Fail-safe index insurance without the cost: a satellite based conditional audit approach*

Jon Einar Flatnes†, Michael R. Carter‡

Draft 2/17/2017

Abstract

While index insurance offers a compelling solution to the problem of covariant risk among smallholder farmers in developing countries, most weather based contracts suffer from poor quality due to a low correlation between the index and farmer losses. Moreover, area yield contracts are generally infeasible due to high data collection costs. This paper proposes and analyzes an alternative index insurance contract, which combines a satellite based index with the potential for a second-stage audit. The satellite index is created using a model which takes publicly available high-resolution satellite data and converts them into an index. Using plot level panel data from a retrospective yield survey among smallholder rice farmers in Tanzania, we estimate a response function mapping actual yields to this index. If the index fails to predict losses, an audit, in the form of a crop-cutting exercise, can be invoked at farmers' request. Our results suggest that the index explains approximately 60% of the variation in zone level yields across years. Moreover, when combining this index with a possibility of an audit, this contract closely mimics the payouts of an area yield contract, but at a fraction of the cost, as audits are estimated to occur only 17% of the time. Based on expected utility analysis, we also show that demand for this contract would exceed that of an area yield contract and a pure satellite contract under reasonable loading cost assumptions.

Keywords: []

JEL Classification Codes: []

*We would like to thank Travis Lybbert, Stephen Boucher, Ghada Elabed, Jean Paul Petraud, Isabel Call, Wenbo Zou, Emilia Tjernstrom, Thomas Barre, Eliana Zeballos, Patrick McLaughlin, Jacob Humber, Asare Twum Barima and Abbie Turianski for their insightful comments and feedback. We would also like to thank VisionFund Tanzania, the enumerators, community leaders and participants in the survey for their time and service. This research was made possible, in part, through funding provided by the U.S. Agency for International Development through the BASIS Assets and Market Access Collaborative Research Support Program.

†Assistant Professor, Department of Agricultural Environmental, and Development Economics, The Ohio State University, Columbus, OH 43210; email: flatnes.1@osu.edu

‡Professor, Department of Agricultural and Resource Economics, University of California, Davis, One Shields Avenue, Davis, CA 95616; email: mrcarter@ucdavis.edu

1 Introduction

An overwhelming body of literature has left no doubt that risk poses one of the greatest threats to development in low-income economies. Particularly small scale farmers are plagued by risk, as weather variation is the largest source of risk in agriculture (Cole and Giné 2010; Giné and Yang 2009) and because such risk is spatially correlated, making local risk sharing mechanisms ineffective and affecting everyone in the community. While index insurance offers a compelling solution to the problem of covariant risk, these products have generally suffered from low demand. For example, Cole and Giné (2010) find that the adoption rate for a rainfall based index insurance product offered to smallholder farmers in two regions in India is close to zero. This somewhat disconcerting observation has prompted several empirical and experimental studies attempting to isolate the determinants of the demand for index insurance. While this evidence suggests that price, liquidity, interlinkage with credit, and trust all have an effect on demand, an issue which has only recently received attention in the development literature is the quality of the contract itself. For example, Clarke (forthcoming) develops a theoretical model demonstrating that the presence of basis risk, or the risk that a farmer suffers a loss but the insurance index does not trigger a payout, could make it optimal for a risk averse farmer to decline index insurance coverage. This result stems from the fact that if a farmer suffers a loss but the insurance index does not trigger, she is worse off in the bad state of the world with insurance than without insurance due to the payment of the insurance premium.

The objective of an insurance contract is to maximize the welfare of farmers by smoothing income fluctuations at the lowest possible cost. Traditional individual indemnity contracts do in theory provide perfect coverage against actual losses, but are fraught with problems of moral hazard and adverse selection. Extensive monitoring and verification could reduce the magnitude of these problems, but this is often prohibitively expensive, thus making such contracts infeasible for smallholder agriculture. Index insurance contracts, on the other hand, are based on a verifiable index, that is correlated with, but cannot be influenced by individual outcomes. Hence, index insurance contracts essentially eliminate moral hazard and adverse selection and typically require no monitoring or verification of individual outcomes, making such contracts affordable even when the insured amount is small. By construction, index insurance contracts are not intended to cover idiosyncratic losses but rather to protect farmers against covariant yield or price shocks beyond the control of individual producers. Therefore, index insurance contracts are not insurance contracts per se, but rather hedging instruments, which may or may not help individual farmers smooth income. The residual level of risk faced by the farmer is in the finance literature referred to as basis risk, which if sufficiently high, can significantly reduce the value of index insurance to farmers, even to the point where it is welfare-reducing.

Under an index insurance contract, this basis risk can be decomposed into two types of risk:

1) design risk; and 2) non-insurable idiosyncratic risk. Design risk arises due to the inability of an index to accurately measure covariant losses within a defined insurance zone, and manifests itself as an imperfect correlation between the index and actual average zone-level losses. In contrast, idiosyncratic risk is due to individual shocks that are independent across people within the zone, and is thus uninsurable through index insurance. If we assume that farmers face no price risk and that average yields within a zone can be accurately measured, then an index insurance contract based on area yields will have no design risk. Hence, an area yield contract will offer the best protection that an index insurance contract can have within a given zone; however, such contracts are typically infeasible due to the high cost of collecting yield data for each insurance period.

Instead, index insurance projects in developing countries have almost exclusively used precipitation-based indices, which have been shown to carry a relatively high level of basis risk, with correlations between precipitation and biomass growth at the weather station ranging from .26 to .70 (Sims and Singh, 1978; Price et al., 1998). Even lower correlations have been reported for the link between rainfall indexes and crop production (Martyniak 2007; Staggenborg et al. 2008). Making matters worse, weather stations are typically sparse and the agro-climatic landscape in Sub-Saharan Africa is generally quite heterogeneous, with the result that basis risk increases sharply with distance from the weather station. Smith and Watts (2009) show that if the spatial correlation between yields at two farms is .5, the correlation between the precipitation index and yields at the farm located some distance away from the weather station would be .35 or lower. Moreover, their simulation results indicate that in this scenario, there is approximately a 60% probability that a farmer experiencing a severe yield loss (yields less than 50% of average) will receive no indemnity payout. Moreover, in a study of 270 weather based index insurance products in India over the period 1999-2007, Clarke et al. (2012) shows that when there is a 100% loss at the sub-district level, the average claim payment made was only 12%. These discouraging results prompt the need for more innovative insurance contracts and further empirical research on the impact of basis risk.

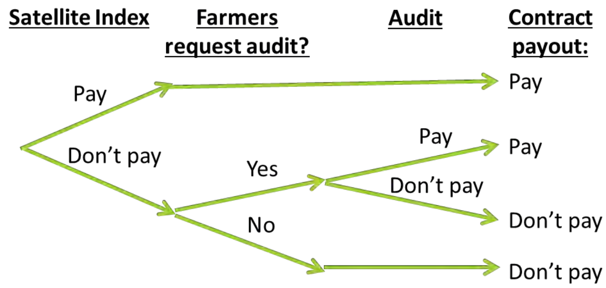
In this paper, we propose and analyze an alternative to the typical weather based contract, with the purpose of improving insurance quality without incurring the costs that area yields contracts typically do. In particular, using yield data from a retrospective yield survey in Tanzania, we design an innovative contract that uses relatively inexpensive satellite data as the primary index but includes a provision which allows farmers to request an audit if the primary index fails to predict area losses. We then ask whether this satellite based conditional audit contract can match the performance of an area yield contract, but without the high costs typically associated with collecting area yield data. Finally, we analyze whether such a contract would reduce risk and improve welfare for smallholder farmers. To our knowledge, this is the first paper to consider an index insurance contract that bases payouts on a combination of satellite data and audits. Moreover, this paper is also one of the first to analyze the impact of basis risk on the welfare of individual farmers using a plot-level panel data. Only one other paper (Jensen et al., 2014) studies basis risk on a household level. In their paper, they study a particular index based livestock insurance (IBLI) made available to pastorarists in Northern Kenya, and find that IBLI reduces exposure to covariate risk due to high loss events by an average of 63%.

