Improving the quality of index insurance with a satellite-based conditional audit contract *

Jon Einar Flatnes[†], Michael R. Carter[‡], Ryan Mercovich[§]

6/15/2018

Abstract

While index insurance offers a compelling solution to the problem of covariant risk among smallholder farmers in developing countries, most weather-based contracts have suffered from poor quality. This paper demonstrates that a contract based on satellite data combined with a second-stage conditional audit has the potential to improve index insurance quality. We develop a welfare-based metric for quantifying index insurance quality and use this to identify an optimal audit trigger. Using plot-level panel data of 323 smallholder rice farmers in Tanzania, we estimate a hypothetical willingness to pay for a set of contracts. Our simulation results indicate that demand for the conditional audit contract is 36% under reasonable assumptions, while demand for a satellite-only contract is 22%. Overall, these results indicate that in data-scarce environments where area-yield contracts are infeasible, a satellite-based conditional audit contract may be superior to standard index insurance contracts.

Keywords: [Index Insurance, Basis Risk, Remote Sensing]

JEL Classification Codes: [G22, O16, Q14]

[†]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

[§]Ph.D., Ramerco, LLC, 19 Morgan Street Phoenixville, PA, 19460; email: ryan@ramer.co

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, and relaxing risk constraints can have large positive impacts on investments (Karlan et. al., 2014), consumption smoothing (Kazianga and Udry, 2006), and credit rationing (Boucher et al., 2008). Particularly small-scale farmers are plagued by risk, as weather variation is the largest source of risk in agriculture (Cole et al., 2013; Dercon and Christiansen, 2011) and because such risk is spatially correlated, making local risk sharing mechanisms ineffective (Skees and Barnett, 1999). Index insurance offers a compelling solution to the problem of covariant risk and avoids problems of moral hazard and adverse selection that plagues conventional indemnity insurance contracts (Stiglitz, 1981). However, with a few notable exceptions (e.g. Karlan et al., 2014), these products have generally suffered from low demand, in large part due to the problem of basis risk. This refers to the portion of the risk that is uninsured by the contract and manifests itself as a lack of a payout despite the farmer experiencing a loss. While index insurance contracts by construction has some level of basis risk due to idiosyncratic losses within the insurance zone, the majority of index insurance contracts also fail to compensate farmers for covarient losses (design risk). In particular, precipitation-based contracts have been shown to carry a high level of design risk, to the point where some contracts resemble lottery tickets rather than insurance. For example, 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 indemnity payment made was only 12%. On the other hand, area-yield contracts, which are based on mean losses across an insurance zone, can in theory eliminate design risk; however, such contracts are typically infeasible due to lack of reliable yield data.

The effect of basis risk on farmer welfare and demand can be large. For example, Clarke (2016) develops a theoretical model demonstrating that risk-averse farmers will optimally decline index insurance coverage if basis risk is sufficiently high. 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. Moreover, Mobarak and Rosenzweig (2012) use an randomized experiment among rural Indian households to show that basis risk reduces demand for index insurance but that this effect is partially mitigated by informal risk sharing.

Despite the potential welfare gains from improving index insurance quality, surprisingly few studies have discussed contractual methods to mitigate basis risk. This paper attempts to fill this gap by proposing and analyzing an index insurance contract that specifically minimizes design risk. In particular, we design a contract that uses inexpensive satellite data as the primary index and includes a provision which allows farmers to request an audit if the primary index fails to predict area losses. The satellite index is constructed by creating a response function which correlates

actual reported yields with publicly available satellite-based measures of vegetation and biomass. The purpose of the audit provision is to eliminate large negative basis risk events; however, since audits are costly, there is a tradeoff between correcting uncompensated losses and the increased premium due to the audit cost. In addition, we develop a welfare-based measure of design risk that expresses the willingness to pay (WTP) to avoid design risk as a fraction of the WTP (beyond the premium) for perfect insurance. Using a unique dataset of historical plot-level yields for rice farmers in a concentrated area of Northeastern Tanzania, we are thus able to quantity the level of design risk and compare the relative quality of different insurance contracts. This also allows us to find the optimal audit trigger depending on the audit cost. Finally, we estimate the hypothetical WTP for an area-yield contract, a satellite-only contract, and the satellite-based conditional audit contract and use these measures to estimate the potential demand for each product under reasonable loading cost assumptions.

