provides information on the distribution of coming average losses and thus the likelihood of indemnity payments, and that information was not incorporated in the product's pricing. Even in cases when the insurer can observe the same information that households can, contracts are not always written with variable premium rates. Rather, insurers and reinsurers often set prices according to historic averages and are commonly reluctant to change premiums season by season.

Such intertemporal adverse selection can be incorporated into the above model. Assume that a household observes a signal before purchasing insurance that provides information on the likelihood of certain end-of-season rangeland conditions that could affect the index for this specific season ($E[\text{I}]$) and/or the mortality rate at the end of this season ($E[L^*]$). Let x^* be the household's interpretation of the signal as an adjustment to the index $[I^*] = E[I] + x^*$ and y^*be the household's interpretation of the signal as an adjustment to her own expected livestock mortality rate ($E[L^*] = E[L] + y^*$) where $x^*, y^* \in [-1,1]$. We can then rewrite 3 as (3').

(3')
$$\widetilde{tlu} = \frac{U''[(\bar{L} + y^*)(\bar{I} + x^* - p) + cov(I,L)] - U'(\bar{I} + x^* - p\,)}{[U''\big((\bar{I} + x^* - p)^2 + Var(I)\big)]}$$

---

12 $\frac{d\,\widetilde{tlu}^*}{dCov(I_i,L)} = \frac{1}{Var(I)+Var(z)+2*Cov(I,z)} < 0$ if $Var(I) + Var(z) + 2 * Cov(I,z) < 0$

13 $\frac{d\,\widetilde{tlu}^*}{dCov(I_i,L)} < \frac{d\,\widetilde{tlu}^*}{dCov(I,L)}$ if $Var(z) + 2 * Cov(I,z) > 0$

14 For the same reasons, index insurance reduces the incentives for moral hazard.

If the signal pertains only to individual losses ($x^* = 0$) , $\frac{d\widetilde{tlu}}{dy^*} = \frac{\bar{I}-p}{((\bar{I}-p)^2+Var(I))}$ which has the sign of $\bar{I} - p$ and is identical to a change in long-run livestock losses ($\bar{L}$). Households that believe they will lose livestock at a greater rate in the following season will increase purchases if premiums are subsidized and reduce purchases if premiums are loaded. This leads directly to our third core, testable hypothesis:

Hypothesis 3: Households will respond to signals of increased losses by increasing purchases if premiums are below the actuarially fair rate.

By contrast, if the signal pertains only to the expected index, the outcome is similar to changes in loadings/subsidies and is not monotonically increasing or decreasing in x^*.[15] But, just as with the ambiguous impact of premium rates on optimal purchases, we can learn about impact of x^* through its impact on $\frac{d\widetilde{tlu}}{dy^*}$. The cross partial, $\frac{\partial^2 \widetilde{tlu}^*}{\partial x^* \partial y^*} = \frac{U''^2[Var(I)-(\bar{I}+x-p)^2]}{[U''((\bar{I}+x^*-p)^2+Var(I))]^2}$, inherits its sign from $Var(I) - (\bar{I} + x - p)^2$. If, for example, $\bar{I} = p$ and the household receives a signal of increased losses and higher index, then $\frac{d\widetilde{tlu}}{dy^*} > 0$ and $\frac{d\widetilde{tlu}}{dy^*}$ increases with x^* until ${x^*}^2 = Var(I)$ and then $\frac{\partial^2 \widetilde{tlu}^*}{\partial x^* \partial y^*} \leq 0$. As with the effects of premiums on demand, the impact of signals that inform on both losses and index levels is an empirical question. If those signals correctly predict coming conditions, such behavior will be evident in a correlation between demand and index value.

A related, spatially defined form of group-level adverse selection can occur when index performance or the difference between the expected index value and the premium varies between distinct geographic regions.[16] Differences between expected indemnity payments and the premium are likely to be common for products with little data with which to estimate the expected indemnity payment. It is, in essence, variance in subsidy/loading rates between divisions caused by error in the provider's estimated expected index values or perhaps intentionally (e.g., variation in state subsidy rates). This type of spatial adverse selection is covered in the above examination of the effects of varying the subsidy/loadings.

A second type of spatial adverse selection can occur if there is variation in the basis risk between index regions. That is, there may be very little basis risk in one division and a great deal in another even as subsidy/loading rates are similar. As was shown above, regions with higher basis risk are expected to have less demand, all else being equal. This generates our fourth core hypothesis:

Hypothesis 4: Division-level variation in basis risk will cause spatial adverse selection apparent in uptake patterns.

---

15 $\frac{\partial \widetilde{tlu}^*}{\partial x^*} = \frac{\{U''\bar{L}-U'\}}{U''((\bar{I}+x^*-p)^2+Var(I))} - \frac{2(\bar{I}+x^*-p)U''\{U''[(\bar{L}+y^*)(\bar{I}+x^*-p)+cov(I,L)]-U'(\bar{I}+x^*-p)\}}{[U''((\bar{I}+x^*-p)^2+Var(I))]^2}$

16 Within geographic regions there may be clusters of households for whom the index performs especially well or poorly. Although the resulting variation in demand would likely have a geographic component, the within-division demand patterns have no impact on provider's profits and thus is not adverse selection.

This simple, static model conforms to our expectations of reduced demand with increased basis risk. It predicts that basis risk will be less important for those who do not understand the product well, and that as basis risk increases, the price response will change. In addition the model is easily extended to include factors that may contribute to spatial or temporally associated adverse selection. It predicts that we should expect to see variation in demand within divisions over time that is correlated with rangeland conditions during the sales windows.

4. Research Design & Data

Before any public awareness campaign began surrounding the January 2010 launch of the IBLI pilot, the IBLI research team began to implement a comprehensive household survey that annually tracks key parameters of interest such as herd dynamics, incomes, assets, livelihood portfolios, market and credit access, risk experience and behavior, demographics, health and educational outcomes, and more. The initial baseline survey was conducted in October of 2009, with households revisited annually thereafter in the same October-November period. A total of 924 households were sampled across 16 sub-locations of Marsabit District, selected to represent a broad variation of livestock production systems, agro-ecology, market accessibility and ethnic composition.[17]

A few key elements of the survey design are important to note. Two randomized encouragement treatments were implemented to help identify and test key program parameters on demand. In the first, a sub-sample was selected to play a comprehensive educational insurance game based on the pastoral production system. Participants used role playing and simulations to help learn how IBLI functions in the face of idiosyncratic and covariate shocks. The game was played in three rounds of increasing complexity to help build up a comprehensive knowledge of IBLI. The game was played in nine of the 16 sites among a random selection of half of the sample households in each selected site, and took place just before the launch of sales in January 2010.[18]

The second encouragement is an ongoing price incentive that introduces exogenous variation in premium rates. Discount coupons were randomly distributed to about 60% of the sample before each sales season. The discount provided by the coupons was evenly distributed among 10%, 20%, 30%, 40%, 50% and 60% discount levels. Upon presentation to insurance sales agents, the coupon entitles the household to the relevant discount on premiums for the first 15 TLU insured during that marketing season.[19] The coupons expire after the sales period immediately following their distribution.

The IBLI team also coordinated survey sites to overlap with the Hunger Safety Net Program (HSNP), cash transfer program launched by the Government of Kenya in April 2009 that provides regular monthly

[17] This sample was distributed across the 16 sub-locations on the basis of proportional allocation using the Kenya 1999 household population census statistics. There were only two exceptions to this rule: a minimum sample size of 30 households and maximum of 100 households per sub-location. In addition, sampling across each sub-location was also stratified by wealth class based on livestock holdings reported by key informants before the selection process.

[18] See McPeak, Chantarat, & Mude (2010) for a description of the IBLI educational game and its implementation in the field.

[19] Of the nine sample households that purchased insurance for more than 15 TLUs, six used a discount coupon for the first 15TLUs.

cash transfers to a select group of target households in the northern Kenya ASAL (Hurrell & Sabates-Wheeler 2013). The regularity and certainty of this cash transfer may impact household liquidity constraints and therefore demand for IBLI. Site selection for IBLI extension encouragement was stratified to include both communities targeted by HSNP and other, nearby communities that were not. Figure 1 displays the project's sample sub-locations across Marsabit and illustrates how they vary in terms of the noted elements of the study design. In a structure ensuring that the joint and individual influence of the IBLI game and HSNP transfers could be empirically identified, sub-locations were grouped into four different categories as shown: one group of HSNP recipients who were "game-encouraged", as explained above, and another group of HSNP recipients that were not selected to play the educational games, and another set of two groups consisting of non-HSNP recipients that were either encouraged with games or not.

This paper uses data from four annual rounds of the IBLI household survey collected in Marsabit region between 2009 and 2012. The attrition rate during this period was less than 4% in each round. An analysis of attrition is found in Appendix B. There are a number of differences between those households who remained in the survey and those who attrited (Table B.1), as well as between those who exited the survey and their replacements (Table B.2). For a discussion of the causes of attrition see ILRI (2012). We control for these characteristics in our analysis to mitigate prospective attrition bias introduced by this selection process, but the rate of exit is low enough and differences small enough that attrition should be of little worry.

It is important to note that analysis of demand is performed seasonally while the survey data were collected annually. Although seasonal data were collected for many variables through recall, some characteristics were collected for only one reference point annually. In those cases, the annual values collected in October/November are used to represent household characteristics during the March-September LRLD insurance season and the current October-February SRSD season. When estimating an average or distribution parameter (e.g., variance, covariance) all eight seasonal observations are used to estimate a single statistic, which is then treated as a constant over all periods. These details are described in more detail in the following section.

5. Discussion of Key Variables

IBLI purchases among those surveyed and within the general population across the Marsabit region were greatest in the first sales window and declined in the following periods (Table 1).[20] About 45% of the balanced panel (N=832) purchased IBLI coverage at least once during the four sales periods covered in these data, a relatively high rate of uptake when compared against other index insurance pilots in the developing world. Conditional on purchasing an IBLI policy, the mean coverage purchased among the same sample was 3.15 TLUs or 24% of the average herd size during the sales windows. Table 2 details the frequencies of observed transitions between purchased coverage, existing coverage, and lapsed

[20] It is important to note that IBLI was not available for purchase during the short rain/short dry (SRSD) 2010 or long rain/long dry (LRLD) 2012 seasons due to logistical failures in the commercial supply channel.

coverage. Figure 2 illustrates the proportion of the sample that purchased IBLI during each sales window and the level of purchase, conditional on purchasing.

Although existing research, which we discuss in detail below, has already provided a framework by which to understand many of the factors of demand, we are in the unique position to empirically examine the role of basis risk and spatiotemporal adverse selection. Both are thought to impact demand but have not yet been tested using observations of household losses. At the same time, we reinforce previous findings in the literature by including factors that have been found to influence demand elsewhere. This section discusses the key variables used in the analysis.

Basis Risk

Low uptake is often thought to be due to basis risk, although no studies to date have had a direct measure of basis risk with which to test that hypothesis. Here it becomes useful to decompose basis risk into its design and idiosyncratic components. Design risk arises due to differences between predicted and actual division-average livestock mortality while idiosyncratic risk is due to differences between the covariate and individual losses.[21]

One might think of design risk as an indicator of contract adherence, so far as it is the result of a deviation between the intended and actual coverage provided by a policy. But, households are unlikely to have information about the accuracy (or inaccuracy) of an index before product introduction. In cases where index products are new, such as in the Marsabit IBLI pilot we study, individuals must learn about design risk as index performance is revealed through observations of published index values. Karlan et al. (2014) find evidence that households change their demand for index insurance as they observe new index values, following a demand pattern consistent with learning about the product.

The difference between the index and covariate losses during seasons that IBLI coverage was available and index values were publicized are used to generate our estimates of perceived or observed design risk. These estimates are a lagged moving average of within-division design error during preceding seasons in which IBLI coverage was available. We assume households expect no design error in the first sales round, which is reasonable in this context considering that extension and education focused on the likelihood of idiosyncratic risk but did not discuss design risk at all. After the first round, households discard their initial naive expectation and update so that their posterior is the average observed design error. They continue to do so in each of the following rounds. Table 3 includes the observed design error estimates as well as the seasons used to make each estimate.

Price surely matters to insurance uptake (Cole et al. 2013, Giné, Townsend & Vickery 2008, Karlan et al. 2014). The effective premium rate is calculated as the natural log of the premium rate after accounting for randomly distributed discount coupons. The effective premium rate is also interacted with observed

[21] We did not distinguish between design and idiosyncratic risk in Section 3 because it is their combined effect that determines the level of risk that an insured individual retains. Because design risk can be corrected through index modification while idiosyncratic risk cannot, this decomposition is useful.

design error to test the hypothesis that the price elasticity of demand changes with basis risk and to determine the sign of that change.

Although households initially have very little information on index accuracy, they are likely to already be quite familiar with their own historical losses and how those losses relate to the average losses within their division — i.e., their risk and idiosyncratic risk. Households that systematically face high losses that are unrelated to covariate losses are less likely to benefit from even an accurate (i.e., no design error) index product. The variance in livestock mortality rate is a measure of the insurable risk that a households faces. The correlation between individual and covariate losses provides a measure of the how well covariate risk matches household risk, providing an indication of the amount of coverage that an index insurance product with zero basis error could provide. A household with a correlation of one could be fully covered by an area average loss index insurance product like IBLI. As correlations fall from one, idiosyncratic risk increases and index insurable risk falls.

The risk (variance in loss rate) and correlation between an individual's loss rate and their division's average loss rate is estimated using all eight observed seasons to provide an indicator of the idiosyncratic risk that each household is exposed to. Figure 3 provides histograms of the estimated correlation between individual losses and covariate losses in each division. There is clearly a great deal of variation within and between divisions in the individual-covariate loss correlation. Indeed, 15.4% of households have a non-positive correlation, implying that even if IBLI suffered from zero design risk, it would be risk-increasing for them despite its insurance label.

In order to accurately incorporate knowledge of idiosyncratic risk into their purchase decision, households must also understand that the IBLI contract is meant to insure only covariate risk. Without that understanding, households might not link purchases with their level of idiosyncratic risk. Ideally an estimate of idiosyncratic risk could be interacted with household understanding of IBLI. Although the IBLI survey does include a simple test of accuracy of IBLI knowledge, that evaluation could not be collected before the first sales period and is likely endogenous to the decision to purchase an IBLI policy. In addition, the survey cycle and IBLI sales windows do not coincide so as to provide unique data for each household in each survey window.