The use of satellite data to estimate losses has received surprisingly little attention in the index insurance literature. Giné et al. (2010) suggest that such an index is likely to have less basis risk than a rainfall based index and a few informal studies provide some evidence of this (e.g. Carter and Laajaj, 2009). In addition, several studies from the remote sensing literature have found varying levels of correlation between satellite based indices and observed crop yields. For example, Rosema (1993) develops a sophisticated model of evapotranspiration to simulate crop yields and tests his predictions against several years of recorded biomass observations for 25 sites in Mali. While he finds a relatively low R^2 (.2) using farm level data, a much stronger correlation (R^2 between .68 and .84) was found when using an improved version of her model to predict maize yields in Zambia and Zimbabwe both at a provincial and at communal level for the years 1994-1997 (EARS, 2012). Using a satellite based measure of water use efficiency (WUE) and radiation use efficiency (RUE) to model wheat yield in India, Bhattacharaya et al. (2011) finds an R^2 of .81 and .64 when comparing predicted yields from the RUE model and the WUE model, respectively, to district level average yield statistics for three seasons (2002/03-2004/05) in twelve selected wheat growing districts within four agro-climatic zones in India. Moreover, while a large number of studies have used various satellite based measures to predict crop yields, most of these are relying on county level data or experimental plot data over a short time period to validate their models. Hence, by using plot-level data that covers up to ten planting seasons, we are able to produce more accurate estimates of the relationship between the satellite based measures and crop yields.

Our results show that both a contract based on a satellite index alone, and the satellite based conditional audit contract exhibit a lower level of design risk than most weather based contracts that exist on the market today. In particular, our satellite measure of Gross Primary Production (GPP) explains 64% of the across-year variance in zone-level yields. Moreover, when adding the possibility of an audit in the case the satellite index fails to predict losses, we show that the design risk is virtually eliminated and that the performance of the satellite based conditional audit contract closely mimics that on an area yield contract, despite the fact that audits would only be requested 17% of time. Our analysis also demonstrate that farmers' willingness to pay for this contract exceeds the actuarially fair premium by approximately 13% for farmers with a normal level of risk aversion, and that under reasonable loading cost assumptions, demand may be as high as 34%.

The rest of the paper is structured as follows. Section 2 lays out the index insurance contracts and the basic framework analyzed in this paper, while section 3 describes the data and the study area. In section 4, we discuss the methodology used to analyze the performance of the index insurance contracts and present the results. Section 5 concludes.

Figure 1: Structure of a satellite based conditional audit contract

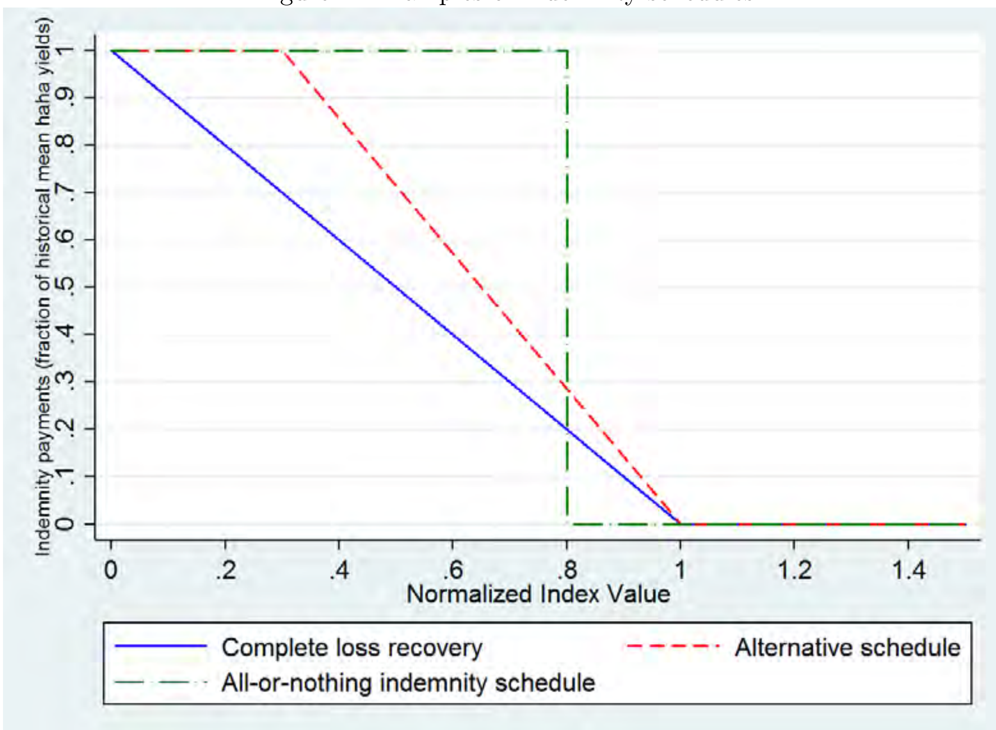


2 A satellite based conditional audit approach

The challenge of improving index insurance quality while keeping costs low has inspired the development of a satellite based index insurance contract which uses audits for verification only if the primary index fails. Under such a contract, the primary insurance index would be based on relatively inexpensive satellite measures of crop health. While this satellite index is designed with the objective of minimizing design risk and performs better than many existing weather based index insurance contracts marketed to farmers, it is still an imperfect predictor of area yield losses, leaving substantial basis risk in the hands of farmers. Hence, to combat this residual design risk, we add a second layer of protection by letting farmers request an audit if they believe the satellite index does not accurately reflect area yield losses. Figure 1 shows how such a contract might be structured. Note that if the satellite index indicates that a payment should be made, the contract will pay out, even if farmers have not experienced a loss. While such false positives might be seen as a windfall to farmers in an otherwise good year, they will still reduce the quality of the contract, as farmers implicitly pay for these windfalls through higher premiums. The idea of using a secondary audit is currently being piloted in Ethiopea (see Berhane et al., (2014)), where it has been termed "gap insurance". In their project, the audit takes the form of a crop-cutting exercise, and payouts are made if the average yield based on this verification falls below the trigger, even if the primary index has not triggered a payout. To prevent unnecessary audits, incentive compatible penalties may be imposed. For example, if farmers request an audit, but measured yields from the crop-cutting exercise are either consistent with or higher than the yields predicted by the index, the farmers would have to bear the cost of this audit. However, if the result of the audit uncovers yield losses beyond those predicted by the satellite index, the insurance company would pay for the audit. This system ensures that farmers would only request an audit when they are confident that the satellite index is incorrect.

To create a framework the remaining analysis, consider an insurance contract, C, which makes indemnity payments I_C according to the function $I_C(X_{zt}, p_C)$, where X_{zt} is a zone-level index and

Figure 2: Examples of indemnity schedules



p_C is a set of parameters specifying the relationship between the index and indemnity payments. The actuarially fair premium can be approximated by $P^{AF}_{C,z} = \frac{1}{T}\sum_{t=1}^{T} I_C(X_{zt}, p_C)$ assuming T is sufficiently large.

For simplicity, consider first an area yield contract which compensates farmers with an amount equal to actual area yield shortfalls below the zone mean. Since we do not have data on prices, we assume that incomes are equal to yields, which implies that all insurance payouts and incomes are denoted in kg/acre. The red solid line in Figure 2 plots the normalized indemnity payment, $\frac{I_{AY}(\bar{y}_{zt}, p_{AY})}{\bar{y}_z}$ against the the normalized index $\frac{\bar{y}_{zt}}{\bar{y}_z}$. Next, consider a satellite based contract. This contract would pay farmers based on predicted zone-level yields, $X_{zt} = \bar{y}^*_{zt}$; however, to allow for easier comparison with the area yield contract, we adjust the indemnity schedule parameters, p_{IS}, such that the actuarially fair premiums are the same: $P^{AF}_{IS,z} = P^{AF}_{AY,z}$. The blue dashed line in figure 2 shows an example of how the indemnity schedule for the satellite contract may be adjusted to ensure that the average indemnity payments are the same between the two contracts.

Finally, consider a satellite based conditional audit contract that pays farmers according to the results of the audit if the actual zone-level yield shortfall is at least higher than 5% of average yields relative to the zone-level yields predicted by the satellite contract. Under such a contract, the indemnity payment would first be calculated according to $I_{IS}(\bar{y}^*_{zt}, p_{IS})$. Then, if farmers request an audit, which they would do if and only if $\frac{\bar{y}_{zt}}{\bar{y}_z} - min(\frac{\bar{y}^*_{zt}}{\bar{y}_z}, 1) < 0.05$ assuming farmers have perfect knowledge about zone level yields, the indemnity payment would equal $\bar{y}_{zt}$. Finally, in order to ensure that the premium is the same as that of the area yield contract, the indemnity payment is adjusted downwards by a constant factor[1].

The three contracts described here are the basis for the analysis in section 4.