Our results show that the audit provision has the potential to reduce our welfare-based design risk measure by more than half, assuming the audit trigger is optimized given the audit cost. Furthermore, the hypothetical WTP for both a satellite-only contract and the satellite-based conditional audit contract exceeds the actuarially fair premium (AFP) for the average risk-averse farmer. In particular, assuming a CRRA of 1.5, the average farmer is willing to pay 12% above the AFP for the satellite-only contract versus 22% when adding the audit provision. Finally, we show that demand for the satellite-only contract would exceed that of a feasible area-yield contract, which fetches an additional premium equal to conducting an audit every season (thus approximating the cost of collecting annual yield data). When adding the conditional audit provision, demand increases further despite the increase in cost due to audits. This result is driven by the fact that under the conditional audit contract, audits are only conducted when they result in a net welfare gain for farmers, which is when the severity of a basis risk event is sufficiently large to justify the cost of the audit. In particular, for farmers with a CRRA of 1.5, demand would be 16%, 27% and 36%, under the area-yield contract, the satellite-only contract, and the satellite-based conditional audit contract, respectively.

Our paper makes four important contributions to the literature on index insurance. First, to our knowledge, it is the first paper to analyze the welfare implications of an insurance contract that combines a standard index with a second-stage audit. Second, it is one of few papers that uses a satellite-based measure of yields as the primary index. Third, it develops a welfare-based, normalized measure of design risk that can be used to quantity the quality of any index insurance contract. As such, this measure can also be used to set minimum standards for index insurance quality. Finally, this paper is also one of the first to analyze the impact of basis risk on the welfare of individual farmers using plot-level panel data. Only a handful other papers (e.g. Jensen et al., 2018; Mubarak and Rosenzweig, 2012) study basis risk on a household level.

The rest of the paper is structured as follows. Section 2 discusses the problem of basis risk

in index insurance and develops the index insurance quality metrics used to quantity design risk. Section 3 describes the data and the study area, while section 4 estimates a satellite-based yield model and compares the performance of the index insurance contracts using data aggregated at the insurance zone level. In section 5, we move beyond zone-level data and study how the contracts compare when considering individual-level data. Section 6 concludes.

2   The problem of basis risk

Insurance contracts are designed to maximize the welfare of farmers by smoothing income fluctuations at the lowest possible cost. In theory, traditional individual indemnity contracts provide perfect coverage against actual losses. However, they are fraught with problems of moral hazard and adverse selection. Extensive monitoring and verification can 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, which 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 not be effective at smoothing income for individual farmers. In the finance literature, the residual level of risk faced by the farmer is referred to as basis risk, which, if sufficiently high, can significantly reduce the value of index insurance for farmers, even to the point where it is welfare-reducing (Clarke, 2016). The magnitude of the idiosyncratic risk depends on the geographical scale of the insurance zone and the homogeneity of yield outcomes within a zone. While smaller zones tend to reduce the level of basis risk faced by farmers, moral hazard and adverse selection limit the ability to shrink idiosyncratic risk by downscaling the contract (Elabed and Carter, 2015).

2.1   Basis risk under area-yield contracts

To understand the effect of basis risk, consider the total risk that an individual farmer faces. Following Miranda (1991), we can decompose a farmer's individual yield y_{izt} for farmer i residing in zone z in year t into a covariant component and an idiosyncratic component:

$$y_{izt} = \mu_{iz} + \beta_{iz} \left(\bar{y}_{zt} - \mu_z \right) + \epsilon_{izt} \tag{1}$$