As a proxy for IBLI knowledge, we include a dummy for participation in the randomized education game described in the research design section. Participation in the game had a strongly positive and significant impact on performance on the IBLI knowledge test (Table 4). There is some prospect that game participation leads to purchasing through a mechanism other than knowledge (e.g., trust, a sense of obligation) so that the above test reported in Table 4 captures an increase in knowledge due to purchase rather than due to the educational component of the game. This can be tested by restricting the analysis to only those households who never purchase IBLI. As reflected in the second row of Table 4, among those who never purchase IBLI, participation in the game increased average IBLI knowledge test scores by nearly 36% (p-value<0.01), providing strong evidence that randomized participation in the extension game directly leads to greater IBLI knowledge. The indicator variable for exogenous game participation is therefore interacted with the idiosyncratic risk estimate in order to test the hypothesis that greater understanding of the IBLI contracts impacts consumer response to basis risk.

Spatiotemporal Adverse Selection

IBLI specifically is susceptible to intertemporal adverse selection because droughts leading to high livestock mortality are often the result of multiple seasons with poor precipitation so that households may wait until conditions are very poor before purchasing insurance. We include two variables—*Pre-Czndvi* and the household's expectation of rangeland conditions in the coming season—to capture ecological conditions that pastoralists may observe while making their purchase decision

Pre-Czndvi is a variable used in the IBLI response function to control for conditions at the beginning of the season and is calculated by summing of standardized NDVI values from the beginning of the previous rainy season until the current sales period. Higher *Pre-Czndvi* indicate greater relative greenness during the rainy season leading up to the current insurance season. Although the index takes *Pre-Czndvi* into account when estimating livestock mortality and premiums could be adjusted to reflect the level of risk at the beginning of a season, the insurer and reinsurer have chosen not to vary premium rates to account for this observed intertemporal variation in livestock mortality risk. *Pre-Czndvi* has a statistically significant and negative relationship with predicted livestock mortality rates (column 1, Table 5). Thus, if households observe the relative greenness that is captured by *Pre-Czndvi*, they could use those observations to help predict coming index values and adjust their purchase accordingly.

A set of dummy variables specify if the household's stated expectations for the coming season's rangeland are good, normal, or bad. Expectation of good or normal rangeland conditions are negatively and statistically significantly correlated with end-of-season index values (predicted livestock mortality rates) as is expected if they correctly predicted coming rangeland conditions (column 2, Table 5). Our model predicts that as long as premium rates are below the expected indemnity rate, households expecting higher livestock mortality rates will increase purchases but is ambiguous about the impact of that expectation if it also suggests higher index values.[22]

Households' expectations of rangeland conditions may have information that is captured by the *Pre-Czndvi* variable, which the IBLI providers could account for by adjusting premium rates to match their risk exposure, or the households may be able to observe additional information that is not captured by the remotely observe NDVI. Regressing predicted livestock mortality onto both *Pre-Czndvi* and households' expectations of coming conditions provides strong evidence that the households have information that is not captured by *Pre-Czndvi*. The implication is the although IBLI providers could reduce the potential for intertemporal adverse selection associated with initial rangeland conditions by adjusting premium rates according to *Pre-Czndvi,* they would continue to face risk of adverse selection from accurate private information held by their potential consumers.

[22] The effective seasonal subsidies (E[indemnity payment rate]-seasonal premium rate) are as follows: Central/Gadamoji 0.0249, Laisamis 0.0171, Loiyangalgani 0.0148, and Maikona 0.017

We also test for spatially defined adverse selection, which could emerge due to variation in the subsidy/loading rate in policies or variation in the quality of the policies. Variation in subsidy/loading rate is the result of the aggregation of index divisions into larger premium regions so that lower risk divisions are inadvertently subsidizing the premium rates of higher risk division in the same premium region. Division-average livestock mortality rate and risk (variance in livestock mortality rate) are used to capture division-level differences in risk, and thus actually fair premium rate of a perfect index product. Division average idiosyncratic risk (correlation between livestock mortality rate and covariate livestock mortality rate) provides an estimate of the average levels of basis risk and its importance relative to total risk within each division.

Additional Key Variables

Within the standard model of insurance, exposure to risk coupled with risk aversion is the fundamental reason for insurance demand. At any level of positive exposure to risk, the benefits of indemnified losses increase with level of risk aversion. But the impact of risk aversion on demand is somewhat ambiguous when market imperfections, such as basis risk or premium loadings, enter the picture. Clarke (2011) shows that for individuals with constant absolute or relative risk aversion, demand for insurance with actuarially unfavorable (favorable) premiums should increase (decrease) and then decrease (increase) as risk aversion increases. Most empirical studies of index insurance demand assume a monotonic relationship between risk aversion and demand, often finding that increased risk aversion is associated with decreased demand (i.e., Giné, Townsend & Vickery 2008; Cole et al. 2013). This negative correlation between risk aversion and demand for insurance has been interpreted as evidence that index insurance uptake in developing countries is more similar to technology experimentation/adoption than to neoclassical models of insurance demand. Hill, Robles, and Ceballos (2013) specifically test for hump-shaped demand across risk aversion as predicted by Clarke (2011), but find no significant difference in demand across the domain of observed risk aversion. In a setup similar to that used by Hill, Robles, and Ceballos, we allow for a non-linear relationship between risk aversion and demand.

Whether households place more importance on absolute or relative risk is an empirical question that has not yet been addressed in the context of index insurance. To determine which is more important, we include total herd size and ratio of income generated from livestock and livestock related activities. Total herd size provides an absolute measure of exposure to asset risk associated with IBLI insurable assets, while the ratio of income that is generated from livestock and livestock related activities approximates the relative income risk associated with livestock mortality.

Theory and empirical evidence are also ambiguous as to how wealth should affect demand for insurance when prices are actuarially unfavorable. Clarke (2011) shows that the relationship between wealth and demand is not monotonic for most reasonable utility functions in such environments. Empirical studies offer contradictory evidence, finding that demand increases (e.g., Cole et al. 2013; Mobarak & Rosenzweig 2012) or decreases (e.g., McIntosh, Sarris, & Papadopoulos 2013) in variables associated with wealth. The literature on poverty traps, as has been demonstrated by multiple studies of east African pastoralists (Lybbert et al. 2004, Barrett et al. 2006, Santos and Barrett 2011), indicates that demand may be non-linear in wealth, changing dramatically across certain asset thresholds as households try to avoid or to break free of a low asset dynamic equilibrium (Chantarat et al. 2009; Janzen, Carter & Ikegami 2012; Lybbert, Just, & Barrett 2013). We summarize household wealth with an asset index generated through factor analysis of an extensive list of household construction materials, productive assets excluding livestock, and other durables (Appendix C).

Lack of liquidity is often found to constrain demand. Mobarak and Rosenzweig (2012) found that lack of cash was the primary reason given by Indian farmers for not purchasing an available index insurance product. Although liquidity is likely correlated with wealth, it can constrain demand at any wealth level (Cole et al. 2013). In order to capture liquidity, we calculate the sum of cash savings on hand or placed within any of several formal and informal savings arrangements. A household's savings are liquid and provide a lower band estimate of access to liquid capital. Descriptive statistics show that those households with savings have substantial amounts but that most households (74%) have no savings. Rather than use the continuous but highly skewed estimation of total savings, we use a dummy variable that is equal to one if the household has savings sufficient to purchase IBLI insurance for ten TLUs (≥8,250Ksh or ≥4,875 in the upper and lower contract regions, respectively). We also include an estimate of monthly income composed of earnings and the value of in-kind production.

The Hunger Safety Net Program (HSNP), an unconditional cash transfer program, launched in the Marsabit region in 2009, just before IBLI began. HSNP provides transfers every two months to eligible households for at least two years. The size of bi-monthly household transfers increased from 2,150Ksh in 2009 (about USD25) to 3,000Ksh in 2011 and then increased again in 2012 to 3,500Ksh in order to help households cope with a severe drought. 3,500Ksh could have purchased insurance for about 7 cattle in the lower Marsabit region at that time. There was no retargeting of or graduation from HSNP, which could have led to perverse incentives not to purchase IBLI if insurance has a beneficial impact on wealth. Although HSNP participation was not random within communities, we are able to cleanly identify the impact of transfers on demand by controlling for the known and corroborated household selection criteria and HSNP community selection.[23]

Access to informal insurance schemes can be an important factor in demand for formal insurance. Mobarak and Rosenzweig (2012) show that informal risk pools that insure against idiosyncratic shocks complement index insurance with basis risk while informal schemes that protect against covariate shocks act as a substitute. In the pastoral societies of northern Kenya, informal risk sharing through livestock transfers and informal credit appears to be modest at best (Lybbert et al. 2004; Santos & Barrett 2011). Although there does seem to be a relationship between providing and receiving livestock transfers among Kenyan herders, those transfers are not timed so as to reduce the impact of shocks or to protect assets (McPeak 2006). But, because informal risk sharing is extremely relevant to this work and has empirically been found to impact demand for index insurance in India (Mobarak & Rosenzweig

[23] For more details on the HSNP program logistics go to http://www.hsnp.or.ke/ while analysis of impacts can be found in Hurrell & Sabates-Wheeler (2013) and Jensen, Mude and Barrett (2014).

2012), we include the number of informal groups that the household participates in as a coarse indicator of potential access to risk pooling.[24]

Table 6 describes the temporal nature of the data used to construct each variable, how each variable is constructed and which are lagged to avoid capturing changes due to paying the premium or due to behavior responses to having IBLI coverage. Table 7 provides summary statistics, distinguishing between those households who never purchased IBLI over the four sales windows and those who purchased at least once. Differences in unconditional means between the two groups show that IBLI purchasers have lower dependency ratios, face somewhat less livestock mortality risk, are less likely to be extremely risk averse in favor of moderately risk averse, and more likely to have received a discount coupon in at least one of the sales windows. But, the two groups seem to be mostly similar as the discount coupons directly impact effective price and indirectly IBLI coverage via price.

Finally, the IBLI contracts provide coverage for 12 months following the sales window in which they were purchased. If there had been sales windows before each semi-annual rainy season, it would be common for households to enter sales windows with existing coverage for the following season from the preceding season. They would then be making a decision whether to increase coverage by purchasing more or to maintain current coverage by purchasing nothing this period. Logistical problems faced by the insurer did not allow for consistent sales twice a year, but the survey does capture two consecutive sales seasons during which IBLI policies were sold. We expect that existing coverage still in force could impact purchase decisions and so control for existing coverage in that period.[25]

6. Econometric strategy

We seek to identify the factors that influence demand for IBLI. Insurance demand is best modeled as a two stage selection process. Propensity to purchase is first determined as the household decides whether or not to buy IBLI. Those households who choose to purchase then decide how much to buy. Let h_{it}^* and y_{it}^* be latent variables that describe the categorical desire to purchase insurance and the continuous, optimal level of purchase, respectively. If $h_{it}^* > 0$ we observe the positive level of purchase $y_{it} = y_{it}^*$, and if $h_{it}^* \leq 0$, we observe $y_{it} = 0$. We write the process as a function of time invariant individual characteristics (c_i, d_i) including a constant term, time varying individual and division characteristics (x_{it}, z_{it}), and error terms (u_{it}, v_{it}) as follows.

(5)
$$y_{it}^* = c_i'\eta + x_{it}'\beta + u_{it}$$
$$h_{it}^* = d_i'\eta + z_{it}'\gamma + v_{it}$$

[24] Although ethnic group is also likely to be important in determining access to informal insurance, collinearity between ethnicity and location makes that aspect difficult to examine while also examining other variables that are correlated with location, such as the expected subsidy level and HSNP participation.

[25] We use a dummy variable to indicate existing coverage. If households with existing coverage reduced purchases due to their existing coverage, a continuous variable might be more appropriate. That does not seem to be the case. Households with existing coverage are much more likely to purchase additional insurance than those without it (difference = 13.6%, t-statistic=4.265) but existing coverage does not impact level of purchase conditional on purchasing (difference = 0.22, t-statistic=0.387).

$$y_{it} = \begin{Bmatrix} 0 & if\ h_{it}^* \leq 0 \\ c_i'\eta + x_{it}'\beta + u_{it} & if\ h_{it}^* > 0 \end{Bmatrix}$$

If the same process is used to determine the desire to purchase insurance and the level of purchase, then $y_{it}^* \equiv h_{it}^*$ and the model reduces to Tobin's (1958) model for censored data. In the case of IBLI (and for many other cases) there is reason to believe that the two processes may differ. For example, the probability of purchasing any IBLI coverage is likely correlated with the distance that the purchaser must travel to make the purchase. There is little reason to think that the same distance variable would affect the level of purchase. If demand is a two stage process but the two decisions are independent (conditional on observed covariates), each stage can be estimated separately and consistently using a double hurdle model (Cragg 1971).

In this context, the two decisions most likely fall somewhere between Tobin's assumption that they are identical and Cragg's assumption that they are independent. That is, u_{it} and $v_{it}$are not identical but they are correlated so that both the single model and independent models result in biased estimates of β. Heckman (1979) suggests that such bias is due to a missing variable that accounts for selection. To control for selection, Heckman proposed including the ratio of the predicted likelihood of selection to the cumulative probability of selection (the inverse Mills ratio). The inverse Mills ratio is estimated by first using a probit model to estimate $\Pr(s_{it} = 1 | d_{i,}z_{it}) = \Phi(d_{i,}z_{it}, \eta, \gamma)$, where $s_{it} = \begin{Bmatrix} 0 & if\ h_{it}^* \leq 0 \\ 1 & if\ h_{it}^* > 0 \end{Bmatrix}$. The estimates are then used to calculate the inverse Mills ratio $\hat{\lambda}_{it} = \frac{\phi(d_{i,}z_{it}, \hat{\eta}, \hat{\gamma})}{\Phi(d_{i,}z_{it}, \hat{\eta}, \hat{\gamma})}$, where $\phi(d_{i,}z_{it}, \hat{\eta}, \hat{\gamma})$ is the normal density.