3 Background and Data

In order to create an index which minimizes basis risk and to analyze how the proposed contracts would affect the welfare of individual farmers, we need access to historic plot level yields and satellite index data for the same plots. Ideally, we would rely on individual level yield data from longitudinal household surveys; however, to our knowledge, there are no such surveys that cover more than 3-4 years in East Africa. Moreover, none of these existing household surveys include information on the location of plots, which makes it difficult if not impossible to link the high-resolution satellite data to individual farmer yields. Instead, we implemented a retrospective yield survey among smallholder farmers to gather historic data on yields and locations of the corresponding plots. While this approach might be prone to recall error, we feel confident that the data gathered are quite accurate for various reasons. First, we collected data on rice, which is a major cash crop in the study area, and most farmers we interviewed told us that they kept records of historic yields. Second, the interviewers would refer to major events, such as elections and soccer World Cups, to help farmers remember specific years. Finally, farmers were specifically told that there was no penalty for not remembering data. The interview was on purpose made short to keep the focus on the yield data, and we gathered data on rice yields, fertilizer use, acreage, planting/harvest times, and the occurrence and severity of extreme weather events, for the years 2003-2012. Moreover, we asked farmers to indicate on a detailed satellite map the approximate location of their plot(s). A copy of the actual survey is included in Appendix 1.

Our study area consists of four wards located East of the Pare Mountains in the Same district in the Kilimanjaro region of Tanzania (see Appendix 2 for a map of the study area), and covers an area of approximately 20 by 5 kilometers. The main crops grown are paddy (rice) and maize, and rice fields are clustered together with little other vegetation or crops. This makes rice the ideal crop for satellite based index insurance, since pixels will suffer from little contamination from non-rice vegetation. The rice clusters are also used as the basis for the insurance zones, since the

[1]A fourth type of contract was also considered, in which payouts are based on predicted losses for an individual using pixel-level data but only if predicted zone-level yields are below a certain trigger level. The purpose of this contract would be to account for the heterogeneity of weather impacts across a zone while not inducing morally hazardous behavior. However, our analysis of this contract showed that there was little heterogenuity across pixel-specific predicted yields within a zone, resulting in this contract performing nearly identically to the the one using zone-level data.

Table 1: Summary statistics

Variable	Maore N	Maore SE	Maore SW	Ndungu E	Ndungu N	Ndungu S	Ndungu W	Southern Plain N	Southern Plain S	Southern Plain W	Total
Plot size (acres)	1.25	1.49	1.51	0.92	0.96	0.98	1.44	1.18	1.01	0.86	1.18
Land ownership	97%	97%	98%	92%	100%	98%	100%	95%	100%	94%	97%
Irrigation use	84%	58%	87%	100%	94%	97%	75%	100%	100%	97%	89%
Start of season	12-Nov	15-Nov	19-Nov	22-Nov	25-Oct	13-Nov	15-Nov	21-Nov	19-Nov	29-Nov	16-Nov
End of season	20-Mar	27-Mar	1-Apr	17-Apr	26-Mar	15-Mar	15-Apr	14-Apr	9-Apr	7-Apr	31-Mar
Season length (months)	4.3	4.4	4.5	4.9	5.1	4.1	5.1	4.8	4.7	4.3	4.6
Fertilizer use	75%	56%	57%	79%	95%	88%	84%	78%	60%	66%	72%
Rice Yield (kg/acre)	1,653	1,665	1,516	1,721	1,927	1,713	1,630	1,489	1,453	1,545	1,629
Drought Index (1-3)	0.81	1.20	0.68	0.35	0.35	0.36	0.67	0.51	0.75	0.59	0.62
Flood Index (1-3)	0.43	0.37	0.44	0.24	0.18	0.23	0.25	0.61	0.62	0.68	0.42
Experienced drought	37%	56%	33%	18%	18%	20%	38%	27%	33%	30%	30%
Experienced drought	19%	16%	22%	12%	8%	11%	17%	30%	27%	31%	20%
N	37	38	53	12	35	40	4	38	34	32	323

agro-ecological conditions are relatively homogenous within each cluster. In total, we define ten separate insurance zones (see Appendix 2). Given the proximity to the Pare Mountains to the west, most of the water enters the valley through rivers coming down from the mountains. Most of the rice fields are irrigated using canals that link to these rivers. This implies that an index insurance contract based on rainfall from weather stations located in the valley would likely fail to predict drought and flood events. While this area can accommodate up to three growing seasons per year, rice is normally only cultivated once annually, typically between mid-November and mid-March, which covers the short rains period between December and February. While both droughts are floods could impacts crops during any part of the growing season, the rice plants are most vulnerable during the first month and a half following planting, which is also when rains are the most uncertain. There is also significant variation across zones with respect to the drought/flood risk. The three Northern zones are generally most prone to drought as they depend directly on one river coming down from the mountains. The four zones in the Ndungu irrigation scheme face the least risk due to a well-developed canal system. Finally, the three Southern zones, which are located in a flood plain downstream from a major lake, face a smaller drought risk but a high risk of flooding.

Within this study area, we randomly sampled a total of 400 farmers from 10 villages (out of 16) chosen at random from local village lists. Farmers were selected proportional to the population in each village, and were interviewed in a central location in each sub-village. Table 1 shows a set of summary statistics, grouped by each of the ten insurance zones. We note that average yields vary a fair bit across the zones, with the Ndungu irrigation scheme zones having the highest yield, and the Southern flood plains having the lowing average yield. Fertilizer use also varies significantly across zones, but is quite prevalent in this area. Finally, while there is some variation in planting/harvest dates, most zones tend to plant in mid-November and harvest between mid-March and mid-April.

The satellite measures used to create the primary index are based on high-resolution (between 250 and 1000 meter square pixels), high-frequency (daily) remote sensing data, which are publicly available, free of charge, from the NASA website. These data are produced by the MODUS satellite, which started recording data in 2001, and is expected to continue to gather data into the unforeseeable future. In its raw form, the data covers both the infrared and the visual spectrum and capture data in 36 spectral bands ranging in wavelength from $0.4\mu m$ to $14.4\mu m$. Using models that are well-established in the remote sensing literature, these images can be used to create a series of crop health indicators, which have been found to be relatively accurate estimates of the actual measures. In particular, some of the most relevant and commonly used measures of vegetation health are: 1) Normalized Difference Vegetation Index (NDVI); 2) Fraction of Photosynthetic Active Radiation (FPAR); 3) Leaf Area Index (LAI); 4) Evapotranspiration (ET); and 5) Gross Primary Production. While most of these indicators are directly available from the NASA website, converting these data into an index that can be used to predict yields for a particular plot or area requires several steps. In particular, after the index data are either computed using the aforementioned models or downloaded directly from the NASA website, they are adjusted for cloud cover and atmospheric conditions. Then, the daily index data for a season are converted into a single value that can be used to estimate crop yields. This is done by taking the integral of the index data between an assumed planting date and harvest date, but adjusting for the background vegetation at planting. Finally, in order to filter out non-rice pixels and to reduce the impact of "bad" pixels, we apply a crop masking model and a spatial smoothing function. Given the massive data processing and complex modelling required to run the data through these steps, we partnered with the SI Organization, which created a comprehensive model that automatically produces the final index values we need. For a complete description of this work, see Merkovich, 2014.

4 Methodology and Results

In this section, we first develop and estimate a response function mapping the satellite index values to the yield data. This response function is then used as the basis for the satellite based index insurance contracts defined in section 2. We then analyze and compare the the performance of the three different index insurance contracts and present a method for evaluating the welfare implications of each contract.

4.1 Creating a yield response function

In order to minimize the design risk of an index insurance contract, we estimate a response function which maps a satellite index to actual average zone level data. While a non-parametric method would capture the potential non-linearities in the relationship between the index and yields, we have too few zone level data points to get reliable out-of-sample predictions. Instead, we estimate a simple linear model with zone level fixed effects:

$$\bar{y}_{zt} = \alpha_z + \beta * Index_{zt} + \epsilon_{zt} \tag{1}$$

We estimate equation (1) for all five indices, both with and without zone level fixed effects, and as a switching regression with a different intercept and slope for index values below and above the zone mean. Table (2) displays the results for a few select regressions with the best fit. Among the five different indices, the Gross Primary Production index appears to perform the best. This is not unexpected given that GPP is the most direct measure of produced biomass. Moreover, allowing for a different intercept and slope for index values below the zone mean significantly improves the fit. Finally, the zone level fixed effects are jointly significant, which is consistent with the idea that different growing conditions in the different zones affect the relationship between GPP and yield. Comparing the adjusted R^2 of the different models, we find that the model using only GPP as the index, zone level fixed effect, and allowing for a different intercept and slope below the zone mean exhibits the best in-sample fit of the data with an adjusted R^2 of 61.9.