4

Here, $(\bar{y}_{zt} - \mu_z)$ is the zone-level deviation from the mean yield in year t and represents how good (or bad) the year has been in a particular zone, relative to the long-term mean. β_{iz} measures how well the farmer's yield tracks covariant losses, while ϵ_{izt} is an idiosyncratic shock specific to an individual farmer. To simplify the notation, define $y_{izt}^* = \frac{y_{izt} - \mu_{iz}}{\mu_z}$, $\bar{y}_{zt}^* = \frac{\bar{y}_{zt}}{\mu_z} - 1$, and $\epsilon_{izt}^* = \frac{\epsilon_{izt}}{\mu_z}$. We can then write equation 1 as:

$$y_{izt}^* = \beta_{iz}\bar{y}_{zt}^* + \epsilon_{izt}^* \tag{2}$$

Equation 2 expresses the normalized deviation of an individual farmer's yield from her long-term average as a sum of a covariant component and an idiosyncratic shock. The variance of ϵ_{izt}^*, σ_ϵ^2, measures the level of idiosyncratic risk faced by the farmer. If β_{iz} is equal to zero, an individual farmer's yield is completely uncorrelated with zone-level yields. Moreover, if $\beta_{iz} = 1$ and $\sigma_\epsilon^2 = 0$, the farmer's yields perfectly tracks those of the zone. As demonstrated by Miranda (1991), index insurance may only be effective if β_{iz} is sufficiently large. Specifically, for farmers with betas below some critical value, even an index insurance contract that perfectly tracks area yields will not be effective and may actually increase the total risk faced by the farmer.

However, most index insurance contracts offered to farmers in developing countries are not based on area yields and thus do not perfectly track covariant losses. This imperfect correlation introduces another source of basis risk, which is referred to as design risk. Design risk arises due to the inability of the index to accurately measure covariant losses within a defined insurance zone and manifests itself as a lack of insurance payouts despite area yields being low. Assuming farmers face no price risk and that average yields within a zone can be accurately measured, an index insurance contract based on area yields will, by construction, have no design risk. In this case, the total basis risk will be equal to idiosyncratic risk. Hence, an ideal 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, the large majority of index insurance contracts are based on rainfall, which has been shown to be a rather poor predictor of area yields. The following section defines a set of metrics for quantifying design risk in index insurance contracts.

2.2   Quantifying design risk and insurance quality

To understand how design risk affects farmers, we decompose $\bar{y}_{zt}^*$ into a predictable and an unpredictable component. Following Carter et al. (2014), we write:

$$\bar{y}_{zt}^* = f^*(S_{zt}) + \nu_{zt}^* \tag{3}$$

Here S_{zt} is an index, such as rainfall, satellite-predicted yields, or direct measures of average

zone yields, while f^* is a function that maps the index value into normalized zone-level yields deviations, and ν_{zt}^* is the design error. We can now write equation 2 as:

$$y_{izt}^* = \beta_{iz} f(S_{zt}) + \beta_{iz} \nu_{zt} + \epsilon_{izt}^* \tag{4}$$

This equation decomposes the total risk faced by the farmer into an insurable component (the first term), an uninsurable design risk component (the second term), and an uninsurable idiosyncratic risk component (the third term). The purpose of this section is to develop a methodology for quantifying the level of design risk for a given contract.

The simplest measure of design risk is the fraction of the total variance explained by the model, or the R^2 of the regression in 3:

$$D_1 = 1 - \frac{\sigma_\nu^2}{Var(y_{zt}^*)}$$

where σ_ν^2 is the variance of ν_{zt}^*. While the R^2 is a useful and commonly reported measure of design risk, it is less effective at quantifying insurance quality. For example, insurance quality depends primarily on the model's ability to predict losses rather than above-average yields. Moreover, while false negatives result in a non-payment during a bad year, false positives increase the insurance premium, which each have vastly different effects on insurance quality. Hence, to create a more accurate measure of insurance quality, we define a normalized measure of uninsured design risk. First, define an insurance contract C, which makes indemnity payments I_C according to the function $I_C(X)$, where X is a normalized yield measure. Hence, the indemnity payout is $I_{AY}(\bar{y}_{zt}^*)$ for an area-yield contract and $I_{II}(f^*(S_{zt}))$ for an index insurance contract with design risk (such as a rainfall contract). Moreover, define the premium for each contract as π_z^{AY} and π_z^{II}, respectively. Now, define:

$$D_{2,z} = \frac{\sum_{t=1}^{T} max\left(\left[\left(I_{AY}(\bar{y}_{zt}^*) - \pi_z^{AY}\right) - \left(I_{II}(f^*(S_{zt})) - \pi_z^{II}\right)\right], 0\right) 1\left(\bar{y}_{zt}^* < 1\right)}{\sum_{t=1}^{T} \left(I_{AY}(\bar{y}_{zt}^*) - \pi_z^{AY}\right) 1\left(\bar{y}_{zt}^* < 1\right)} \tag{5}$$