Accounting for unobserved household level fixed effects is then a matter of applying panel data estimation methods to Heckman's framework. If intercepts among observations are common ($c_i = c, d_i = d$), pooled models suffice. Random effects probit/linear models are consistent as long as the intercepts are uncorrelated with the covariates (x_{ti}, z_{ti}). But, if the individual fixed effects are correlated with explanatory variables then the random effects estimates are inconsistent. For short panels, the standard fixed effects approaches suffer from the incidental parameters problem when applied to probit models.[26] But, if the data generating process is best described by the fixed effects model, pooled and random effects models will also be biased. Greene (2004) compares the magnitude of the bias introduced by estimating pooled, random effects, and fixed effects probit parameters for data generated by a probit process with fixed effects. At T=3 and T=5, Greene finds the random effects estimates are the most biased, and that the bias associated with the pooled and fixed effects models are similar in magnitude. In addition, standard errors are likely to be underestimated in the fixed effect model. We include pooled estimates in this analysis, acknowledging their likely bias but appealing to Greene's (2004) result that these are likely least bad estimates.

[26] Because the probit model is non-linear the parameters must be estimated using within household observations, of which we have a maximum of four.

As an alternative, we also follow a procedure developed by Wooldridge (1995), which builds off of earlier work by Mundlak (1978) and Chamberlain (1980), to allow for correlation between the fixed effects and a subset of within-household mean characteristics ($\bar{x}_i^{FE}$) but assume independence conditional on the mean. In addition the errors are assumed to be distributed normally.

(6)
$$c_i = \bar{x}_i^{FE\prime}\gamma_1 + e_{it}^c, e_{it}^c|\bar{x}_i^{FE} \sim N(0, \sigma_e^2)$$

$$d_i = \bar{x}_i^{FE\prime}\delta_1 + e_{it}^d, e_{it}^d|\bar{x}_i^{FE} \sim N(0, \sigma_e^2)$$

$$\bar{x}_i^{FE} = \frac{1}{T}\sum_T x_{it}^{FE}, x_{it}^{FE} \subseteq x_{it}, z_{it}$$

As with the Heckman selection process described above, a probit model is used estimate the inverse Mills ratio, but in this case the estimate is a function on household average characteristics and period specific characteristics $\hat{\lambda}_{it} = \frac{\phi(\bar{x}_i^{FE\prime}, z_{it}, \hat{\delta}_1, \hat{\eta}, \hat{\gamma})}{\Phi(\bar{x}_i^{FE\prime}, z_{it}, \hat{\delta}_1, \hat{\eta}, \hat{\gamma})}$. In order to add more flexibility, and thus accuracy, to the first stage estimations, the probit model is estimated separately for each period.

Within household mean characteristics are estimated using all eight seasonal observations while s_{it} and y_{it} are only estimated during the four seasons in which there were sales. For those variables that appear in our estimates twice, as a household mean and a period specific observation, we use the deviation from the mean as the period specific observation to facilitate interpreting the estimates.

We report the pooled and the conditionally independent fixed effects estimates, while relying primarily on the latter as the preferred estimates. If the data generating process does include unobserved individual effects that are correlated with our outcome variables and the covariates, our pooled estimates are likely to be biased but perform better than either random or fixed effects models (Greene 2004). The conditionally independent fixed effects should generate estimates that are at the very least, less biased than those from the pooled model.

Both models are estimated using maximum likelihood. Although effective (discounted) price is included in both selection and demand equations, a dummy variable indicating that the household randomly received a discount coupon is included in the selection equation but is excluded from the demand equation. The discount coupon serves merely as a reminder of the product availability and thus should affect the dichotomous purchase decision but have no effect on the continuous choice of insurance coverage conditional on purchase once we control for the effective discounted price. Although there is no agreed upon exclusion test for selection models, we perform two exploratory tests that support the exclusionary restriction on the coupon dummy variable in the demand equation, as reported in Appendix D.

7. Results and Discussion

Wooldridge (1995) describes a test for selection that assumes conditionally independent fixed effects in the selection stage but relaxes the conditional assumption in the outcome stage. That test does not reject the null hypothesis of an independent second stage at the standard 10% level of statistical significance (*F-stat*=2.06, *p-value*=0.1522), but is near enough to indicate that caution is advised. Thus, we proceed as though demand for IBLI can only be understood by first examining the factors that determine who purchases IBLI and then what drives the levels of purchases conditional on purchasing.[27] In the following discussion we focus on the estimates generated from the conditional fixed effects model while also reporting the pooled estimates. The average marginal effects (AME) estimates are provided in tables 8 and 10 while the regression coefficient estimates can be found in Appendix E.[28]

Determinants of IBLI uptake

The relationship between wealth, access to liquidity, investments in livestock, and uptake are predictably complicated (Table 8). Herd size and HSNP transfers are positively related to IBLI purchase while asset wealth is negatively related to purchases. Although these estimates may seem superficially contradictory, in the context of a new technology in a pastoral region they strike us as intuitive. Households with larger herds have the greater potential absolute gains from the IBLI product. Large herds also require mobility to maintain access to forage (Lybbert et al. 2004) and many of the larger assets included in the asset index (e.g., TV, tractor, plow) are likely to be less appropriate for mobile, livestock-dependent households for whom IBLI should be most valuable.

There is weak evidence of intertemporal adverse selection and strong evidence of spatial adverse selection. Households in divisions with greater average livestock mortality rate, lower variation in that rate (risk), and less idiosyncratic risk (as captured by greater average correlation between losses and the index) are more likely to purchase IBLI. The negative relationship between idiosyncratic risk and uptake is consistent with the predictions made by our analytic model. The fact that greater variation in livestock losses is associated with reduced uptake requires a closer look at the data. One likely explanation is that there is greater idiosyncratic risk (and thus basis risk) in divisions with more variation in losses. We test for a positive correlation between division average variance in livestock mortality rate and division average idiosyncratic risk, finding that the correlation is positive and significant (rho=0.98, p-value=0.004, N=4).

Observed design error has a significant and negative AME on uptake. Although the estimated AME of price is statistically insignificant, the coefficient estimates (Table E.1) show that the interaction between price and observed design error is important. Examining the impact of design error across a range of observed IBLI prices reveals that AME of observed design error is negative and increases in both significance and magnitude as prices increase (Table 9). The same test for price response at various

[27] Analysis of uptake and level of purchase separately provides estimates that are very similar to those described in this paper. Importantly, our findings concerning the importance of basis risk and adverse selection are the same.

[28] The second stage of the conditional fixed effects model is estimated using inverse Mills ratios generated by estimating the first stage probit model separately for each period. In Tables 8 and E.1, we present the average coefficient estimates generated by pooling the four periods, including both time specific and household average characteristics.

levels of observed design error shows that at low levels of design error uptake is does not respond strongly to prices, while at higher levels of design error price plays a much more significant role in determining uptake. When observed design error is one standard deviation above the mean, the average effect of a one unit increase in prices is to reduce uptake by 7.9% (AME=-0.079, t-statistic=-1.68).

Households with consistently high participation in social groups have a greater propensity to purchase IBLI (Table 8). Although participating in social groups could be endogenous to purchasing IBLI, we find that lagged participation in the pooled model (column 1, Table 8) and household's average participation (including 3 seasons before the first sales season., column 3, Table 8) has a positive and significant impact on uptake. Plausible explanations for the positive relationship between social group participation and IBLI uptake include the complementarities between index insurance and informal idiosyncratic risk pooling described by Mobarak and Rosenzweig (2012) and learning through social networks (Cai, de Janvry & Sadoulet 2011).

Randomized exposure to the IBLI educational game allows us to look more closely at the impact of learning. Here we see that increased IBLI knowledge associated with participating in the game has no discernible impact on the decision to purchase IBLI (Table 9), although we know it does have a strong impact on understanding of the IBLI product (Table 4). In that case, it seems less likely that the pathway by which participation in social groups impacts demand is through increased understanding of the product and the argument that social group linkages stimulate IBLI uptake due to complementarities with informal insurance is stronger.

The discount coupon, which is excluded in the second stage, has an AME of +17% on the likelihood of purchasing insurance and is statistically significant at the one percent level. Quite apart from the price effect of the discount coupon, it seems to serve a useful role as a visible reminder to households of the availability of insurance.

Quantity of Insurance Purchased

The continuous IBLI purchase decision reveals some of the same patterns evident in the decision to purchase (Table 10). Larger herds are again associated with increased demand.[29] But, among those purchasing, demand increases with greater asset wealth, greater income, and income diversification into non-livestock related activities (nearly all of which is generated earnings). Jointly, these results provide strong evidence that demand is liquidity constrained among those seeking to purchase IBLI.[30] Referring back to our model of household demand for insurance, we could not analytically sign many of the relationships between household financial characteristics and demand because of the ambiguity of

[29] The AME of herd size is positive but less than one, revealing that households with larger herds do insure more animals but they are insuring a smaller portion of their total herd.

[30] All household income was derived from livestock in about 53% of the household observations during sales season. During the same periods, 47% of the households that purchased insurance generated all of their income from livestock in the period that they purchased. Non-livestock income sources captured in the survey are from sale of crops, salaried employment, pensions, casual labor, business, petty trading, gifts, and remittances.

the wealth effect on demand. Empirically we also find mixed responses, such as asset wealth reducing the likelihood of uptake but increasing coverage levels conditional on uptake, while livestock wealth is associated with increases in both uptake and conditional coverage levels.

There is evidence of both inter-temporal and spatial adverse selection in IBLI purchases conditional on positive demand. For households that purchase insurance, the AME of expecting good rangeland conditions represents an 11.2% reduction in coverage from the mean coverage purchased.[31] The coefficient estimate for *Pre-Czndvi* (a division level proxy for rangeland conditions at the time of sale) is also negative and statistically significant. Division level risk has a positive impact on level of purchase so that households in divisions with high average risk are less likely to purchase but buy more coverage, conditional on purchasing. In addition, those divisions with higher average livestock mortality rates are more likely to purchase IBLI, but purchase less coverage.

The correlation between individual and covariate losses plays a role in determining level of demand, although its impact is somewhat obscured by interactions (Table E.2). Separating purchasers by game play, the estimated AME of the correlation between an individual's losses and the covariate losses of their division is negative and significant for households who did not participate in the IBLI extension game (Table 11). Although this does not confirm our hypothesis on the interaction between understanding the IBLI product and the impact of basis risk on demand, it does point to a grave misunderstanding of the product among those that did not received product education via the extension game. As discussed in Section 4, participation in the IBLI game was randomized and has a large and significant impact on understanding of the IBLI product (Table 4). Here we see that purchase levels among those with less understanding of the product are higher among those with less covariate (insurable) risk.[32]

Price is a significant factor influencing demand conditional on uptake, but demand is rather price inelastic, with an AME -0.43, lower than any of the other estimates we find in the literature. Examining the impact of observed design error on the price elasticity of demand, we find that the elasticity of demand and statistical significance of premium rates increases at higher levels of observed design error (Table 11). But, there is no direct negative effect of design error on level of purchase even at high premium levels. Jensen, Barrett and Mude (2014) shed some light on why households may not have responded to design risk directly; in most cases design risk is minor when compared to idiosyncratic risk. Hence our findings that demand is much more closely linked with indicators of adverse selection make perfect sense.

Concluding Remarks

[31] The AME of expecting good rangeland conditions is -0.2709 while the averse coverage purchased is 2.429 TLUs.

[32] Household level risk is accounted for in the risk variable so that this effect is not due to level of covariate risk picking up the effects of total risk. In addition, very few households ever purchase coverage for more animals than they hold so that this is unlikely to be the result of households (mistakenly) over-insuring to make up for uninsured idiosyncratic risk.

The above analysis provides strong empirical evidence that in addition to price and household demographic characteristics, adverse selection and basis risk play economically and statistically significant roles in determining demand. The point estimates from our analysis (Table E1 and E2) predict the changes in IBLI purchases over time rather well, showing a reduction in uptake after the first period and a small upturn in the final period (Figure 4 and Figure 5).

A Shapley's R^2 decomposition can shed some light on which factors contribute most to explaining variation in IBLI uptake and level of purchase. After grouping the covariates into several categories, we re-estimate the uptake and demand equations separately and decompose their goodness of fit measures using the user-written command *shapely2* (Juárez 2014), which builds off earlier work by Kolenikov (2000) and theory by Shapley (1953) and Shorrocks (2013).[33] The Shapley R^2 decompositions reported in Appendix F should be interpreted as the ratio of the model's goodness of fit (R^2 or Pseudo R^2) that can be attributed to the groups of variables. For both uptake and level of demand, the total role of adverse selection and product related variables in explaining demand is similar but larger than that of household characteristics (demographics and financial), providing strong evidence that product design and the nature of the insured risk are at least as important as household demographic characteristics. The Shapley values indicate that the three variables associated with design risk and price are responsible for 21% of our goodness of fit measure for the uptake model, a considerable achievement considering that there are more than 25 other covariates and that the discount coupon accounts for 35% of the model's fit. The role of design risk and price falls by about 5 points when examining level of purchase, where spatial and temporal adverse selection become increasingly important. Together the two groups of adverse selection variables account for 32% of the models goodness of fit for level of purchase. The importance of idiosyncratic risk to the fit of the model is fairly low and consistent in both uptake (5.46%) and level of purchase (5.42%).

With the model estimates and Shapely values in mind, it is clear that both product and household characteristics play an important role in determining demand for index insurance. Insurance products can do little to change household characteristics; but it may be possible to lessen adverse selection and idiosyncratic risk through improved contract design. For example, IBLI no longer aggregates index divisions into premium regions, removing one source of adverse selection. Adjusting premium rates dynamically to account for initial season conditions is an additional step that could be taken to reduce adverse selection. Idiosyncratic risk limits the potential impact of even a perfect index product, but is in part a construct of the index division, which could be adjusted to increase the importance of covariate risk. And finally, reducing design risk is likely to be relatively simple if household-level data is collected and used to improve the performance of the index.

[33] The variable categories are demographic, financial, intertemporal adverse selection, spatial adverse selection, idiosyncratic risk and knowledge, design risk and price, other, and the instrument variable.

References

Barnett, B. J., & Mahul, O. (2007). Weather index insurance for agriculture and rural areas in lower-income countries. *American Journal of Agricultural Economics*, *89*(5), 1241-1247.