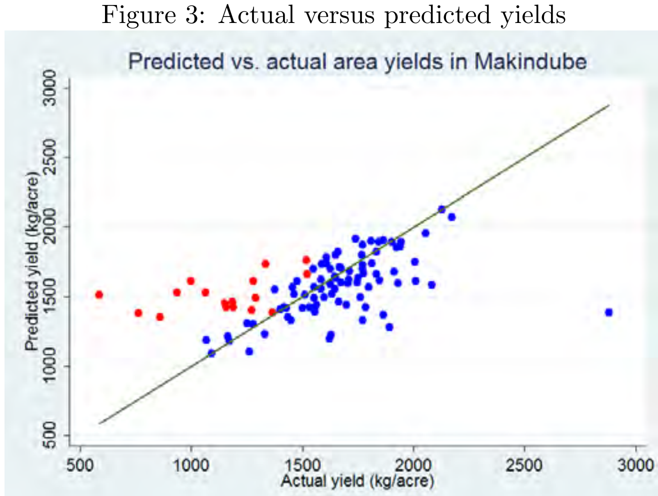
Based on the estimates from this model, we can create predictions of zone level yields. These predictions are then used as the primary index for the satellite based insurance contract. Figure (3) shows a scatter plot of actual zone level yields versus predicted yields with a 45 degree line superimposed. It is clear from the plot that the model is effective at distinguishing good years from bad years. In particular, consider all the zone years for which actual yields were lower than 1,300kg/acre (corresponding to a 20% loss relative to the overall mean). All these data points have predicted yields lower than the overall mean of 1,600kg/acre. Hence, an index insurance contract which pays out for any year for which the predicted yield based on the satellite index is lower than the mean would correctly target years when losses have actually occurred. However, the model is less effective at predicting the severity of losses. In particular, for zone years where actual yields were lower than 1,300kg/acre, there is no correlation between actual yields and predicted yields. While we have no good explanation for this result, we believe it might stem from the imperfect biological relationship between produced biomass and yields under extreme conditions. For example, during a drought, the plant will first reduce the production of grains, while using the available water to maintain the health of the plant itself, thereby resulting in a yield shortfall while the biomass remains the same.

Finally, the red-colored points indicate zone-years for which farmers would request an audit under audit rule described in section 2. Out of a total of 100 zone-year observations, only 17 would result in an audit.

Table 2: Regression results for select yield response functions. Cluster robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Dependent variable is zone level yield	[1]	[2]	[3]	[4]
GPP Index	0.131***	0.166***	0.105**	
	(0.0246)	(0.0231)	(0.0461)	
NDVI Index				117.6***
				(19.14)
GPP<Mean(GPP)			-168.0*	
			(86.37)	
GPP*(GPP<Mean(GPP))			-0.0178	
			(0.0581)	
Constant	1,565***	1,491***	1,575***	1,253***
	(29.30)	(66.36)	(82.94)	(87.50)
Zone Fixed Effects	NO	YES	YES	YES
Observations	70	70	70	70
R-squared	0.294	0.594	0.619	0.537

Figure 3: Actual versus predicted yields



4.2 Relative contract performance

When evaluating the quality of index insurance products, it is useful to first consider objective measures of insurance quality such as actual payouts and the distribution of losses under various contracts. The advantage of this approach is that it does not require any assumptions about how farmers value such payouts or the distribution of final incomes with and without insurance. In the next section, we move beyond objective distributions and estimate the welfare impact of different contracts by assuming an underlying utility function.

We begin by studying the design risk of each of the three contracts described in section 2. By construction, index insurance cannot protect farmers against idiosyncratic shocks; however, as we show in the subsequent section, the magnitude of idiosyncratic risk relative to covariant shocks still affects the value of index insurance to farmers. To study design risk, we assume that all farmers within a zone are identical and that the only risk facing farmers is covariant risk. This is equivalent to saying that farmers only care about zone-level incomes, which would be true under a perfect risk sharing arrangement within the zone. We then calculate zone-level payouts, $I_C(X_{zt}, p_C)$, and resulting incomes, π_{zt}, for the three different contracts and study payouts as a function of actual normalized losses, and the distribution of zone level incomes after insurance relative to no insurance.

Figure (4) shows a scatter plot of zone level insurance payouts, $I_C(X_{zt}, p_C)$, as a function of normalized zone-level yields, $\frac{\bar{y}_{zt}}{\bar{y}_z}$ for the satellite only contract relative to the optimal area yield contract. This plot illustrates clearly the finding from the previous section that while the satellite model is effective at separating good years from bad years, it is less effective at estimating the severity of losses. In particular, 23 out of 25 zone-years with less than 95% of average yields would receive some kind of payout under the satellite only insurance contract. However, out of those 25 zone-years, 10 will receive insurance payouts that do not cover at least 95% of losses. With regards to false positives, out of 28 zone-years with a higher than 105% of average yields, only 3 would falsely receive an insurance payout under this contract.

Next, consider the impact on design risk when farmers have the option to request an audit. Figure (5) shows the same scatter plot for the satellite based conditional audit contract. By the setup of the audit rule, this contract can do nothing about the false positives, but it essentially eliminates the cases when the contract severely underpays losses. During an audit year, the contract still pays less than the area yield contract, but this is a result of the assumption that all the contracts should have the same actuarially fair premium. If instead, the premium of the satellite based conditional audit contract is not adjusted to account for the additional indemnity payouts caused by the audits, the false negatives would be eliminated completely.

Another, and perhaps more useful, way to view these data, is to look at the distributions of final zone-level incomes after indemnity payments are received and premiums are paid. In particular, final incomes, π_{zt}, are calculated as:

Figure 4: Indemnity payments by zone-level yields for satellite only contract relative to area-yield contract

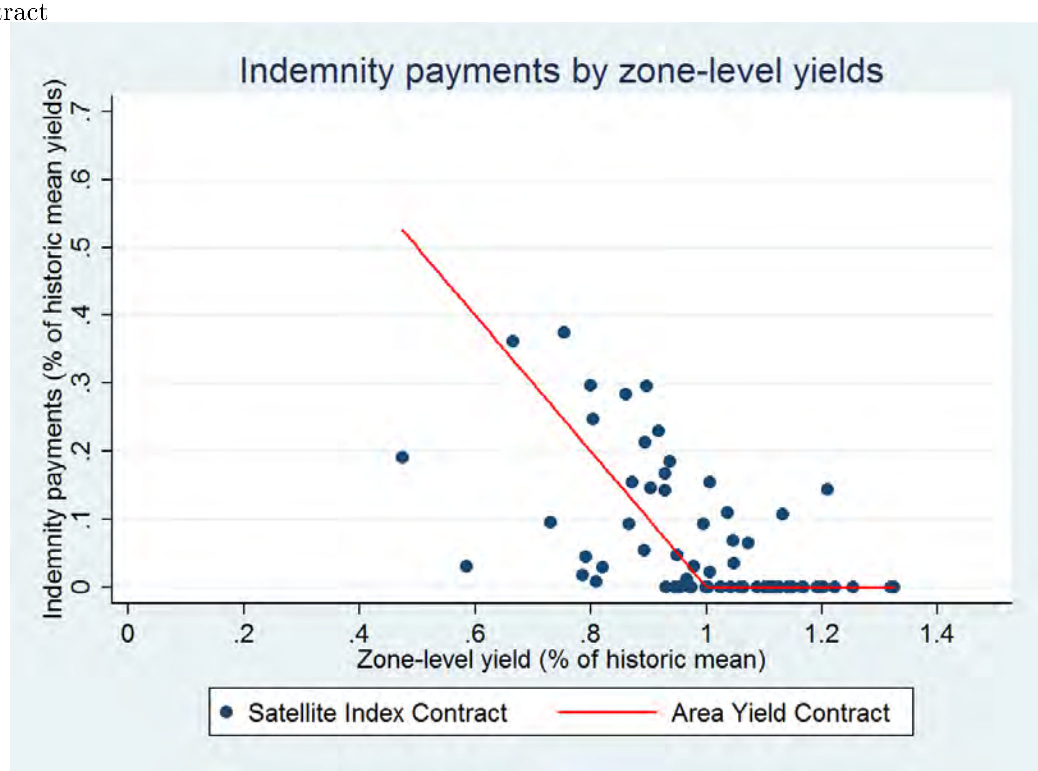


Figure 5: Indemnity payments by zone-level yields for satellite based conditional audit contract relative to area-yield contract