$D_{2,z}$ expresses the mean gap between ideal net payouts under an area-yield contract and actual net payouts under an imperfect index insurance contract as fraction of total net indemnity payments under the area-yield contract. Hence, an area-yield contract or an index insurance contract with no design risk will have $D_{2,z} = 0$, while an index insurance contract with so much design risk that uninsured losses are equivalent to having no insurance at all will yield a value of $D_{2,z} = 1$. Note that it is theoretically possible for $D_{2,z}$ to be greater than 1, if the design risk is so large that insurance actually increases total losses relative to having no insurance. This could happen because the insurance premium may make net payouts negative (see Clarke, 2016).

While $D_{2,z}$ provides a useful measure of insurance quality by quantifying the total fraction of uninsured losses due to design risk, it does not capture the welfare losses experienced by farmers.

Hence, we introduce another measure of insurance quality that measures the willingness to pay (WTP) to avoid design risk as a fraction of the WTP (beyond the premium) for perfect insurance. In order to estimate WTP, we assume that all farmers have the same utility function, $U(W_{izt}^C)$, where W_{izt}^C is their final wealth after their yield has been realized and insurance payouts from contract C have been made ($W_{izt}^C = \frac{\mu_{iz}}{\mu_z} + y_{izt}^* - \pi_z^C + I_C(X_{zt}))$), and $U()$ has the usual properties ($U'() > 0, U''() < 0$). The expected utility for an individual farmer can thus be approximated as[1]:

$$EU_{iz}^C = \frac{1}{T} \sum_{t=1}^T U(W_{izt}^C) \tag{6}$$

Now, the WTP to avoid the design risk under an index insurance contract, P_{iz}^{DR}, is defined by $\frac{1}{T} \sum_{t=1}^T U(W_{izt}^{AY}) = \frac{1}{T} \sum_{t=1}^T U(W_{izt}^{II} + P_{iz}^{DR})$ and the WTP for perfect insurance beyond the premium, P_{iz}^{AY}, is defined by $\frac{1}{T} \sum_{t=1}^T U(W_{izt}^{AY}) = \frac{1}{T} \sum_{t=1}^T U(\frac{\mu_{iz}}{\mu_z} + y_{izt}^* + P_{iz}^{AY})$. We can now define two welfare-based measures of index insurance quality, $D_{31,z}$ and $D_{32,z}$. $D_{31,z}$ assumes idiosyncratic risk is zero and can thus be calculated using only zone-level data:

$$D_{31,z} = \frac{P_z^{DR}}{P_z^{AY}} \tag{7}$$

For an area-yield contract, $D_{31,z} = 0$, and for an index insurance contract with so much design risk that farmers are indifferent between such a contract and having no insurance at all, $D_{31,z} = 1$. Note that again, it is possible for $D_{31,z} > 1$, if the design risk is so high that farmers are better off without insurance than with the index insurance contract.

Finally, $D_{32,z}$ relaxes the assumption that idiosyncratic risk is zero and requires individual-level data. Specifically, it is defined as:

$$D_{32,z} = \frac{1}{N} \sum_{i=1}^N \frac{P_{iz}^{DR}}{P_{iz}^{AY}} \tag{8}$$

The interpretation of $D_{32,z}$ is the same as that of $D_{31,z}$.

3   Data

To create an index which minimizes design risk and to analyze how the proposed contracts would affect the welfare of individual farmers, we need 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

---

[1]Since the objective of this metric is to estimate design risk using finite data rather than hypothetical yield distributions, we use summations instead of integrals.

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. The subsequent sections discuss the yield survey and the high-resolution satellite data, respectively.

Our study area (Figure 1) consists of four wards (counties) located East of the Pare Mountains in the Same district in the Kilimanjaro region of Tanzania and covers an area of approximately 20 by 5 kilometers. The main crops grown are paddy and maize, and paddy fields are clustered together with little other vegetation or crops. This makes paddy the ideal crop for satellite-based index insurance, since pixels will suffer from little contamination from non-paddy vegetation. The paddy clusters are also used as the basis for the insurance zones, since the agro-ecological conditions are relatively homogenous within each cluster. In total, we define ten separate insurance zones (see Figure 1). 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, paddy 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 and floods could impacts crops during any part of the growing season, the paddy plants are most vulnerable during the first month and a half following planting, which is also when rains are the most uncertain.