Barnett, B.J., Barrett, C.B., & Skees, J.R. (2008). Poverty Traps and Index-Based Risk Transfer Products. *World Development*, 36(10), 1766-1785.

Barrett, C., Barnett, B. J., Carter, M., Chantarat, S., Hansen, J., Mude, A, Osgood, D., Skees, J., Turvey, C., & Ward, M. (2008). Poverty Traps and Climate Risk: Limitations and Opportunities of Index-Based Risk Financing, IRI Technical Report 07-02.

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

Barrett, C., Barnett, B. J., Carter, M., Chantarat, S., Hansen, J., Mude, A., Osgood, D., Skees, J., Turvey, C., & Ward, M. N. (2008). Poverty traps and climate risk: limitations and opportunities of index-based risk financing.

Binswanger-Mkhize, H. P. (2012). Is there too much hype about index-based agricultural insurance? *Journal of Development Studies*, *48*(2), 187-200.

Boucher, S. R., Carter, M. R., & Guirkinger, C. (2008). Risk rationing and wealth effects in credit markets: Theory and implications for agricultural development. *American Journal of Agricultural Economics*, *90*(2), 409-423.

Cai, J., de Janvry, A., & Sadoulet, E. (2011). Social networks and insurance take up: Evidence from a randomized experiment in china. *ILO Microinsurance Innovation Facility Research Paper*, (8).

Carriquiry, M. & Osgood, D. (2012). Index insurance, probabilistic climate forecasts, and production. *Journal of Risk and Insurance* 79(1), 287-300.

Chamberlain, G. 1980. Analysis with qualitative data. *Review of Economic Studies*, 47, 225-238.

Chantarat, S., Mude, A. G., Barrett, C. B., & Carter, M. R. (2013). Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya. *Journal of Risk and Insurance*, 80(1), 205-237.

Chantarat, S., Barrett, C. B., Mude, A. G., & Turvey, C. G. (2007). Using weather index insurance to improve drought response for famine prevention. *American Journal of Agricultural Economics*, *89*(5), 1262-1268.

Chantarat, S., Mude, A., Barrett, C., & Turvey, C. (2009). The performance of index based livestock insurance: ex ante assessment in the presence of a poverty trap. *Available at SSRN 1844751*.

Clarke, D. J. (2011). A Theory of Rational Demand for Index Insurance. Department of Economics Discussion Paper Series, Oxford University, No. 572.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R., & Vickery, J. (2013). Barriers to Household Risk Management: Evidence from India. American Economic Journal: *Applied Economics*, 5(1), 104-135.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica: Journal of the Econometric Society*, 829-844.

Dercon, S., Hill, R., Clarke, D., Outes-Leon, I., & Taffesse, A. (2014). Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia. *Journal of Development Economics,* 106(1), 132-143.

Food and Agriculture Organization (FAO). 1999. Livestock and Environment Toolbox. Found at http://www.fao.org/ag/againfo/programmes/en/lead/toolbox/Index.htm. Accessed on May 2014.

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

Giné, X., & Yang, D. (2009). Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics*, *89*(1), 1-11.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal*, *7*(1), 98-119.

Hellmuth M.E., Osgood D.E., Hess U., Moorhead A., &Bhojwani H. (eds) (2009). *Index insurance and climate risk: Prospects for development and disaster management.* Climate and Society No. 2. International Research Institute for Climate and Society (IRI), Columbia University, New York, USA.

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

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.

Hess, U., Skees, J. R., Stoppa, A., Barnett, B. J., & Nash, J. (2005). Managing agricultural production risk: Innovations in developing countries. *Agriculture and Rural Development (ARD) Department Report*, (32727-GLB).

Hill, R. V., Robles, M., & Ceballos, F. (2013). *Demand for Weather Hedges in India: An Empirical Exploration of Theoretical Predictions* (No. 1280). International Food Policy Research Institute (IFPRI).

Hurrell, A., & Sabates-Wheeler, R. 2013. Kenya Hunger Safety Net Programme Monitoring and Evaluation Component: Quantitative Impact Evaluation Final Report: 2009 to 2012. Oxford Policy Management. Report can be found at http://www.hsnp.or.ke/

International Livestock Research Institute (IRLI). (2012). Index Based Livestock Insurance for northeastern Kenya's arid and semi-arid lands: The Marsabit Pilot Project-Codebook for IBLI evaluation baseline survey. Nairobi, Kenya: ILRI.

Janzen, S. A., Carter, M. R., & Ikegami, M. (2012). Valuing Asset Insurance in the Presence of Poverty Traps. Unpublished. http://www.dartmouth.edu/~neudc2012/docs/paper_270.pdf

Jensen, N., Barrett, C., & Mude, A. (2014). Basis Risk and the Welfare Gains from Index Insurance: Evidence from Northern Kenya. Unpublished.

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

Juárez, F. W. (2014). skapely2. STATA user-written command.

Karlan, D., Osei, R. D., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *Quarterly Journal of Economic*s, in press.

Leblois, A., & Quirion, P. (2010). Agricultural Insurances Based on Meteorological Indices: Realizations, Methods and Research Agenda. *Fondazione Eni Enrico Mattei Working Papers.* Paper 460. http://services.bepress.com/feem/paper460

Liu, Y., & Myers, R. J. (2012). The dynamics of insurance demand under liquidity constraints and insurer default risk. International Food Policy Research Institute (IFPRI).

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

Lybbert, T. J., Just, D. R., & Barrett, C. B. (2013). Estimating risk preferences in the presence of bifurcated wealth dynamics: can we identify static risk aversion amidst dynamic risk responses?. *European Review of Agricultural Economics*, *40*(2), 361-377.

Mahul, O., & Skees, J. (2007). Managing agricultural risk at the country level: The case of index based livestock insurance in Mongolia. *World Bank Policy Research Working Paper*, (4325).

McIntosh, C., Sarris, A., & Papadopoulos, F. (2013). Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia. *Agricultural Economics*, *44*(4-5), 399-417.

McPeak, J. (2006). Confronting the risk of asset loss: What role do livestock transfers in northern Kenya play?. *Journal of Development Economics*, *81*(2), 415-437.

McPeak, J. G., & Barrett, C. B. (2001). Differential risk exposure and stochastic poverty traps among East African pastoralists. *American Journal of Agricultural Economics*, *83*(3), 674-679.

McPeak, J., Chantarat, S., & Mude, A. (2010). Explaining index based livestock insurance to pastoralists. *Agricultural Finance Review*, *70*(3), 333-352.

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., & 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).

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica,* 46, 69-85.

Pratt, A., Suarez, P., & Hess, U. (2010). How do small-holder farmers understand insurance, and how much do they want it? Evidence from Africa. *Global Environmental Change*, *20*(1), 153-161.

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

Santos, P., & Barrett, C. B. (2011). Persistent poverty and informal credit. *Journal of Development Economics*, 96(2), 337-347.

Shapley, L. (1953). A value for n-person games. In: Kuhn, H.W., Tucker, A.W. (eds.) *Contributions to the Theory of Games,* vol. 2. Princeton University Press.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality*, 1-28.

Skees, J. R., & Collier, B. (2008). The potential of weather index insurance for spurring a green revolution in Africa. *Global Ag Risk Inc.*

Skees, J. R., Hartell, J., & Hao, J. (2006). Weather and index-based insurance for developing countries: Experience and possibilities. In A. Sarris, & D. Hallam (Eds.), *Agricultural commodity markets and trade: New approaches to analyzing market structure and instability.* Northampton, MA: Edward Elgar.

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

Kolenikov, S. (2000). shapely. STATA user-written command.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica,* 26(1), 24-36.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics*, *68*(1), 115-132.

Figures

Figure 1. Survey design, participation in IBLI game and HSNP target sites

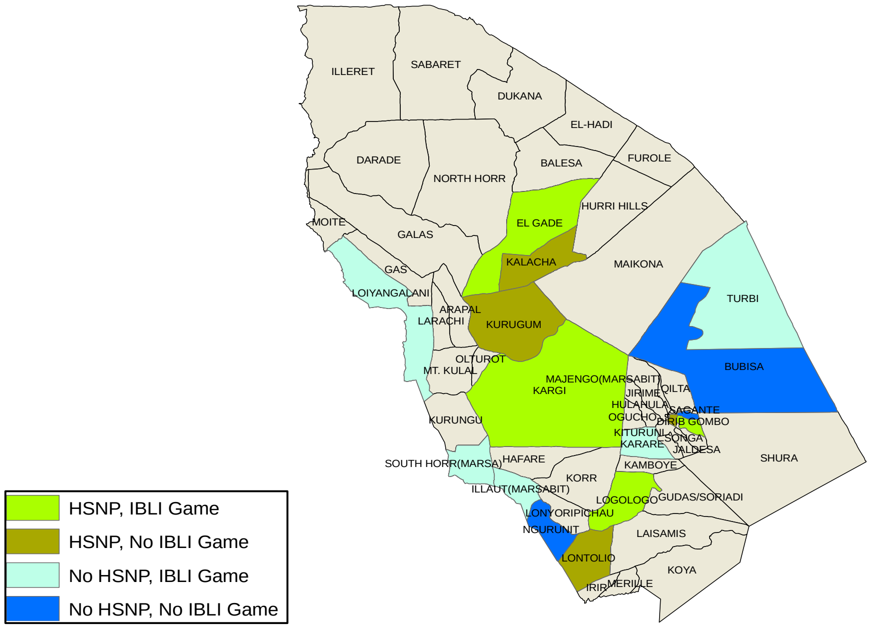


Figure 2. IBLI purchasing behavior during each sales window

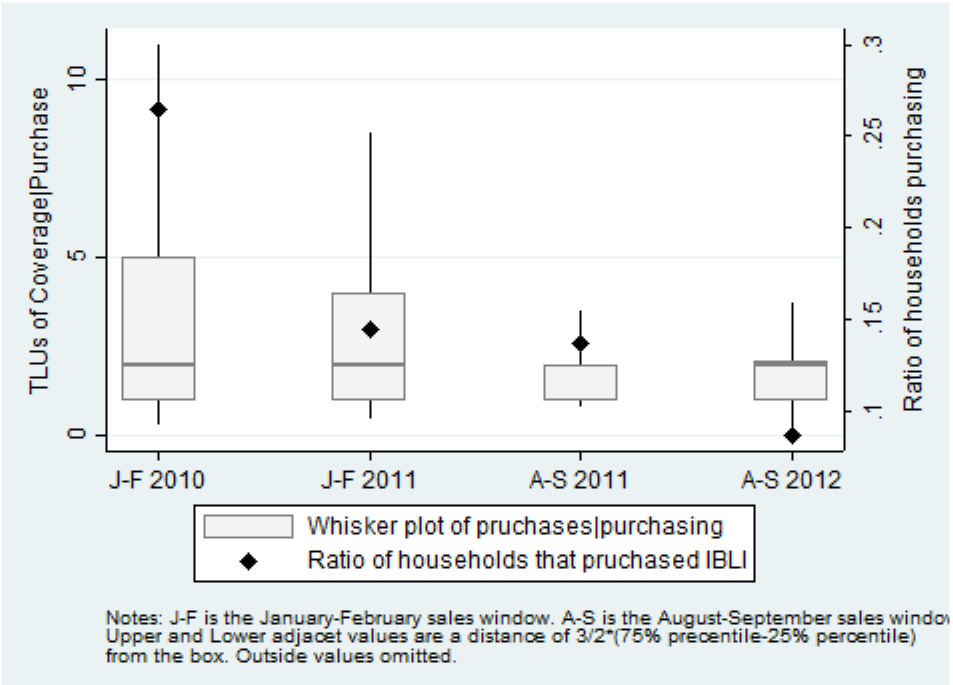


Figure 3. Histograms of the correlation between individual and covariate livestock mortality rates

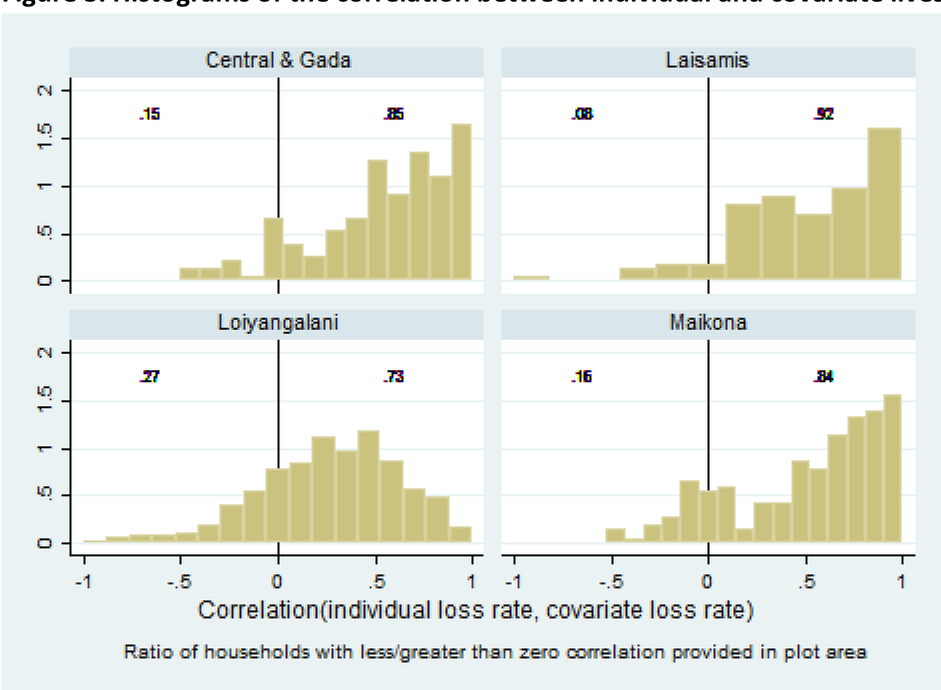


Figure 4. Unconditional observed and predicted (Conditional FE) likelihood of purchasing IBLI