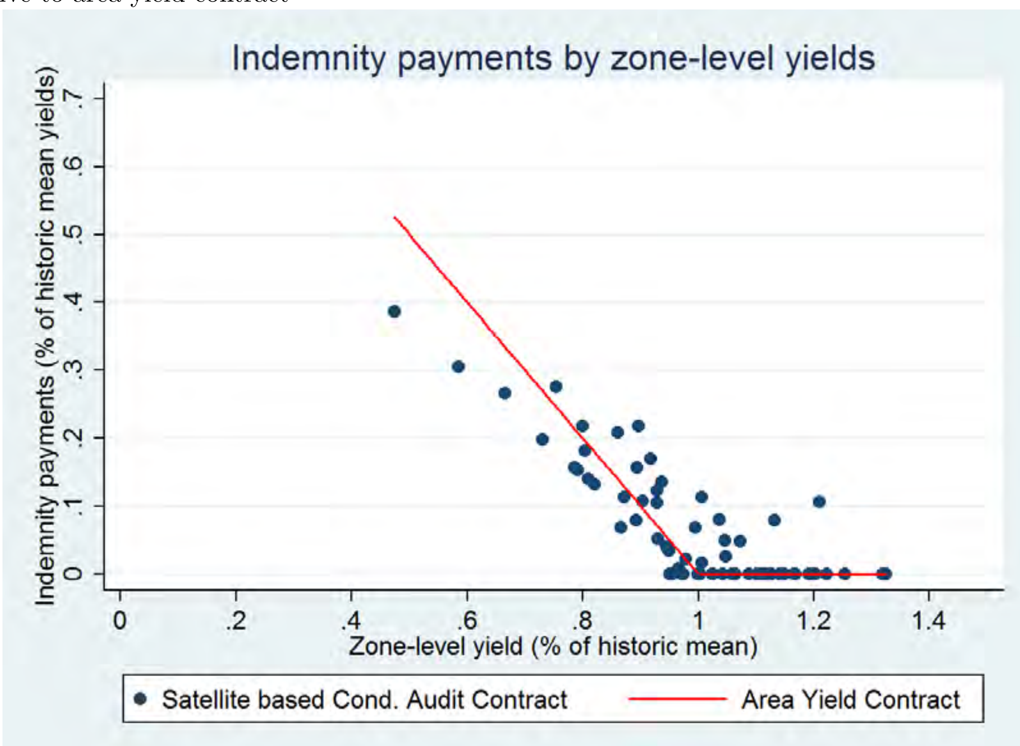
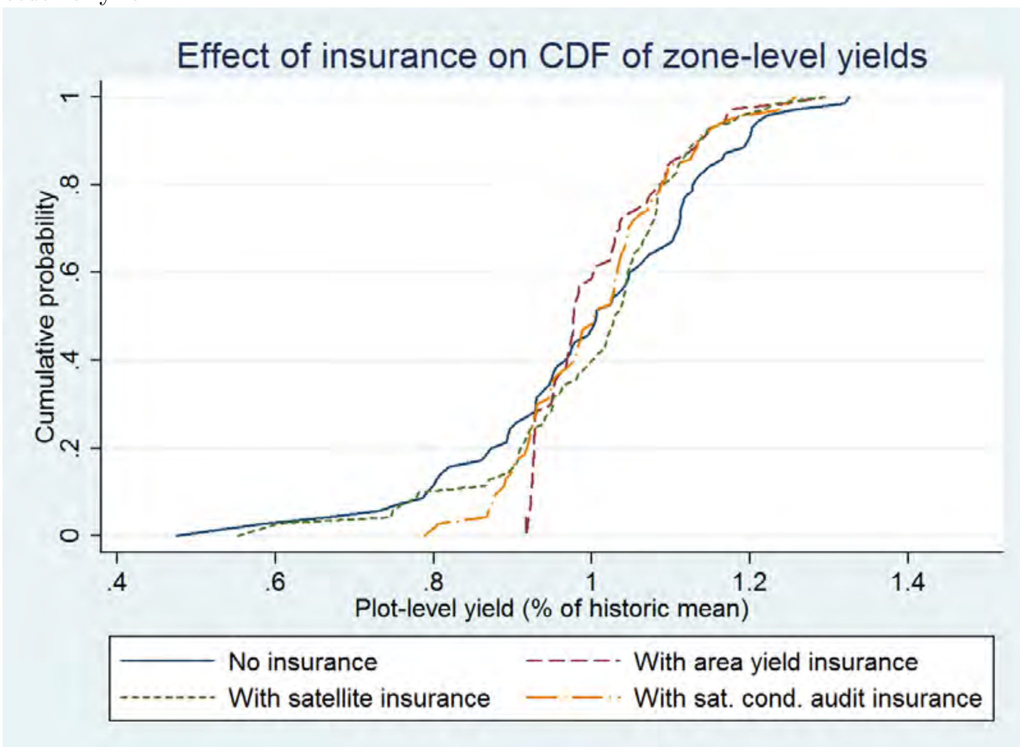


Figure 6: CDF of zone-level yields with and without insurance. Distributions assume that premiums are actuarially fair

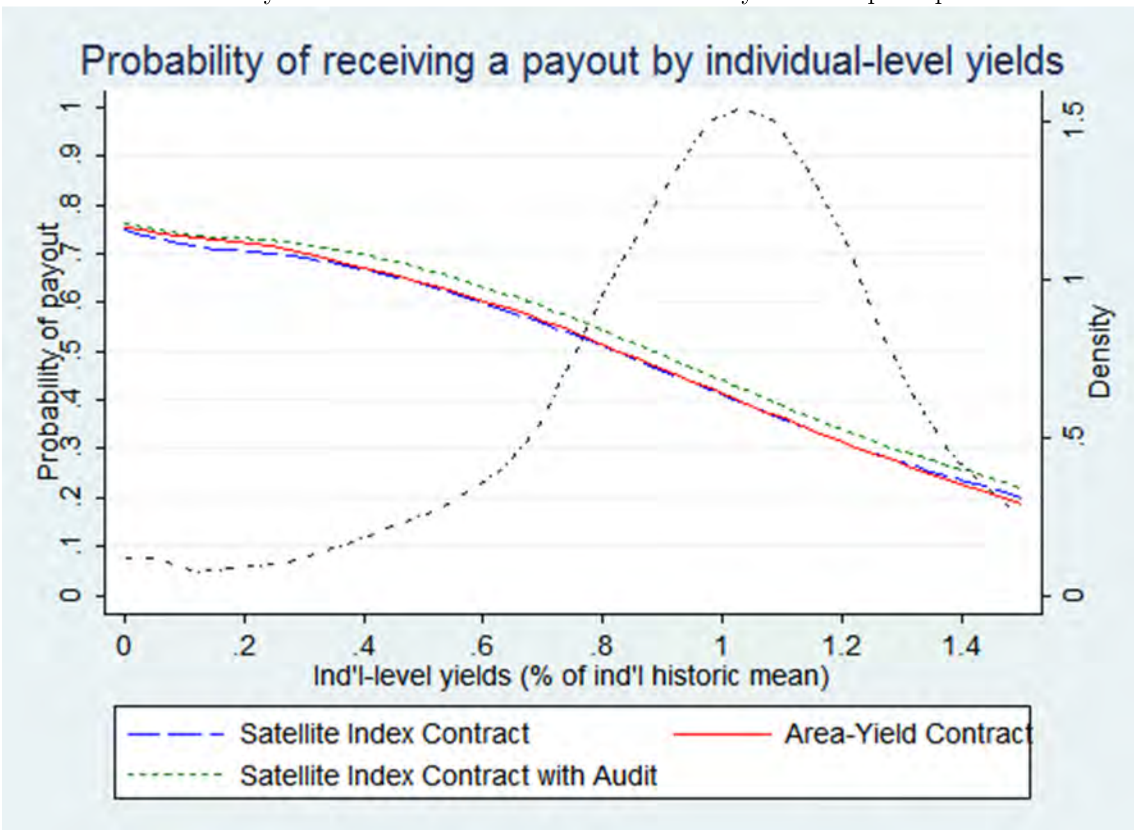


$$\pi_{zt} = \bar{y}_{zt} + I_C(X_{zt}, p_C) - P_{C,z} \tag{2}$$

For ease of comparison, we assume that premiums are all actuarially fair. Figure (6) shows the cumulative distribution function (CDF) of normalized incomes under no insurance, and under each of the three different insurance contracts. Even after premiums are paid, all three contracts offer an improvement over no insurance at the lower end of the tail. In particular, all three contracts second order stochastically dominate the no-insurance option, which implies that if farmers only care about covariant risk (which is a strong assumption), all risk-averse farmers would prefer to buy either of the insurance contracts. Moreover, while the satellite based conditional audit contract still has some design risk, no zone will suffer losses greater than approximately 20% of average yields after insurance payments are made under this contract.