Figure 1: Map of study area
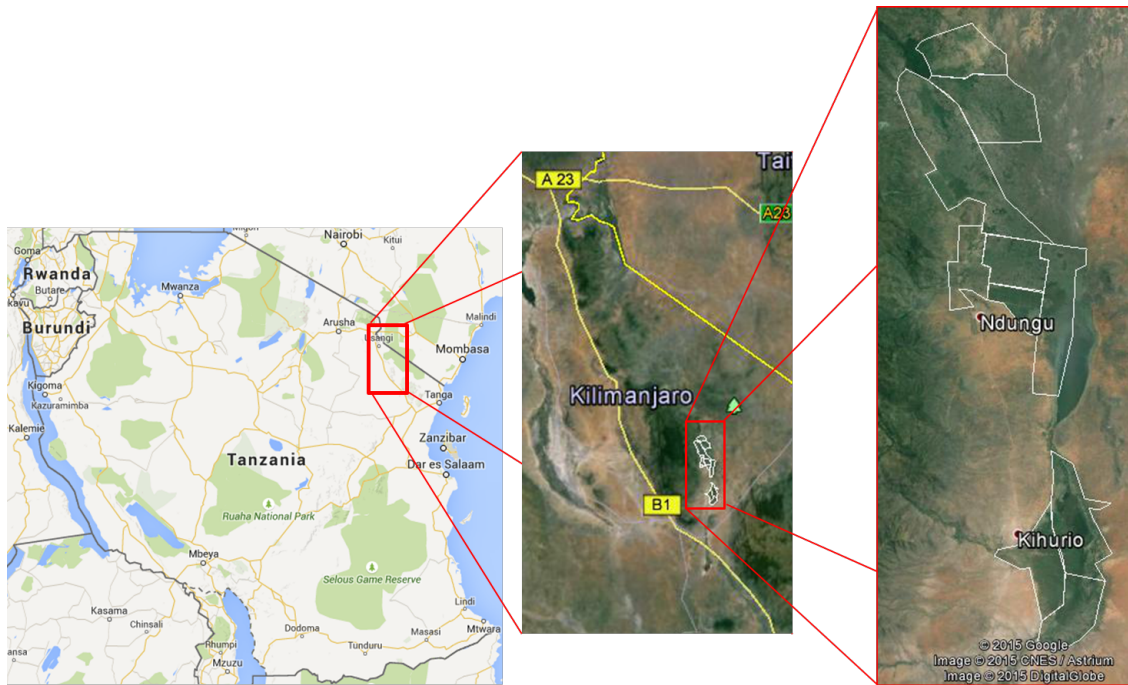
3.1 Survey data

The retrospective yield survey was conducted in May and June, 2013. Within our study area, we randomly sampled a total of 323 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. The survey was on purpose made short to keep the focus on the yield data, and we gathered data on paddy 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)[2].

Table 1 shows a set of summary statistics, grouped by each of the ten insurance zones. We note that average yields vary somewhat across the zones, with the Ndungu irrigation scheme zones (Ndungu N/S) having the highest yield, and the Southern flood plains (Southern Plain N/S/W) having the lowest average yield. There is also significant variation across zones with respect to the

---

[2]While our 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 paddy, 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 would be no penalty for not remembering 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	0.97	0.97	0.98	0.92	1.00	0.98	1.00	0.95	1.00	0.94	0.97
Irrigation use	0.84	0.58	0.87	1.00	0.94	0.97	0.75	1.00	1.00	0.97	0.89
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	0.75	0.56	0.57	0.79	0.95	0.88	0.84	0.78	0.60	0.66	0.72
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 severity idx. (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 severity idx. (1-3)	0.43	0.37	0.44	0.24	0.18	0.23	0.25	0.61	0.62	0.68	0.42
Reported drought	0.37	0.56	0.33	0.18	0.18	0.20	0.38	0.27	0.33	0.30	0.30
Reported flood	0.19	0.16	0.22	0.12	0.08	0.11	0.17	0.30	0.27	0.31	0.20
N	37	38	53	12	35	40	4	38	34	32	323

drought/flood risk, ranging from 18% indicating in Ndungu East and North to 56% in Maore East. The three Northern zones (Maore N/SE/SW) are generally most prone to drought as they depend directly on one river coming down from the mountains, while 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.