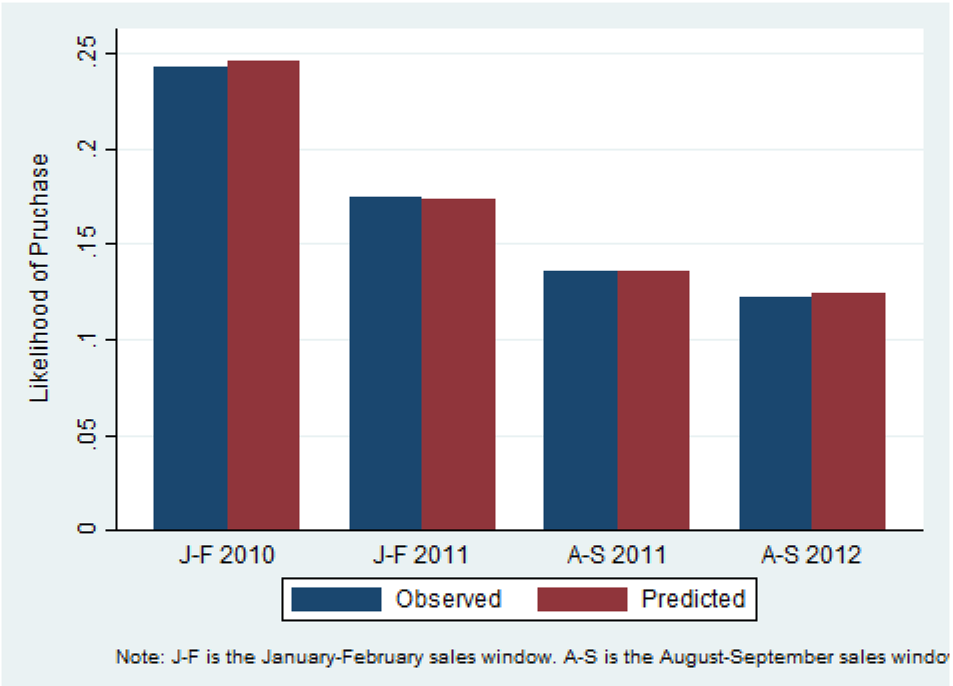
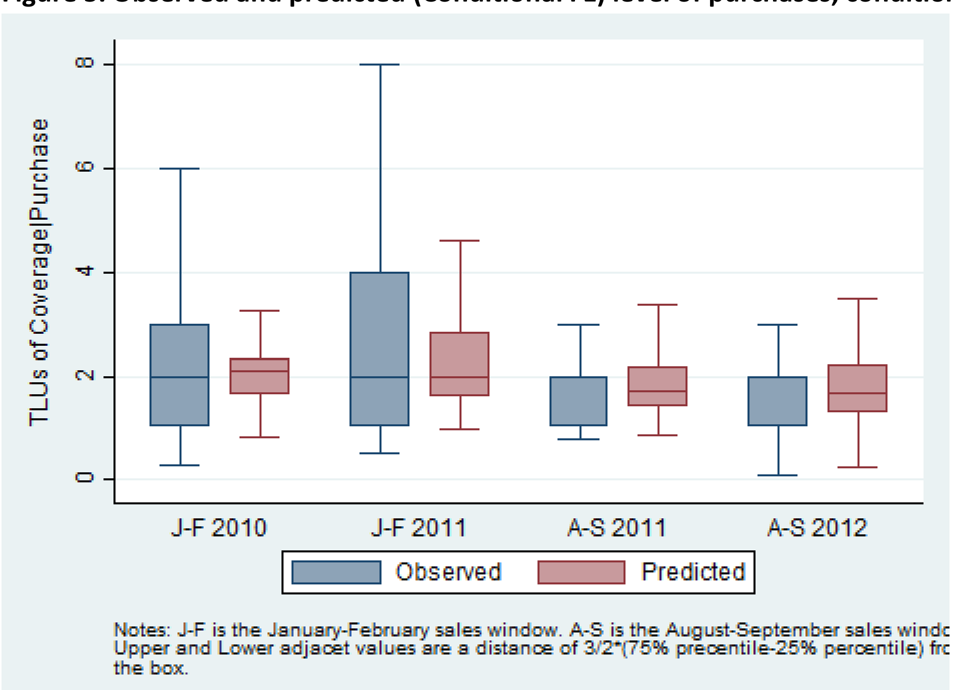


Figure 5. Observed and predicted (Conditional FE) level of purchases, conditional on being a purchaser



Tables

Table 1. Representation of demand for IBLI in the survey sample

Survey[#]	Sales Window	IBLI Coverage Period		Total Contracts Sold	IBLI survey households: Did Not Purchase	IBLI survey households: Purchased
R1	-	-		-	-	-
(2009)	-	-		-	-	-
R2	J-F 2010	LRLD10/SRSD10	(N)	1,974	679	245
(2010)			(Mean) &	3.0	-	(3.94)
	-	-	(N)	-	-	-
			(Mean) &	-	-	-
R3	J-f 2011	LRLD11/SRSD11	(N)	595	790	134
(2011)			(Mean) &	(2.1)	-	(3.05)
	A-S 2011	SRSD11/LRLD12	(N)	509	797	127
			(Mean) &	(1.6)	-	(2.39)
R4	-	-	(N)	-	-	-
(2012)			(Mean) &	-	-	-
	A-S 2012	SRSD12/LRLD13	(N)	216	844	80
			(Mean) &	(1.9)	-	(2.64)

LRLD and SRSD refer to the long rain/long dry and short rain/short dry season respectively. There were no sales during the Aug/Sept 2010 and Jan/Feb 2011 sales periods due to supply channel failures. Jan/Feb 2010, Jan/Feb 2011 & Aug/Sept 2011 were sold under UAP. Aug/Sept 2012 was sold under APA. [#]Surveys were collected during October and November of each year. [&]*Mean* is the unweighted mean coverage purchased in TLUs, conditional on purchasing IBLI.

Table 2. Household IBLI purchase patterns, by sales window

Sales window	New[1]	Replacement[2]	Augmenting[3]	Holding[4]	Reenter[5]	Lapsed[6]	Total[7]
J-F 2010	225	0	0	0	0	0	225
J-F 2011	67	60	0	0	0	165	292
A-S 2011	66	0	31	96	21	144	358
A-S 2012	19	25	0	0	33	300	377

We use the balanced panel of 832 households in this table to track household purchase behavior over time. Therefore, columns do not sum to the totals reported in Table 1. [1]First time purchasers. [2]Replaced a policy about to expire. [3]Purchased additional coverage that overlapped with existing coverage. [4]No purchase but had existing coverage. [5]Let policy lapse for at least one season but purchased this season. [6]Past policies have lapsed and did not purchased additional coverage. [7]Total number of households that have purchased to date.

Table 3. The average observed design error in each division at each sales period

Sales Seasons	Design Risk Observations	Observed Average Estimated Design Error (%)			
		Central/Gadamoji	Laisamis	Loiyangalani	Maikona
J-F 2010	-	0	0	0	0
J-F 2010	LRLD 2010	4.50	11.73	10.22	3.34
A-S 2011	LRLD 2010, SRSD 2010	7.20	11.07	12.90	5.22
A-S 2012	LRLD 2010, SRSD 2010, LRLD 2011, SRSD11	2.07	1.24	7.45	1.91

LRLD and SRSD refer to the long rain/long dry and short rain/short dry season respectively. The observed average estimated design error is the mean difference between covariate loss rate and the predicted loss rate (index) during previous seasons with potential IBLI coverage.

Table 4. The impact of the randomized extension game on understanding of the IBLI contracts

	Not game participant		Game participant			
IBLI Knowledge:	Mean	Std. Err.	Mean	Std. Err.	Difference	t-test
Full Sample	1.72	0.065	2.22	0.086	0.50	4.60***
Never Purchase	1.50	0.085	2.05	0.123	0.54	3.63***

The game was played in January 2010. The scores above reflect the number of survey questions, which tested household understanding of IBLI contract details, correctly answered. Significance is indicated by: *** p<0.01, ** p<0.05, * p<0.1

Table 5.Rangeland conditions during each sales window as predictors of final index value

Variable	Index	Index	Index
Pre-Czndvi	-0.0067**		-0.0059***
	(0.0026)		[0.0001]
Expected Rangeland Condition[1]:			
Good		-0.0790***	-0.0525***
		[0.0042]	[0.0038]
Normal		-0.0470***	-0.0387***
		[0.0040]	[0.0037]
District Fixed Effects:			
Laisamis	-0.0399	-0.0015	-0.0291***
	(0.0698)	[0.0019]	[0.0014]
Loiyangalani	-0.0473	-0.0368***	-0.0424***
	(0.0692)	[0.0016]	[0.0011]
Maikona	-0.0114	-0.0026*	-0.0134***
	(0.0694)	[0.0016]	[0.0011]
Constant	0.1009*	0.1727***	0.1365***
	(0.0502)	[0.0031]	[0.0029]
Observations	16	3,696	3,669
R-squared	0.3945	0.1167	0.3910

Four seasons' data for four divisions with Central & Gadamoji division dummy omitted. [1]The expected conditions variables are the division-season average of a set of dummy variables for expected conditions are: good, normal, or bad. *Expected conditions: Bad* is the omitted category. Standard errors in parentheses. Robust and clustered standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 6. Description of key variables

Variable	Data Frequency	Description
Male	Annual	Sex of the head of household (1=male).
Age of Head	Annual	Age of the head of household (years).
Education	Annual	Maximum education level achieved within the household (years).
Risk Aversion: Neutral	Constant	Following Binswanger (1980), households were allowed to choose from a menu of real gambles in which level of risk and expected outcome were positively correlated. Each household participated in the experiment once during their first survey round. Households are then placed into a risk aversion category according to the lottery that they choose. The categories are risk neutral, moderately risk averse, and extremely risk averse.
Risk Aversion: Moderate	Constant	
Risk Aversion: Extreme	Constant	
Dependency Ratio	Annual	Ratio of members that are younger than 15 years, older than 55 years, disabled, or clinically ill.
Social Groups	Annual	A count of the number of informal groups in which the household participates. This variable is lagged by one period in the analysis.
Asset Index	Annual	The asset index is generated by a factor analysis performed on more than 30 variables capturing asset ownership from the following categories: productive assets, household construction materials, household facilities, cooking and lighting fuels, and consumer durables. This variable is lagged by one period in the analysis.
Ln income	Seasonal	Ln(1+ average monthly income) where income is the sum of the value of earnings, milk production, livestock slaughter, and livestock sales. Earnings include earnings from sale of crops, salaried employment, pensions, casual labor, business, petty trading, gifts, and remittances, expressed in Kenyan shillings (Ksh). This variable is lagged by one period in the analysis.
Ratio Livestock Income	Seasonal	Ratio of income that is generated through milk production, livestock slaughter or livestock sales. This variable is lagged by one period in the analysis.
Herd Size	Seasonal	Average herd size during the sales window (1 TLU=0.7 camels=1 cattle=10 sheep=10 goats). This variable is lagged by one period in the analysis.
Livestock Mortality Rate	Seasonal	Seasonal livestock mortality rate is calculated by dividing total losses within a season by the total herd owned within that season. Total herd owned is the sum of beginning herd size and all additions to the herd during the season. This variable is lagged by one period in the analysis.
Risk	Constant	Within household variance in livestock mortality rate
Savings	Annual	A dummy variable that is equal to one if the household has cash savings sufficient to purchase IBLI insurance for ten TLUs. Savings are estimated by summing the total monies held at home, in merry-go-round groups, in micro-finance institutions, in savings and credit cooperatives, in bank accounts, with traders or shops, and in M-Pesa (a mobile-based micro-finance institution) accounts. This variable is lagged by one period in the analysis.
HSNP	Seasonal	Participation in HSNP (1=participant). This variable is lagged by one period in the analysis.
HSNP Community	Seasonal	Community is an HSNP target community (1=target community).
Expected Rangeland: Good/Normal/Poor	Annual	A set of three dummy variables reflecting that the respondent's prediction of coming season's rangeland conditions were: much above normal or above normal (Good=1), normal (Normal=1), or somewhat below normal or much below normal (Poor=1).
Ln(Effective Price)	Seasonal	Log of the price for one TLU of coverage after coupon discounts (ln(Ksh)).
Observed Design Error	Seasonal	The mean observed design error (%).
Correlation(M,CL)	Constant	The correlation between individual and covariate seasonal livestock mortality rates. For households with no variation in livestock mortality rate, this is set to zero.
IBLI game	Constant	Household participated in the IBLI educational game in 2010 (1=participant).
IBLI coverage	Seasonal	Household has existing IBLI coverage (1=true).
Coupon	Seasonal	Household received a discount coupon (1=true).
Pre-Czndvi	Seasonal	Preceding season's cumulative standardized normalized difference vegetation index.
Division Livestock Mortality	Division Constant	The eight-period average loss rate of all households within each division.
Division Risk	Division Constant	The within-household variance in loss rate averaged across all households in each division.
Division Correlation	Division Constant	The within-household correlation between individual loss rate and covariate loss rate averaged across all households in each division.