Although index insurance cannot address idiosyncratic risk, it is still informative to study how these contracts affect the total risk faced by individual farmers. For a farmer who is not at all

Figure 7: Probability of receiving a payout by individual-level yields under each of the three contracts. A kernel density function of individual-level normalized yields is superimposed.



exposed to covariant risk, perhaps because she has invested in irrigation, even the "optimal" area yield contract would not provide any benefit. Conversely, a farmer whose total risk is mostly tied to covariant events would greatly value an area yield contract, and may also benefit from a less optimal contract, provided that the design risk is not too large. Figure (7) shows the probability of receiving a payout as a function of normalized individual yields. The function is created by first calculating whether or not each individual in the sample receives an insurance payment in a given year, and then fitting a local polynomial regression to these data. The plot shows that conditional on losses, both the satellite based contracts pay out just as often as the area yield contract. For example, a farmer with a complete yield loss will have a nearly 80% chance of receiving some amount of insurance payout under all three contracts. The plot also demonstrates the relative importance of idiosyncratic risk: during an average year, a given farmer would still have approximately a 40% chance of receiving a payout.

Next, consider how actual indemnity payouts vary with losses under the three contracts. Figure

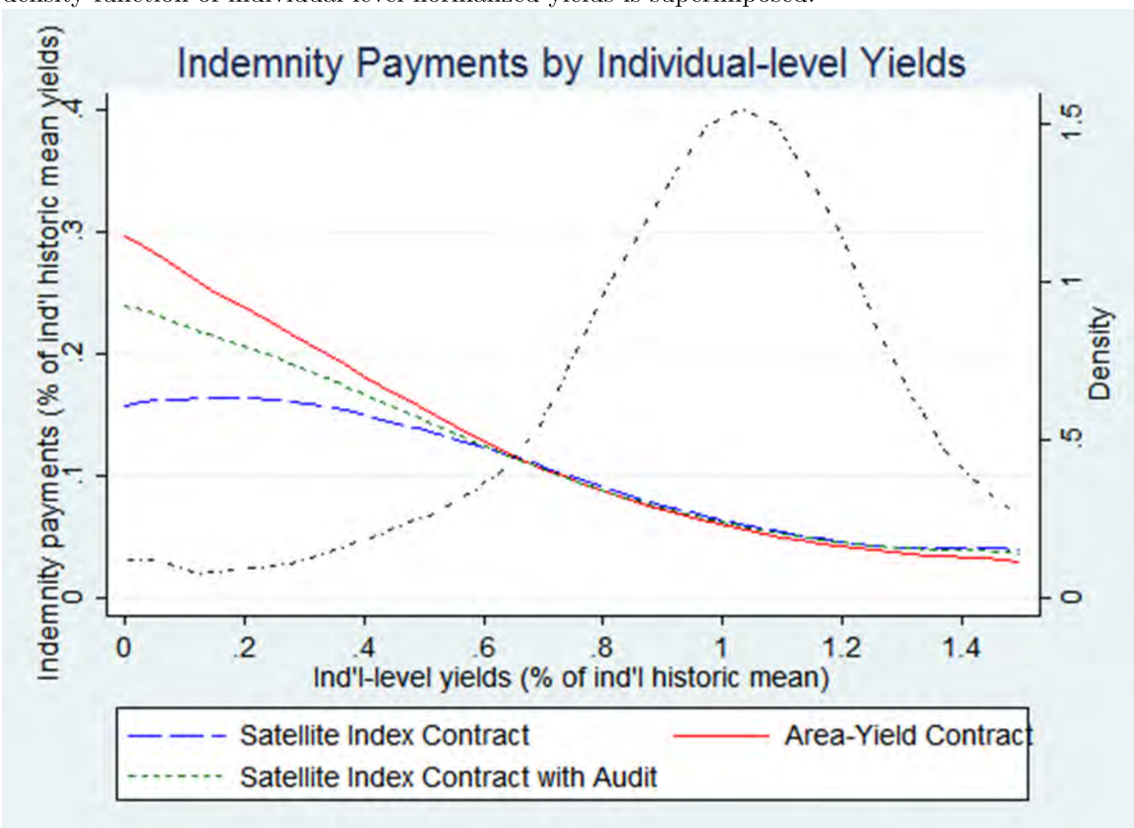
(8) shows the results of fitting a local polynomial regression to individual indemnity payment data (expressed in percent of individual historic mean yields) as a function of normalized individual yields. First, note that the satellite only contract fails to distinguish high-loss events from medium-loss events, as average payouts are roughly the same (17%) under a 60% loss as it is under a 100% loss. The satellite based conditional audit contract corrects this shortcoming, and average payouts are only slightly lower than those made under the area yield contract when individual losses are high. However, due to the high level of idiosyncratic risk, neither contract offers great protection against the total risk faced by farmers. For example, a farmer with a complete yield loss would only receive an average payment of 30% of normal yields under an area yield contract. Under the satellite only contract, the average payment is only 17%, while the satellite based conditional audit contract would pay 24% on average. While these figures seem low, each of these index insurance contracts may still provide valuable benefits to farmers. The next section explores the welfare implications of index insurance, and analyzes whether farmers would be better off under each contract than they would without insurance.

4.3 Welfare effects

While the previous analysis allows for an objective comparison of different contracts using actual yield data, it provides no indication of whether any of these contracts would actually enhance the welfare of individual farmers. As shown in Clarke (forthcoming), even an actuarially fair index insurance contract might be welfare-reducing for a risk-averse individual if the basis risk is sufficiently high. Understanding the welfare implications of index insurance is important because it allows us to determine whether a particular contract should even be marketed to farmers. In particular, if the design risk is sufficiently high, as is the case with many weather based index insurance contracts, or the importance of idiosyncratic risk relative to covariant risk is sufficiently large, index insurance may not improve the welfare of farmers even if they face significant production risk. In this section, we use expected utility analysis to analyze how each contract would affect the welfare of farmers. In particular, we assume a constant relative risk aversion (CRRA) utility function, which is consistent with similar analysis of utility gains from index insurance (Woodard et al, 2012; Clarke, 2014). Using the data from our yield study, we first estimate a model of zone-level yields assuming an underlying Weibull distribution. Specifically, if $\bar{y}_{zt}$ is average yield is zone z in year t, and $\bar{y}_z$ is average yield in zone z across all years, we estimate the parameters a_1, a_2, b_1, b_2 of the Weibull pdf using maximum likelihood estimation:

$$f\left(\bar{y}_z, a, b\right) = \frac{a}{b}\left(\frac{\bar{y}_{zt}}{b}\right)^{a-1} e^{-\left(\frac{\bar{y}_{zt}}{b}\right)^a} \tag{3}$$

Figure 8: Indemnity payments by individual-level yields under each of the three contracts. A kernel density function of individual-level normalized yields is superimposed.



where:

$$a = a_1 + a_2 * \bar{y}_z \tag{4}$$

$$b = b_1 + b_2 * \bar{y}_z \tag{5}$$

Based on these estimated parameters, we can simulate draws of $\bar{y}^s_{zt}$ for each of the ten insurance zones. To calculate insurance payouts under the area yield contract, we compute $I_{AY}(\bar{y}^s_{zt}, p_{AY})$ as defined in section 2. For the satellite based contracts, in order to simulate insurance payouts, we first need to create a response function linking $\bar{y}^s_{zt}$ to an index value, or predicted area yield, $\bar{y}^{p;s}_{zt}$. To do that, we estimate a switching regression allowing for a different intercept and slope if average yields are below the zone mean:

$$\bar{y}^p_{zt} = (\gamma_{1zt} + d\gamma_{2zt}) + (\delta_{1zt} + d\delta_{2zt})\,\bar{y}_{zt} + \epsilon_{zt} \tag{6}$$