In the analyses that follow, we use both the plot-level data and aggregated zone-level data. To create the zone-level yield data, we calculate $\bar{y}_{zt} = \sum_{i=1}^{N} y_{it}$, and drop any observations for which $N < 10$. This leaves us with 74 zone-level data points.

3.2 High-resolution satellite data

Earth science data has been collected from satellites for many decades, and it has found increasing utility in agriculture over time. Presently, dozens of satellites orbit Earth with the right characteristics to support analysis of agricultural production using vegetation indices. The methods used here relate passively sensed electromagnetic radiation with actual on the ground crop production. When solar irradiance illuminates the earth and is reflected back to the sensor, multiple filters allow light from different spectrum to pass through to form a multi-band image. The relative strength of the reflected radiance in these various spectral bands are the primary input to crop health indices. The following discusses the vegetation indices used in our study and the processing methods we apply to create the final index values.

3.2.1 Vegetation Indices

It is well established that satellite remote sensing systems can collect very precise and accurate physical quantities which relate to vegetation health (Rosema, 1993; EARS, 2012; Bhattacharaya et al., 2011). The most common and well-known vegetation health index from remotely sensed imagery is the Normalized Difference Vegetation Index (NDVI). Healthy vegetation has stronger reflectivity in the near-infrared compared to biologically stressed vegetation. NDVI is also simple to calculate, and it is robust to slight differences in sensor bandpass and to absolute calibration error; the latter aspects make this a valuable index for cross-sensor studies. NDVI is expressed as a band ratio, NDVI=(NIR-VIS)/(NIR+VIS), where NIR and VIS represent the spectrally integrated reflectance over the near infrared and visible parts of the spectrum. The Enhanced Vegetation Index (EVI), is a similar measure of health, but it uses additional data from green and blue parts of the spectrum to improve the dynamic range and sensitivity in the healthier side of the index (i.e., EVI will generally not saturate as quickly as NDVI). For NDVI and the EVI, this study uses the MCD43A4 data product (Schaaf and Wang, 2015). This product utilizes an empirically driven model of the surface bidirectional reflectance distribution function (BRDF) to remove the majority of the view-angle dependent variability in the measured radiance and derived reflectance.

The NASA MODIS sensor is particularly well suited for this study because of its spectral, spatial, and temporal collection characteristics; the wide field of view sensor is flown aboard two satellites resulting in twice daily daytime coverage of nearly the entire globe. The atmospheric calibration factors and geolocation data are used as provided.

Another measure useful in this work is known as Gross Primary Productivity (GPP) . Primary production is the process of creation of organic compounds from carbon dioxide. The GPP index data used for this study comes from the MODIS MOD 17 data products (Running et al., 2015). The index is not a direct measure of GPP; instead, it attempts to compute the daily average GPP based on relationships between the Fraction of Photosynthetically Active Radiation (FPAR) and the total incident PAR to determine the Absorbed PAR (APAR). This absorbed radiation in the photosynthetic spectrum is well correlated to primary production. The authors of the algorithm to produce the MODIS MOD17 products describe the primary production as designed to provide an accurate, regular measure of the production activity or growth of terrestrial vegetation (Running et al., 1999). The sum of the GPP index during the growing season period has been shown to be one of the best indicators of the amount of new biomass (Prince, 1991; Gitelson, 2006), and should correlate well with rice yield, i.e., it should relate strongly to the number of panicles per unit area (Bastiaanssen, 2003). GPP is most closely related to the cumulative sum of either the NDVI or EVI values over the applicable time period, with healthy and growing vegetation showing a strong GPP signal.