Table 7. Summary statistics

Variable	Never Purchase (N=450) Mean	Never Purchase (N=450) Std. Err.	Did Purchase (N=382) Mean	Did Purchase (N=382) Std. Err.	Difference	t-stat
Male	0.57	0.03	0.63	0.04	0.06	1.09
Age	47.10	1.05	48.67	1.80	1.57	0.76
Education	3.75	0.25	4.01	0.37	0.26	0.59
Risk Aversion:						
Neutral	0.26	0.03	0.25	0.04	-0.01	-0.14
Moderate	0.41	0.04	0.53	0.05	0.12	2.05 **
Extreme	0.33	0.03	0.22	0.04	-0.11	-2.28 **
Dependency Ratio	0.63	0.01	0.58	0.02	-0.04	-2.08 **
Social Groups	0.51	0.04	0.66	0.05	0.15	2.42 **
Asset Index	-0.14	0.05	-0.11	0.06	0.03	0.32
Income	7,190	465	6,997	454	-193	-0.30
Ratio Livestock Income	0.63	0.02	0.63	0.03	0.00	0.11
Herd Size	15.01	1.02	13.06	1.00	-1.94	-1.35
Livestock Mortality Rate	0.15	0.01	0.13	0.01	-0.02	-2.80 ***
Savings	0.08	0.01	0.08	0.01	0.00	-0.16
HSNP	0.27	0.02	0.26	0.03	-0.01	-0.27
HSNP Community	0.74	0.03	0.65	0.04	-0.08	-1.80 *
Expected Rangeland Conditions:						
Good	0.45	0.01	0.45	0.02	0.00	0.08
Normal	0.34	0.01	0.31	0.02	-0.03	-1.34
Poor	0.21	0.02	0.24	0.02	0.03	0.97
Pre-Czndvi	-2.77	0.09	-2.99	0.12	-0.22	-1.43
IBLI Coverage	0.00	0.00	0.13	0.00	0.13	35.24 ***
Risk (X 100)	5.84	0.52	4.05	0.33	-1.79	-2.92 ***
Correlation(M, CL)	0.44	0.02	0.46	0.03	0.02	0.43
IBLI Game	0.24	0.03	0.24	0.03	0.00	-0.02
Ln(Effective Price)	6.22	0.01	6.14	0.01	-0.07	-4.09 ***
Observed Design Error (%)	2.29	0.07	2.51	0.07	0.21	2.12 **
Coupon	0.55	0.02	0.64	0.02	0.09	2.78 ***

This table only includes the 832 balanced panel households in order to correctly categorize the “Never Purchase” households and maintain consistency in the periods and shocks captured in the summary statistics. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8. Average marginal effects (AME) on IBLI uptake, from probit

	Pooled		Conditional FE	
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.
Household Period-Specific Characteristics:				
Male	0.0327	(0.0220)	0.0302	(0.0218)
Dependency Ratio	-0.0922*	(0.0536)	0.0121	(0.0888)
Social Groups [L]	0.0270**	(0.0105)	-0.0004	(0.0118)
Asset Index [L]	-0.0683***	(0.0261)	-0.0729**	(0.0290)
Ln(Income) [L]	0.0039	(0.0077)	-0.0005	(0.0060)
Ratio income livestock [L]	-0.0667*	(0.0341)	-0.0335	(0.0331)
TLU [L]	0.0002	(0.0010)	0.0019**	(0.0010)
Livestock Mortality Rate [L]	0.0033	(0.0378)	0.0289	(0.0386)
Savings (10TLU) [L]	-0.0358	(0.0345)	-0.0481	(0.0402)
HSNP [L]	0.0525**	(0.0231)	0.0558**	(0.0224)
Household Average Characteristics:				
Dependency Ratio			-0.1434**	(0.0583)
Social Groups			0.0570***	(0.0176)
Asset Index			-0.0089	(0.0204)
Ln(Income)			0.0041	(0.0089)
Ratio Income Livestock			-0.0207	(0.0454)
TLU			-0.0010	(0.0007)
Livestock Mortality Rate			-0.1084	(0.2105)
Savings (10TLU)			-0.0840	(0.0690)
Expected Rangelands: Good [#]			-0.0599	(0.0666)
Expected Rangelands: Normal [#]			-0.0746	(0.0657)
Prospective Adverse Selection:				
Expected conditions: Good [#]	-0.0514**	(0.0231)	-0.0311	(0.0243)
Expected conditions: Normal [#]	-0.0219	(0.0218)	0.0007	(0.0225)
Pre-CZNDVI	-0.0010	(0.0014)	-0.0009	(0.0013)
Division Livestock Mortality	0.0570***	(0.0203)	0.0577***	(0.0198)
Division Risk	-0.0522**	(0.0210)	-0.0589***	(0.0207)
Division Correlation	0.2814*	(0.1636)	0.2964*	(0.1671)
Product Related Characteristics :				
Existing IBLI Coverage	0.0239	(0.0482)	0.0172	(0.0497)
Risk	-0.6914***	(0.2416)	-0.3257	(0.3505)
Correlation	0.0064	(0.0267)	0.0090	(0.0272)
Extension Game	0.0043	(0.0220)	0.0104	(0.0215)
Ln(price)	-0.0234	(0.0473)	-0.0289	(0.0437)
Observed Design Error (ODE)	-0.0080***	(0.0025)	-0.0073***	(0.0025)
Coupon Dummy	0.1779***	(0.0341)	0.1715***	(0.0304)
Observations	3,292		3,292	
F-statistic	4.11		5.10	
P-value (model)	0.00		0.00	

Additional covariates not listed above include age, age^2, average age (for the Conditional FE model), education, level of risk aversion, HSNP Village and a constant. [L] Variable is lagged one period. [#] Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 9. AME of the interacted variables on the likelihood of purchasing IBLI

	AME	Std. Err.	t	P>t	Confidence Interval	
Price =	**Observed Design Error**					
Mean-1 SD	-0.003	0.003	-1.150	0.252	-0.009	0.002
Mean Price	-0.009	0.003	-3.540	0.000	-0.014	-0.004
Mean +1SD	-0.014	0.004	-3.770	0.000	-0.022	-0.007
Observed Design Error=	**Price**					
Mean-1 SD	0.048	0.052	0.930	0.353	-0.054	0.150
Mean ODE	-0.025	0.043	-0.580	0.565	-0.108	0.059
Mean +1SD	-0.079	0.047	-1.690	0.092	-0.171	0.013
Extension Game	**Correlation(M,CL)**					
No	0.002	0.030	0.060	0.949	-0.058	0.062
Yes	0.033	0.053	0.610	0.539	-0.071	0.136

Table 10. Average marginal effects (AME) on level of purchase, conditional on purchase

	Pooled		Conditional FE	
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.
Household Period-Specific Characteristics:				
Male	0.0597	(0.0829)	0.0883	(0.0829)
Dependency Ratio	-0.3643	(0.2270)	-0.0773	(0.6169)
Social Groups [L]	0.0400	(0.0464)	-0.0082	(0.0487)
Asset Index [L]	0.1514	(0.1021)	0.2329**	(0.1019)
Ln(Income) [L]	0.0230	(0.0319)	0.0419*	(0.0240)
Ratio Income Livestock [L]	-0.3690***	(0.1289)	-0.4424***	(0.1605)
TLU [L]	0.0050	(0.0040)	0.0106**	(0.0054)
Livestock Mortality Rate [L]	0.0061	(0.1817)	-0.0091	(0.1696)
Savings (10TLU) [L]	0.1428	(0.1466)	0.2676	(0.1982)
HSNP [L]	-0.0684	(0.0932)	-0.1545	(0.1076)
Household Average Characteristics:				
Dependency Ratio			-0.4075*	(0.2449)
Social Groups			0.0998	(0.0681)
Asset Index			0.1970**	(0.0813)
Ln(Income)			0.0189	(0.0330)
Ratio Income Livestock			-0.0111	(0.1900)
TLU			0.0004	(0.0049)
Livestock Mortality Rate			0.0523	(0.7496)
Savings (10TLU)			0.0861	(0.2566)
Expected Rangelands: Good[#]			-0.6382***	(0.2469)
Expected Rangelands: Normal[#]			-0.5968**	(0.2455)
Prospective Adverse Selection:				
Expected Conditions: Good[#]	-0.3915***	(0.0869)	-0.2709***	(0.0946)
Expected Conditions: Normal[#]	-0.3270***	(0.0933)	-0.2118**	(0.0842)
Pre-CZNDVI	-0.0037	(0.0047)	-0.0100*	(0.0056)
Division Livestock Mortality	-0.1804**	(0.0829)	-0.1865**	(0.0767)
Division Risk	0.1682*	(0.0886)	0.2002**	(0.0842)
Division Correlation	-0.7921	(0.8214)	-0.9204	(0.7385)
Product Related Characteristics :				
Existing IBLI Coverage	-0.1203	(0.0979)	-0.2114**	(0.1037)
Risk	-1.1626	(1.0886)	-0.0205	(1.3111)
Correlation	-0.1599*	(0.0935)	-0.1397	(0.1035)
Extension Game	0.0156	(0.0686)	0.0267	(0.0707)
Ln(Price)	-0.4790***	(0.1214)	-0.4275***	(0.1342)
Observed Design Error (ODE)	-0.0070	(0.0124)	-0.0062	(0.0127)
Observations	3,292		3,292	
F-statistic	4.11		5.10	
P-value (model)	0.00		0.00	

Additional covariates not listed above include age, age^2, average age (for the Conditional FE model), education, level of risk aversion, HSNP Village, the Inverse Mills ratio, and a constant. [L] Variable is lagged one period. [#]Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 11. AME of the interacted variables on IBLI purchase level, conditional on purchasing

	AME	Std. Err.	t	P>t	Confidence Interval	
Price =	**Observed Design Error**					
Mean-1 SD	-0.003	0.014	-0.250	0.799	-0.030	0.023
Mean Price	-0.006	0.013	-0.500	0.619	-0.032	0.019
Mean +1SD	-0.009	0.017	-0.560	0.575	-0.042	0.023
Observed Design Error=	**Price**					
Mean-1 SD	-0.393	0.191	-2.050	0.040	-0.768	-0.017
Mean ODE	-0.424	0.138	-3.070	0.002	-0.695	-0.153
Mean +1SD	-0.455	0.129	-3.520	0.000	-0.709	-0.201
Extension Game	**Correlation(M,CL)**					
No	-0.247	0.131	-1.880	0.060	-0.505	0.010
Yes	0.218	0.153	1.420	0.155	-0.082	0.517

Appendix A: Key Features of Index Based Livestock Insurance (IBLI) Contract

The risk:
Index based Livestock Insurance (IBLI) is a product that is designed to protect against drought-related livestock mortality.

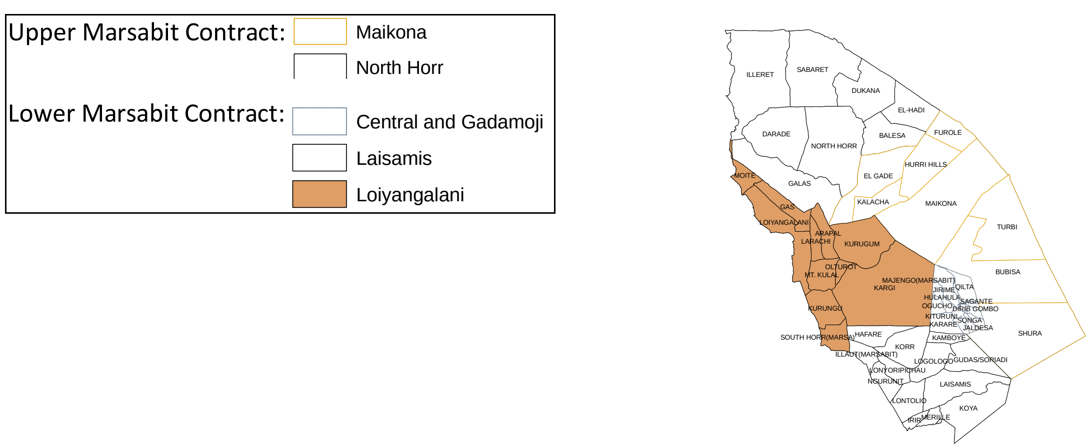
The index:
As described in Chantarat et al. (2013), the index in IBLI is the predicted livestock mortality rate. It is calculated by using a measure of vegetation coverage that is measured by satellite-based sensors, called the Normalized Difference Vegetation Index (NDVI). This vegetation measure is fed into a statistical response function that was constructed by relating historic drought related livestock mortality data to various transformation of the historic NDVI. The parameters estimated from the historic data are used to predict drought related livestock mortality from sequences of observed NDVI values.

Contract strike level:
The index threshold above which payouts are made is called the strike level. The strike level for IBLI is 15%. In other words, IBLI will compensate if predicted livestock mortality is above 15%.

Geographical coverage of contract and the index:
Marsabit District is covered by two separate contracts. There is an Upper Marsabit contract consisting of Maikona and North Horr divisions, and a Lower Marsabit contract consisting of Central, Gadamoji, Laisamis, and Loiyangalani divisions (Figure A.1).

The index – predicted livestock mortality – computed and reported at the division level. The five division—North Horr, Maikona, Loiyangalani, Laisamis and Central—could each have a different index level. Because insurance payments are made according to the index level, this means that IBLI may make different indemnity payments across divisions. Every insurance policy holder within the same division, however, will receive the same rate of insurance payment, provided that the index is above the strike.

Figure A1. IBLI Geographical Coverage



Contract premium rates and indemnity payments:

Premiums are different between the two contract regions to reflect their differences in historical risk of livestock mortality. Premium rates are reported as a percent of the value of insured livestock. From first initial sales in January of 2010 through 2012, the unsubsidized and loaded premiums were 5.4% and 9.2% in the lower and upper IBLI contract regions, respectively. At that time, those premiums were subsidized by about 40% so that pastoralists in the lower and upper regions purchased IBLI coverage at a rate of 5.5% and 3.25%, respectively.

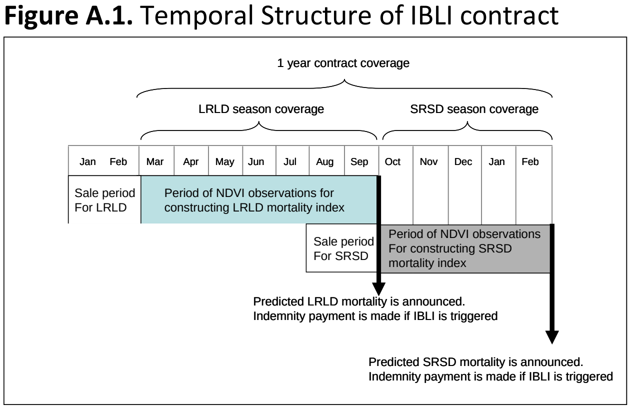
The standard livestock types for a pastoral herd will be covered: camels, cattle, sheep and goats. To arrive at a value for the insured herd, the four livestock types will be transformed into a standard livestock unit known as a Tropical Livestock Unit (TLU). TLU is calculated as follows: 1 Camel = 1.4 TLU, 1 Cattle = 1 TLU and 1 goat/sheep = 0.1 TLU. Once total TLU are calculated, the value of the total herd is computed based on average historical prices for livestock across Marsabit, at a set price per TLU insured of Ksh 15,000. The premiums are then applied to the insured value to arrive at the amount one pays for IBLI coverage for the year.

There are no indemnity payments if the index falls below the strike. If the index exceeds the strike, indemnity payments are calculated as the product of the value of the insured herd and difference between the predicted livestock mortality and the deductible.