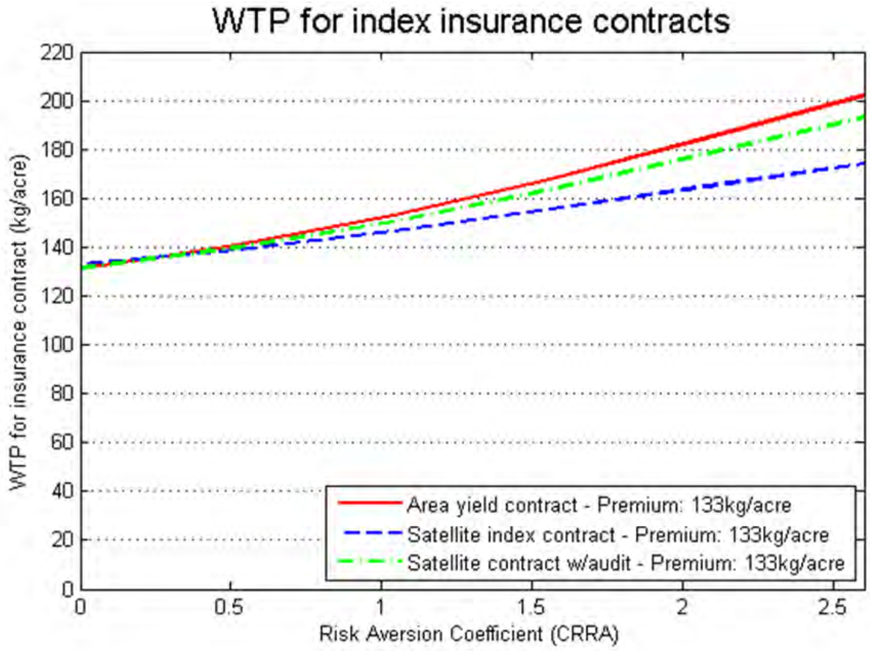
where d is a dummy which equals one if $\bar{y}_{zt} < \bar{y}_z$, and zero otherwise. Now, $\bar{y}^{p;s}_{zt}$ are simulated as random draws from the estimated model, and indemnity payments for the satellite based contract and the satellite based conditional audit contract can be computed as $I_{SI}(\bar{y}^{p;s}_{zt}, p_{SI})$ and $I_{SA}(\bar{y}^{p;s}_{zt}, p_{SA})$, respectively. To calculate actuarially fair insurance premiums for the different contracts, we compute $P^{AF}_{C,z}$ =for $C \in (AI, SI, SA)$ as defined in section 2. Also, as described in section 2, the parameters of the indemnity schedule for the satellite based contracts are adjusted such that their premiums are equal to that of the area yield contract.

Next, to simulate individual yields, we first estimate the following model based on Miranda (1991) as a random intercepts and coefficients model:

$$y_{izt} = \alpha_{iz} + \beta_{iz}\,(\bar{y}_z - \bar{y}_{zt}) + \epsilon_{izt} \tag{7}$$

We assume that α_{iz} and β_{iz} are distributed jointly normal and that ϵ_{izt} are independent and normally distributed. Using the coefficients and the variance-covariance matrix of the random effects parameters from estimating equation (7), we can simulate individual yields y^{sim}_{izt} by making draws from the resulting distributions.

Figure 9: WTP for index insurance contracts



Using the simulated individual yields and insurance payouts under the three different contracts, we calculate the expected utility of an individual, both with and without insurance:

$$EU_{iz,NI} = \sum_{t=1}^{T} U\left(w + y_{izt}^{sim}\right) \tag{8}$$

$$EU_{iz,C} = \sum_{t=1}^{T} U\left(w + y_{izt}^{sim} + I_{C,zt} - P_{C,z}\right) \tag{9}$$

where $U(z_{izt}) = \frac{1}{1-\gamma}(z_{izt})^{1-\gamma}$, γ is the coefficient of relative risk aversion, and$w$ is initial wealth, which we assume to be constant across individuals and time, and equal to the lowest value such that final income is still positive for all individuals after insurance premiums are paid. Now, in order to estimate the welfare impact of index insurance, we calculate an individual's willingness to pay (WTP) for each insurance contract under a range of risk preferences. In particular, we solve for an individual's premium $P_{C,iz}$ such that $EU_{iz,NI} = EU_{iz,C}$, and compute the mean WTP across all individuals.

Figure (9) shows the average WTP among farmers for the three different contracts over a reasonable range of risk preferences (0-2.5). As expected, when individuals are risk neutral, the WTP simply equals the actuarially fair premium of 133kg/acre. As farmers become more risk averse, the benefit of all the insurance contracts increases, which is evidence that if priced at actuarially fair rates, each of the insurance contracts improves the welfare of the average farmer. Moreover, the area yield contract provides the greatest value to farmers at all levels of risk aversion, which is not surprising given that this contract has no design risk, by construction. Finally, it is worth noting that farmers' WTP for the satellite based conditional audit contract is only marginally less than that of the area yield contract. In particular, for an individual with a CRRA of 1, the WTP for the area yield contract is 14% over the actuarially fair price, versus a 12.5% premium for the satellite based conditional audit contract. Given that the latter contract would likely be significantly cheaper to provide, this might in fact be a more viable contract than an area yield contract.

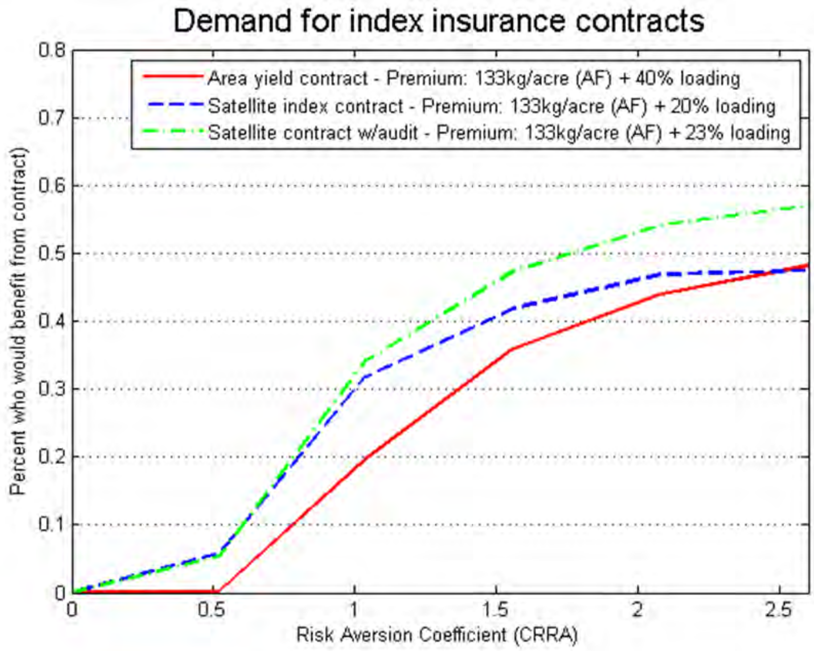
The above analysis only considers the average WTP across all farmers. However, given that there is significant heterogeneity between farmers in how closely their yields are correlated with area yields (as measured by β_{iz} in equation (7)), there will also be heterogeneity in farmers' WTP for index insurance. To estimate the demand for each of the three index insurance contracts, we therefore calculate the percentage of farmers in the simulated population that would benefit from the insurance, given certain assumptions about loading costs. In particular, we assume that all three contracts have a base loading cost of 20% above actuarially fair prices. Moreover, complete verification of area yields in all years would add an extra 20% premium, resulting in a loading cost of 40% for the area yield contract. The satellite based conditional audit contract, however, only requires verification of area yields 17% of the years, which assuming a linear cost structure, would result in a premium of 23% above the actuarially fair price for this contract. Using these premiums, we calculate the proportion of farmers benefiting from insurance as:

$$\frac{1}{N}\sum_{i=1}^{N} 1\left(EU_{iz,C} > EU_{iz,NI}\right) \tag{10}$$

where 1() is an indicator function which equals one if the expression is true, and zero otherwise, and N is the total number of simulated individuals.

Figure (10) displays the proportion of individuals in the simulated sample who would benefit from each of the three different index insurance contracts under a range of risk preferences. While an area yield contract commands the highest WTP, the demand for such a contract will likely be low due to the high data collection costs. Both the pure satellite based contract and satellite based conditional audit contract would have approximately the same demand for risk aversion coefficients

Figure 10: Demand for index insurance contracts



less than 1; however, more risk averse farmers would prefer the satellite based conditional audit contract, as it provides better coverage at a slightly higher premium. For farmers with a CRRA of 1, demand would be 18%, 32% and 34%, under the area yield contract, the satellite based contract, and the satellite based conditional audit contract, respectively. These findings are obviously dependent on the loading cost assumptions; however, these assumptions are consistent with the premiums charged by many microinsurance institutions in developing countries.

5 Conclusion and Discussion

While index insurance may in theory offer a promising solution to the problem of covariant risk in smallholder agriculture, its impact has so far been limited in part due to the poor quality of many index insurance contracts. In particular, the correlation between the index and farmer losses has often been close to zero, implying that farmers are purchasing a lottery ticket rather than actual insurance. While the impact of insurance quality (basis risk) has been studied extensively in the literature on index insurance in a developed country context, very little research has focused on the impact of poor insurance quality in developing countries, and even fewer make use of plot level panel data to analyze the welfare implications of index insurance on a household level.

In this paper, we use a unique panel dataset of plot-level rice yields in Tanzania to study the welfare implications of an innovative index insurance contract that uses satellite data as the primary index but contains a provision that allows famers to request an audit if the primary index fails to predict yields. In particular, we compare the performance of such a contract to an "optimal" area yield contract, which by construction has no design risk. We then use expected utility analysis to estimate farmers' willingness to pay (WTP) and demand for an area yield contract, a pure satellite contract, and the satellite based conditional audit contract, under a range of risk preferences. Our results demonstrate that by combining a primary satellite index with the possibility of an audit, we can almost completely eliminate design risk at a very modest cost. In particular, under this contract, we estimate that audits would be requested only 17% of the time, which implies that audit costs would be approximately 1/6th of those incurred under an area yield contract, which would require an audit for every season. Moreover, we show that farmers' WTP for this contract exceeds the actuarially fair price by approximately 12.5% for a farmer with a CRRA of 1, and that demand might be higher than 30% even under unsubsidized commercial rates.

These findings may have important implications for the future of index insurance in developing countries. Since most index insurance products in the developing world today are based on precipitation measures, which have been shown to correlate poorly with actual farmer losses, the contract we propose may offer a significant improvement over these contracts and might significantly improve the welfare of farmers by more accurately insuring against covariate risk. However, further research is needed to test the viability of this contract in a real-world setting, and a next step would be conduct a small-scale impact evaluation by offering this contract to farmers though an insurance company.

Finally, it is important to consider some of the limitations of these results. For example, the satellite based yield response function has only been estimated for rice, which is clustered together with little contamination from other vegetation or non-rice crops. If fields are more scattered, as is the case with maize and sunflower and several other crops in developing countries, the correlation between the satellite based measures and yields might be lower. Also, this contract would be most appropriate in data scarce environments, such as African small-scale agriculture, and should not be seen as a replacement for an area yield contract in places where area yield data are already available at a low cost.

References

[1] Barhane, Guush; Clarke, Daniel; Dercon, Stefan; Vargas Hill, Ruth; Taffesse, Alemayehu S.; "Insuring against the weather: Addressing the Challenges of Basis Risk in Index Insurance using Gap Insurance in Ethiopea", BASIS I4 Update, April, 2013

[2] Bhattacharya BK, Mallick K, Nigam R, Dakore K, Shekh AM (2011). Efficiency based wheat yield prediction in semi-arid climate using surface energy budgeting with satellite observations. Agric Forest Meteorol 2011;151:1394–408.

[3] Carter, Michael (2011). "Innovations for Managing Basis Risk under Index Insurance for Small Farm Agriculture". FERDI memo

[4] Carter, Michael R., Lan Cheng, and Alexander Sarris (2011). "The Impact of Interlinked Index Insurance and Credit Contracts on Financial Market Deepening and Small Farm Productivity," unpublished manuscript.

[5] Carter, Michael R., and Rachid Laajaj (2009). "Using Satellite Imagery as the Basis for Index Insurance Contracts in West Africa," http://i4.ucdavis.edu/projects/contracts/files/laajaj-using-satelliteimagery.pdf

[6] Clarke, Daniel (forthcoming) "A theory of Rational Demand for Index Insurance", forthcoming in the American Economics Journal.

[7] Clarke, Daniel; Mahul, Oliver; Rao, Kolli N.; Verma, Niraj; "Weather Based Crop Insurance in India", Policy Research Working Paper 5985, March 2012

[8] Clarke, Daniel (2010). "A Theory of Rational Hedging". Mimeo. Oxford University.

[9] Cole, S., X. Giné, et al. (2010) "Barriers to Household Risk Management: Evidence from India." Harvard Business School Working Paper.

[10] De Bock, Ombeline, Michael Carter, Catherine Guirkinger, Rachid Laajaj (2010) "Feasibility Study: Which micro-insurance mechanisms are most beneficial to cotton growers in Mali?" (english translation). Centre de Recherche en Economie du Développement, Universitaires Notre-Dame de la Paix, Namur, Belgium.

[11] EARS 2012, Environmental Analysis and Remote Sensing, Monitoring and Early Warning Validation Results: http://www.earlywarning.nl/frames/Frame_val.htm (crop yield in Africa link)

[12] Gine, Xavier, Lev Menand, Robert Townsend and James Vickery (2010). "Microinsurance, A Case Study of the Indian Rainfall Index Insurance Market", World Bank Policy Research Working Paper no. 5459. 6

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

[14] Jensen, Nathaniel D.; Barrett, Christopher B.; Mude, Andrew G. (2014), "Basis Risk and the welfare gains from index insurance: Evidence from Norrthern Kenya", Working paper

[15] Martyniak, Ludwika (2007). "Response of Spring Cereals to a Deficit of Atmospheric Precipitation in the Particular Stages of Plant Growth and Development." Agriculture Water Management; 95 (3): 171-178.

[16] Merkovich, Ryan (2014), "Tanzania Crop Yield Study", internal report, available upon request

[17] Miranda, Mario J., "Area-Yield Crop Insurance Reconsidered," American Journal of Agricultural Economics, 73 (1991), 233-242.

[18] Prince, S. D., E. Brown de Coltound, and L. L. Kravitz (1998). "Evidence from rain use efficiencies does not indicate extensive Sahelian desertification." Global Change Biology; 4: 359-374.

[19] Rosema A, 1993, Using METEOSAT for operational evapotranspiration and biomass monitoring in the Sahel region Remote Sens. Environ., 46 (1993), pp. 27–44

[20] Smith, Vince and Myles Watts (2009), "Index Based Agricultural Insurance in Developing Countries: Feasibility, Scalability and Sustainability," 2009.

[21] Sims, P. L., and J. S. Singh (1978). "The Structure and Function of Ten Western North American Grasslands: Intra-seasonal Dynamics and Primary Producer Compartments." Journal of Ecology; 66: 547-572.

[22] Staggenborg, S. A., K. C. Dhuyvettere, and W. B. Gordon (2008). "Grain Sorghum and Corn Comparisons: Yield, Economic and Environmental Responses." Agricultural Experiment Station Working Paper, Kansas State University, Manhattan, Kansas.

[23] Woodard, Joshua D., Alexander D. Pavlista, Gary D. Schnitkey, Paul A. Burgener, and Kimberley A. Ward, "Government Insurance Program Design, Incentive Effects, and Technology Adoption: The Case of Skip-Row Crop Insurance," American Journal of Agricultural Economics, 94 (2012), 823-837.

6 Appendix

6.1 Appendix 1: Survey instrument

Figure 11: Survey recording sheet

Gathering historic yield data from paddy farmers in Makindube ADP, Same District, Tanzania

Plot/Farmer specific information

Enumerator number: Farmer Number: Plot number:

Enumerator name:

Date of interview:

Name of ward:

Name of village:

Name of subvillage:

Name of area where plot is located:

Plot size: acres

Rent/Own? (1=own; 2=rent) Years owned/rented same plot: years (if more than 9, write 9)

Is plot irrigated by canals? (1=yes; 2=no)

Size bag used for paddy? kg

Seasons: Start Vuli: End Vuli:

Yield codes

X. Don't know/remember
NP. Not planted
20+: No less than 20 bags
20-: No more than 20 bags
20-30: 20-30 bags

Drought/Flood Severity codes

1. Mild (drought/flood)
2. Moderate (drought/flood)
3. Severe (drought/flood)

Yield History:

			Yield					
Year of harvest	Season	Used fertilizer (1=Y; 2=N)	Total Yield (bags)	Acres planted	Yield/acre (bags/acre)	Drought (see codes above)	Flood (see codes above)	Comments (e.g. duration of flood/drought etc.)
2012	Masika							
2012	Vuli							
2011	Irrigation							
2011	Masika							
2011	Vuli							
2010	Irrigation							
2010	Masika							
2010	Vuli							
2009	Irrigation							
2009	Masika							
2009	Vuli							
2008	Irrigation							
2008	Masika							
2008	Vuli							
2007	Irrigation							
2007	Masika							
2007	Vuli							
2006	Irrigation							
2006	Masika							
2006	Vuli							
2005	Irrigation							
2005	Masika							
2005	Vuli							
2004	Irrigation							
2004	Masika							
2004	Vuli							
2003	Irrigation							
2003	Masika							
2003	Vuli							

6.2 Appendix 2:

Figure 12: Map of study area