Time Coverage of IBLI:

The figure below presents the time coverage of the IBLI. The annual contract begins at the close of a marketing window, either March 1st or October 1st. Contracts are sold only within a two month (January-February of August-September) time frame as the rainy season that typically begins right after that window may give the potential buyer information about the likely range conditions of the season to come that would affect purchase decisions. This annual contract has two potential payout periods: at the end of the long dry season based on the October 1st index reading and at the end of the short dry season based on the March 1st index readings. At these points of time, if the index exceeds 15%, active policy holders receive an indemnity payment.

Figure A.1. Temporal Structure of IBLI contract



Appendix B: Analysis of Attrition

Attrition rates averaged about 4% per year and the rate of attrition was similar between survey rounds. Table B.1 provides details on the differences between full balanced panel households and those that left. Refer to Table 6 for the variable descriptions. Note that participation in the IBLI extension game, the discount coupon, effective price, expected conditions, and design error are all related to time so that we expect there to be systematic differences in those variables between those whom we observe in all periods and those that exit, due purely to exogenous factors.

Table B.1. Summary statistics for those that stayed and those that left/entered the survey

	Full Panel (N=832)		Left/Entered (N=94)			
Variable	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Difference**	**t-stat**
Gender	0.60	0.03	0.67	0.06	0.07	1.11
Age	4.78	0.10	4.73	0.18	-0.05	-0.24
Education	3.86	0.21	4.24	0.49	0.37	0.70
Risk Aversion: Neutral	0.26	0.02	0.33	0.05	0.08	1.35
Risk Aversion: Moderate	0.46	0.03	0.35	0.05	-0.11	-1.96 *
Risk Aversion: Extreme	0.28	0.03	0.32	0.06	0.04	0.63
Dependency Ratio	0.61	0.01	0.58	0.02	-0.02	-1.03
Social Groups	0.58	0.03	0.57	0.08	-0.01	-0.17
Asset Index	-0.12	0.04	-0.18	0.08	-0.06	-0.68
Income (Kshs monthly)	7,103	325	7,266	939	164	0.16
Ratio Livestock Income	0.63	0.02	0.52	0.04	-0.11	-2.43 **
Herd Size	14.13	0.72	20.55	2.68	6.42	2.31 **
Livestock Mortality Rate	0.14	0.00	0.16	0.01	0.02	1.67 *
Savings	0.08	0.01	0.10	0.03	0.02	0.65
HSNP	0.26	0.02	0.17	0.03	-0.09	-2.36 **
Expected Conditions: Good	0.45	0.01	0.68	0.03	0.22	6.77 ***
Expected conditions: Normal	0.32	0.01	0.18	0.02	-0.14	-5.87 ***
Expected Conditions: Poor	0.22	0.01	0.14	0.02	-0.08	-3.06 ***
Risk	0.05	0.00	0.06	0.01	0.01	1.46
Correlation(M, CL)	0.45	0.02	0.68	0.05	0.23	4.38 ***
IBLI game	0.24	0.02	0.22	0.05	-0.02	-0.28
Ln(effective price)	6.18	0.01	6.31	0.03	0.12	4.04 ***
Observed Design Error (%)	2.39	0.05	1.34	0.20	-1.05	-5.16 ***
Coupon	0.59	0.02	0.50	0.05	-0.10	-1.79 *

The survey teams used a census of households with herd sizes in order to replace exit households with households from the same wealth stratum. Thus we hope that the exiting and replacement households are similar. Their descriptive statistics are found in Table B.2. As above, most of the systematic differences are likely due to duration of survey participation and likelihood of participating during certain periods rather than actual differences between households. The variables that are most worrisome are herd size, education and ratio of income from livestock, which indicate that replacement households are less educated, have much smaller herds, and are more dependent on those herds than those that left. This is most likely a result of over-sampling in the wealthy household strata, which leaves fewer eligible replacements for attrited wealthy households.[34]

Table B.2. Summary statistics for entry vs. exit households

	Exit *(N=91)*		**Enter** *(N=91)*				
Variable	**Mean**	**Std. Err.**	**Mean**	**Std. Err.**	**Difference**	**t-stat**	
Gender	0.63	0.07	0.75	0.09	0.12	1.09	
Age	4.69	0.23	4.85	0.33	0.16	0.40	
Education	4.71	0.59	2.96	0.97	-1.75	-1.55	
Risk Aversion: Neutral	0.28	0.06	0.48	0.11	0.20	1.61	
Risk Aversion: Moderate	0.35	0.06	0.35	0.10	0.00	0.02	
Risk Aversion: Extreme	0.37	0.07	0.17	0.07	-0.20	-2.02	**
Social Groups	0.60	0.03	0.53	0.04	-0.07	-1.34	
Dependency Ratio	0.54	0.09	0.63	0.14	0.08	0.49	
Asset Index	-0.15	0.10	-0.29	0.14	-0.14	-0.83	
Income (Ksh monthly)	7,804	1,304	5,829	722	-1,975	-1.32	
Ratio livestock Income	0.46	0.05	0.69	0.05	0.23	3.13	***
Herd Size	24.18	3.57	10.84	1.89	-13.34	-3.29	***
Livestock Mortality Rate	0.16	0.02	0.15	0.02	-0.01	-0.44	
Savings	0.12	0.05	0.05	0.03	-0.06	-1.15	
HSNP	0.15	0.04	0.23	0.07	0.08	0.96	
Expected Conditions: Good	0.70	0.04	0.62	0.06	-0.09	-1.19	***
Expected Conditions: Normal	0.17	0.03	0.20	0.04	0.03	0.51	***
Expected Conditions: Poor	0.13	0.02	0.19	0.05	0.06	1.04	
Risk	0.07	0.01	0.05	0.01	-0.02	-0.89	
Correlation(M, CL)	0.73	0.06	0.54	0.10	-0.18	-1.56	
Ln(Effective Price)	0.30	0.07	0.01	0.01	-0.29	-4.24	***
Observed Design Error	6.35	0.04	6.18	0.05	-0.17	-2.80	***
IBLI Game	0.61	0.14	3.31	0.37	2.70	6.85	***
Coupon	0.65	0.07	0.30	0.04	-0.35	-4.32	***

[34] Large portions of the middle and high wealth strata were sampled in some smaller communities. In such cases, finding within strata replacement households can be difficult. Pastoral mobility and demand for herding labor far from households and community centers further exacerbates the challenges of replacing households from an already attenuated roster.

Appendix C: Asset Index

The asset index is constructed by performing a factor analysis on a set of variables meant to capture variation in household wealth. This approach is discussed in Sahn and Stifle (2000). The variables focus on five general categories: household construction materials, household facilities, cooking and lighting fuels, and household durables. Because the list of possible durables is extremely long (more than 70), they are aggregated by value (small, medium, large) and use (productive, other) except for large assets which are divided into those with motors and those without. Categorization was performed by the authors and is clearly not the only method for dividing or aggregating the long list of assets. When in doubt as to which category to place an item, we relied on the frequency of ownership to guild our decision. Table C.1 includes the descriptions of each variable. Table C.2 provides the factor loadings, which were estimated using the variables described in table C.1 and including division year fixed effects.

Table C.1 Variable used in the factor analysis to generate an asset index

Improved Wall	=1 if walls are stone, brick, cement, corrugated iron, mud plastered with cement, or tin
Improved Floor	=1 if floor is cement, tile, or wood
Improved Toilet	=1 if toilet is flush or covered latrine
Improved Light	=1 if main source of lighting is electricity, gas, solar
Improved cooking appliance	=1 if main cooking appliance is jiko, kerosene stove, gas cooker, or electric cooker
Improved Fuel	=1 if main cooking fuel is electricity, paraffin, gas or charcoal
Improved furniture	Total number of the following assets: metal trunks, mosquito nets, modern chairs, modern tables, wardrobes, mattresses and modern beds
Water Source: Open	=1 if main water source is river, lake, pond, unprotected well or unprotected spring
Water Source: Protected	=1 if main water source is protected spring or protected well
Water Source: Borehole	=1 if main water source is a borehole
Water source: Tap	=1 if main water source is a public or private tap
Water Source: Rainwater catchment	=1 if main water source is a rainwater catchment (usually cement or plastic)
Water Source: tanker	=1 if main water source is water tanker (usually associated with NGO and food aid activities during drought)
Education	Maximum household education
Total cash savings	Total monies held at home, in merry-go-round groups, in micro-finance institutions, in savings and credit cooperatives, in bank accounts, with traders or shops, and in M-Pesa (a mobile-based micro-finance institution) accounts.
Land	Hectares owned
Irrigation	=1 if household owns irrigated land
Poultry	Number of chickens
Donkeys	Number of donkeys

Table C.1 continues

Table C.1 continued

Very small	Total number of the following assets: gourds, cups, scissors, and needle and thread sets.
Small tools	Total number of the following assets: anvils, panier, sickle, pickaxe, hoe, spade, machetes, spears, bows, club, chisels, hammers, files, fishing lines.
Small other	Total number of the following assets: musical instruments, traditional tools, bells, knifes, basins, sufirias, thermoses, buckets, wristwatches, jewelry
Medium tools	Total number of the following assets: Wheelbarrows, fishing nets, mobile phones, washing machines, spinning machines, weaving machines, sewing machines, bicycles, and plows.
Medium other	Total number of the following assets: water tank, jerry can, paraffin lamp, water drum, kerosene stove, charcoal stoves, ovens and radios.
Large	Total number of the following assets: animal carts, shops, stalls and boats.
Large with motor	Total number of the following assets: cars, motorbikes and tractors.

Table C.2 Factor loadings used to generate the asset index

Variables	Factor Loading
Improved Wall	0.132
Improved Floor	0.130
Improved Toilet	0.128
Improved Light	0.118
Improved cooking appliance	0.077
Improved Fuel	0.064
Improved furniture	0.165
Water Source: Open	0.004
Water Source: Protected	0.004
Water Source: Borehole	-0.008
Water source: Tap	0.040
Water Source: Rainwater catchment	0.079
Water Source: Tanker	0.021
Education	0.121
Total cash savings	0.085
Land	0.051
Irrigation	0.033
Poultry	0.081
Donkeys	0.018
Very small	0.040
Small tools	0.126
Small other	0.053
Medium tools	0.164
Medium other	0.135
Large	0.037
Large with motor	0.089

Division*period dummies included in the factor analysis.

Appendix D: Validity of Excludable Variable.

We include a dummy variable to indicate that the household received a discount coupon in the first stage selection equation but exclude it from the demand equation. The selection equation estimates found in Table E1 and Table 8 clearly indicate that receiving a coupon has a large, positive, and statistically significant impact on the likelihood of purchasing IBLI, even after accounting for size of the discount the coupon offered ($\beta=0.9866$, $p<0.01$). This effect seems purely a randomized treatment that should be irrelevant to purchase volume conditional on uptake. So that variable seems a strong candidate for exclusion from the second stage estimation of uptake volume.

Although there is no agreed upon method for testing excludability of a candidate instrument and it seems to rarely be done with selection models, we venture to provide some statistical support that the exclusion of that indicator variable does not cause bias in the demand equation estimates. Because we only have one exclusion variable, our tests rest on identification through nonlinearity on the probit model, which is likely to be very weak. First, we include the coupon dummy variable in the second stage regression. The coefficient on the coupon dummy is negative and statistically insignificant (beta=-0.122, p-value=0.366). Comparing this set of estimates with those estimated with the coupon dummy excluded, we fail to reject the null hypothesis that the joint change to remaining estimates is zero ($\chi^2(45)=1.62$, p-value=1.00). More specifically, we would expect a large change between the two models in the estimated parameter on the effective price if the receiving a coupon played an important role in determining levels of demand beyond providing a price discount. Testing for a difference in the two price parameter estimates, we cannot reject the null of no change ($\chi^2(1)=0.87$, p-value=0.352). Of course, this does not mean that the variable should be omitted, only that it has little independent effect on the level of purchase and does not result in large shifts in parameter values when included.

We can also check if the errors estimated by the demand equation without the coupon dummy vary by coupon status. Because selection is controlled for through the inverse Mills ratio and coupons were randomly distributed, there should be no omitted variable bias in the demand equation parameter estimates except potentially in effective price, but that bias was ruled out in step one. A t-test of the demand residuals over the coupon status does not reject the null of equal errors between those who received a coupon and those who did not (difference=0.054, t-statistic(529)=0.756).

Appendix E. Coefficient Estimates of Uptake and Demand for ILBI

Table E1. Coefficient estimates for probit selection

	Pooled		Conditional FE	
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.
Household Period-Specific Characteristics:				
Male	0.1493	(0.1006)	0.1412	(0.1018)
Dependency Ratio	-0.4211*	(0.2449)	0.0565	(0.4149)
Social Groups ᴸ	0.1232***	(0.0475)	-0.0018	(0.0553)
Asset Index ᴸ	-0.2876***	(0.1069)	-0.3400**	(0.1323)
Asset Index^2 ᴸ	0.0728***	(0.0266)	0.1703*	(0.0938)
Ln(income) ᴸ	-0.0632	(0.0561)	-0.0049	(0.0264)
Ln(income)^2 ᴸ	0.0053	(0.0052)	0.0078	(0.0053)
Ratio income livestock ᴸ	-0.3043**	(0.1550)	-0.1567	(0.1553)
TLU ᴸ	0.0030	(0.0058)	0.0087*	(0.0045)
TLU^2 ᴸ	-0.0001	(0.0001)	-0.0002	(0.0001)
Livestock Mortality Rate ᴸ	0.0149	(0.1728)	0.1351	(0.1795)
Savings (10TLU) ᴸ	-0.1635	(0.1579)	-0.2248	(0.1879)
HSNP ᴸ	0.2397**	(0.1061)	0.2608**	(0.1057)
Household Averages Characteristics:				
Dependency Ratio			-0.6699**	(0.2723)
Social Groups			0.2663***	(0.0811)
Asset Index			-0.0418	(0.0950)
Ln(income)			0.0193	(0.0414)
Ratio income livestock			-0.0967	(0.2127)
TLU			-0.0046	(0.0032)
Livestock Mortality Rate			-0.5062	(0.9822)
Savings (10TLU)			-0.3923	(0.3225)
Expected Rangelands: Good#			-0.2798	(0.3097)
Expected Rangelands: Normal#			-0.3487	(0.3070)
Prospective Adverse Selection:				
Expected conditions: Good#	-0.2348**	(0.1040)	-0.1451	(0.1135)
Expected conditions: Normal#	-0.0999	(0.0991)	0.0034	(0.1050)
Pre-CZNDVI	-0.0045	(0.0062)	-0.0041	(0.0061)
Division Livestock Mortality	0.2603***	(0.0907)	0.2693***	(0.0913)
Division Risk	-0.2381**	(0.0945)	-0.2749***	(0.0959)
Division Correlation	1.2845*	(0.7437)	1.3848*	(0.7789)
Product Related Characteristics :				
Existing IBLI Coverage	0.1089	(0.2207)	0.0806	(0.2326)
Risk	-3.1565***	(1.0903)	-1.5218	(1.6403)
Correlation	0.0008	(0.1388)	0.0091	(0.1428)
Extension Game	-0.0360	(0.1444)	-0.0168	(0.1452)
Correlation X Game	0.1197	(0.2712)	0.1401	(0.2716)
Ln(price)	0.2732	(0.2254)	0.2419	(0.2181)
Observed Design Error (ODE)	0.4001**	(0.1592)	0.3990**	(0.1727)
Ln(price) X ODE	-0.0736***	(0.0269)	-0.0730**	(0.0292)
Coupon Dummy	0.8123***	(0.1472)	0.8011***	(0.1384)
Observations	3,292		3,292	
F-statistic	4.11		5.10	
P-value (model)	0.00		0.00	

Additional covariates not listed above include age, age^2, average age (for the Conditional FE Model), education, level of risk aversion, HSNP Village, and a constant. ᴸ Variable is lagged one period. #Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E2. Estimated demand coefficients, conditional on purchase

	Pooled		Conditional FE	
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.
Household Period-Specific Characteristics:				
Male	0.0597	(0.0829)	0.0883	(0.0829)
Dependency Ratio	-0.3643	(0.2270)	-0.0773	(0.6169)
Social Groups ᴸ	0.0400	(0.0464)	-0.0082	(0.0487)
Asset Index ᴸ	0.1500	(0.0961)	0.2368**	(0.1036)
Asset Index^2 ᴸ	-0.0046	(0.0235)	-0.0659	(0.0687)
Ln(income) ᴸ	0.0821	(0.0592)	0.0454**	(0.0226)
Ln(income)^2 ᴸ	-0.0038	(0.0050)	-0.0044	(0.0054)
Ratio income livestock ᴸ	-0.3690***	(0.1289)	-0.4424***	(0.1605)
TLU ᴸ	0.0066	(0.0053)	0.0099*	(0.0056)
TLU^2 ᴸ	-0.0001	(0.0001)	-0.0002**	(0.0001)
Livestock Mortality Rate ᴸ	0.0061	(0.1817)	-0.0091	(0.1696)
Savings (10TLU) ᴸ	0.1428	(0.1466)	0.2676	(0.1982)
HSNP ᴸ	-0.0684	(0.0932)	-0.1545	(0.1076)
Household Averages Characteristics:				
Dependency Ratio			-0.4075*	(0.2449)
Social Groups			0.0998	(0.0681)
Asset Index			0.1970**	(0.0813)
Ln(income)			0.0189	(0.0330)
Ratio income livestock			-0.0111	(0.1900)
TLU			0.0004	(0.0049)
Livestock Mortality Rate			0.0523	(0.7496)
Savings (10TLU)			0.0861	(0.2566)
Expected Rangelands: Good#			-0.6382***	(0.2469)
Expected Rangelands: Normal#			-0.5968**	(0.2455)
Prospective Adverse Selection:				
Expected conditions: Good#	-0.3915***	(0.0869)	-0.2709***	(0.0946)
Expected conditions: Normal#	-0.3270***	(0.0933)	-0.2118**	(0.0842)
Pre-CZNDVI	-0.0037	(0.0047)	-0.0100*	(0.0056)
Division Livestock Mortality	-0.1804**	(0.0829)	-0.1865**	(0.0767)
Division Risk	0.1682*	(0.0886)	0.2002**	(0.0842)
Division Correlation	-0.7921	(0.8214)	-0.9204	(0.7385)
Product Related Characteristics :				
Existing IBLI Coverage	-0.1203	(0.0979)	-0.2114**	(0.1037)
Risk	-1.1626	(1.0886)	-0.0205	(1.3111)
Correlation	-0.2779**	(0.1205)	-0.2472*	(0.1312)
Extension Game	-0.2157*	(0.1106)	-0.1840*	(0.1113)
Correlation X Game	0.5103**	(0.2057)	0.4648**	(0.2139)
Ln(price)	-0.4311**	(0.1829)	-0.3880*	(0.2016)
Observed Design Error (ODE)	0.0465	(0.1347)	0.0380	(0.1222)
Ln(price) X ODE	-0.0089	(0.0224)	-0.0073	(0.0205)
Observations	3,292		3,292	
F-statistic	4.11		5.10	
P-value (model)	0.00		0	

Additional covariates not listed above include age, age^2, average age (Conditional FE Model), education, level of risk aversion, HSNP Village, inverse Mills Ratio, and a constant. ᴸ Variable is lagged one period. #Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix F: Shapley Goodness of Fit Decomposition

A Shapley's goodness of fit (GOF) decomposition is used to determine the level of variation in demand that is captured by categories of variables (Kolenikov 2000; Shapley 1953; Shorrocks 2013).[35] The variable categories include: household demographics, household finances, prospective intertemporal adverse selection, prospective spatial adverse selection, idiosyncratic risk & knowledge, design risk & price, other, and the instrument variable. A two-stage Heckman approach, rather than the maximum likelihood approach used in the main body of the paper, is used here in order to examine the contributions of the variable groups in both the uptake and demand analysis. In addition, we use the pooled, rather than conditional fixed effects, approach here in order to reduce the computational burden. Notice that the pooled and conditional fixed effects estimates are generally very similar.

Tables F1 and F2 include the two-stage estimates and estimated group contributions to each stage's (uptake and level of purchase) GOF. The pooled maximum likelihood estimates from the Heckman selection model (from Table E1 and Table E2) are also included as evidence that the two models result in very similar estimates and that the decomposition of the two-stage estimates are likely to be reflective of the contributions in the maximum likelihood Heckman model.[36]

Household characteristics clearly play a role in uptake but are unable to account for even half of the variation captured by the model (Table F1). Temporal and spatial adverse selection provide similar contributions and their combined impacts are similar to that of the relative importance of covariate risk. The three design risk and price variables account for 18% of the Pseudo R^2 measure, more than any other group except for our instrumental variable.

The role of adverse selection in the fit of our model is greater for level of demand than uptake. Conversely, the role of design risk and price has fallen considerably. In addition, income and wealth have become much more important while the importance of covariate risk has changes very little.

In summary, the total contribution made by adverse selection and product related characteristics towards the GOF are greater than that of a large set of familiar household characteristics in both uptake and level of demand models. Our models would perform much worse with these crucial estimates of basis risk and adverse selection.

[35] We use the STATA user-written command *shapley2* (Juárez 2014).

[36] The ML Heckman estimates are generated in a single step so that we cannot examine the goodness of fit contributions in each process separately.

Table F1. Decomposition of Pseudo R^2 for uptake probit

	Heckman ML Probit		2 Step Probit		Shapley Decomposition of
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.	Pseudo $R^{2\ A}$
Household Period-Specific Characteristics:					
Demographics: [B]					12.40%
Male	0.1493	(0.1006)	0.1492	(0.1038)	
Dependency Ratio	-0.4211*	(0.2449)	-0.4214*	(0.2533)	
Social Groups [L]	0.1232***	(0.0475)	0.1230**	(0.0492)	
Financial:					14.42%
Asset Index [L]	-0.2876***	(0.1069)	-0.2875***	(0.1104)	
Asset Index^2 [L]	0.0728***	(0.0266)	0.0728***	(0.0276)	
Ln(income) [L]	-0.0632	(0.0561)	-0.0631	(0.0579)	
Ln(income)^2 [L]	0.0053	(0.0052)	0.0053	(0.0054)	
Ratio income livestock [L]	-0.3043**	(0.1550)	-0.3044*	(0.1594)	
TLU [L]	0.0030	(0.0058)	0.0030	(0.0060)	
TLU^2 [L]	-0.0001	(0.0001)	-0.0001	(0.0001)	
Livestock Mortality Rate [L]	0.0149	(0.1728)	0.0147	(0.1781)	
Savings (10TLU) [L]	-0.1635	(0.1579)	-0.1631	(0.1645)	
HSNP [L]	0.2397**	(0.1061)	0.2398**	(0.1097)	
Prospective Adverse Selection:					
Intertemporal:					2.46%
Expected conditions: Good[#]	-0.2348**	(0.1040)	-0.2346**	(0.1081)	
Expected conditions: Normal[#]	-0.0999	(0.0991)	-0.0997	(0.1022)	
Pre-CZNDVI	-0.0045	(0.0062)	-0.0045	(0.0064)	
Spatial:					5.12%
Division Livestock Mortality	0.2603***	(0.0907)	0.2603***	(0.0937)	
Division Risk	-0.2381**	(0.0945)	-0.2381**	(0.0974)	
Division Correlation	1.2845*	(0.7437)	1.2838*	(0.7666)	
Product Related Characteristics :					
Idiosyncratic Risk & Knowledge:					5.46%
Risk	-3.1565***	(1.0903)	-3.1567***	(1.1241)	
Correlation	0.0008	(0.1388)	0.0007	(0.1439)	
Extension Game	-0.0360	(0.1444)	-0.0359	(0.1509)	
Correlation X Game	0.1197	(0.2712)	0.1197	(0.2814)	
Design Risk & Price:					21.13%
Ln(price)	0.2732	(0.2254)	0.2733	(0.2326)	
Observed Design Error (ODE)	0.4001**	(0.1592)	0.4000**	(0.1645)	
Ln(price) X ODE	-0.0736***	(0.0269)	-0.0736***	(0.0278)	
Instrumental Variable:					35.32%
Coupon Dummy	0.8123***	(0.1472)	0.8124***	(0.1515)	
Observations	3,292		3,292		
F-statistic [Wald χ^2]	4.11		[165.48]		
P-value (model)	0		0		
Pseudo R2			0.135		

[A] The Shapley decomposition is performed on eight groups of variables indicated by the bold labels on the left using the 2-stage probit estimates. A group containing existing IBLI coverage and an indicator that the household is in an HSNP targeted community was also included in the regressions and decomposition; its Shapley contribution was 3.82%. [B] Additional covariates in the demographics group include age, age^2, education, level of risk aversion, and existing coverage. [L] Variable is lagged one period. [#]Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table F2. Decomposition of R^2 for level of purchase, conditional on purchase

	Demand (MLE)		Demand (2 Step)		Shapley
VARIABLES	Coefficient	Std. Err.	Coefficient	Std. Err.	Decomposition of Pseudo R2 [A]
Household Period-Specific Characteristics:					
Demographics: [B]					14.08%
Male	0.0597	(0.0829)	0.0602	(0.0889)	
Dependency Ratio	-0.3643	(0.2270)	-0.3650	(0.2443)	
Social Groups [L]	0.0400	(0.0464)	0.0405	(0.0499)	
Financial:					30.85%
Asset Index [L]	0.1500	(0.0961)	0.1489	(0.1058)	
Asset Index^2 [L]	-0.0046	(0.0235)	-0.0042	(0.0258)	
Ln(income) [L]	0.0821	(0.0592)	0.0821	(0.0631)	
Ln(income)^2 [L]	-0.0038	(0.0050)	-0.0038	(0.0053)	
Ratio income livestock [L]	-0.3690***	(0.1289)	-0.3705***	(0.1408)	
Livestock Mortality Rate [L]	0.0066	(0.0053)	0.0066	(0.0057)	
TLU [L]	-0.0001	(0.0001)	-0.0001	(0.0001)	
TLU^2 [L]	0.0061	(0.1817)	0.0068	(0.1952)	
Savings (10TLU) [L]	0.1428	(0.1466)	0.1417	(0.1596)	
HSNP [L]	-0.0684	(0.0932)	-0.0674	(0.1004)	
Prospective Adverse Selection:					
Intertemporal:					18.39%
Expected conditions: Good[#]	-0.3915***	(0.0869)	-0.3924***	(0.0947)	
Expected conditions: Normal[#]	-0.3270***	(0.0933)	-0.3273***	(0.1005)	
Pre-CZNDVI	-0.0037	(0.0047)	-0.0037	(0.0051)	
Spatial:					13.84%
Division Livestock Mortality	-0.1804**	(0.0829)	-0.1798**	(0.0897)	
Division Risk	0.1682*	(0.0886)	0.1674*	(0.0956)	
Division Correlation	-0.7921	(0.8214)	-0.7894	(0.8859)	
Product Related Characteristics :					
Idiosyncratic Risk & Knowledge:					5.42 %
Risk	-0.1203	(0.0979)	-1.1753	(1.1859)	
Correlation	-1.1626	(1.0886)	-0.2778**	(0.1302)	
Extension Game	-0.2779**	(0.1205)	-0.2156*	(0.1194)	
Correlation X Game	-0.2157*	(0.1106)	0.5100**	(0.2216)	
Design Risk & Price:					15.68%
Ln(price)	-0.4311**	(0.1829)	-0.4326**	(0.1981)	
Observed Design Error (ODE)	0.0465	(0.1347)	0.0480	(0.1462)	
Ln(price) X ODE	-0.0089	(0.0224)	-0.0091	(0.0244)	
Observations	3,292		547		
F-statistic	4.11		4.49		
P-value (model)	0.00		0.00		
R^2			0.2582		

[A] The Shapley decomposition is performed on seven groups of variables indicated by the bold labels on the left. A group containing existing IBLI coverage, an indicator that the household is in an HSNP targeted community, and the inverse Mills Ratio was also included in the regressions; its Shapley contribution was 2.45%. [B] Additional covariates in the demographics group include age, age^2, education, level of risk aversion. [L] Variable is lagged one period. [#]Omitted variable is E*xpected conditions: poor*. Robust and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1