3.2.2 Area Aggregation and Crop Masking

While the geographically aligned data are exploited in this work, variability in pixel geometry can lead to false variability in temporal index data when looking at very small areas, such as the fields in this study. In addition, the MODIS pixels are large compared to the plots of interest for this work so the impact of adjacent landcover can be significant. To combat these issues and isolate the crop signature, a spatial aggregation process is used. Specifically, a Gaussian smoothing kernel function and area averaging are applied and a weighting is used based on the likelihood that the pixel contains rice. For a given pixel, the average likelihood of the pixel including rice is calculated based on a modification of the Xiao (Xiao et al., 2006) rice mapping method. Xiao's work utilizes the EVI, NDVI, and LSWI to compute a series of masks and indicator variables. The method masks out permanent water and evergreen using thresholds on the value and abundance of measures such as NDVI. For example, evergreen pixels are those which are above NDVI of 0.6 for at least 180 days per year. Once masks are applied, the likelihood index of rice planting is calculated as a weighted linear model relationship between a water index and a vegetation health index with a trigger based on the growth observed in the vegetation index following a peak in the water index. For some analysis, this index is not applied, and in these cases, the masking to remove non-cropping regions remains. Although not truly a probability, this likelihood index varies from 0 to 1 with 0 representing non-cropping areas and 1 representing a high likelihood that rice is being planted on the given date at the given location. Figure 2 shows examples of detected versus non-detected rice and high-resolution images to display associated ground cover. The plot in the right column of the figure demonstrates a challenge for the algorithm, in which a low water index value during flooding stage – likely due to turbidity, or no flooding for transplant in this area – leads to the rice index value (blue line in each plot) not reaching a strong peak.

3.2.3 Temporal Aggregation

To produce seasonal metrics from spatially aggregated data pixels, a temporal aggregation is applied. For this study, the area under the curve, calculated using trapezoidal numerical integration, is used. Each index, a measure of the vegetation health, also has a threshold applied at the start of the season to eliminate the effects of standing vegetation interfering with the signal. By zeroing out the index at the start of the season, each year can be more easily compared to other years. The area under the curve becomes area between the curves, where the second curve is the background vegetation. This modified metric examines more closely the health of the crop relative to the background. Further enhancement could be performed by allowing the background signal to vary throughout the season to isolate crop changes in mixed vegetation areas. For any missing data, temporal interpolation is preferred compared to spatial interpolation because of the highly varying surface material composition at the spatial scales examined in this project and the predicted rate

Figure 2: Rice detection (center column) and missed detections (left and right)
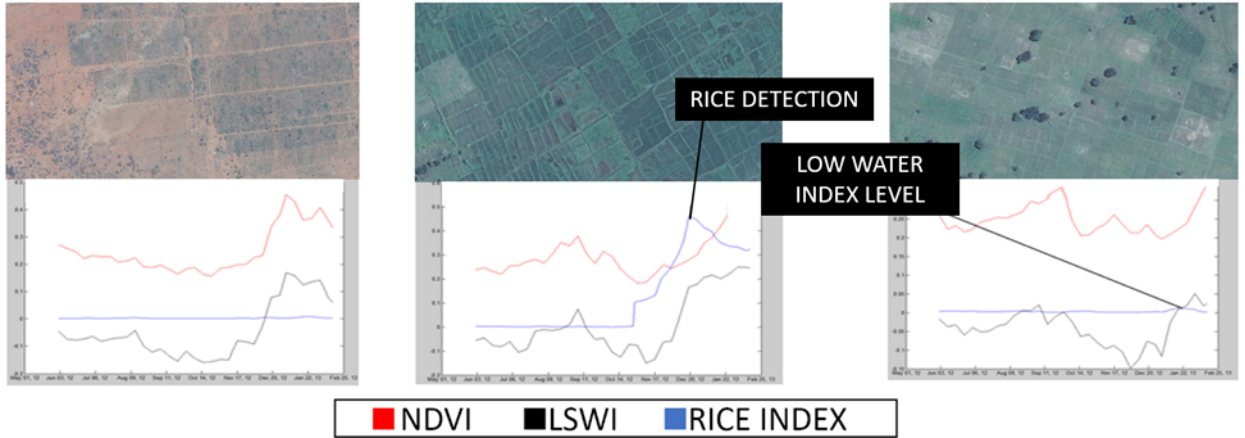
of change in temporal data. We expect the data from before and after a given date in time to correlate better to a particular missing measurement than the spatially adjacent data at the time of the missing measurement. Splined interpolation using Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) is utilized when missing data is filled.
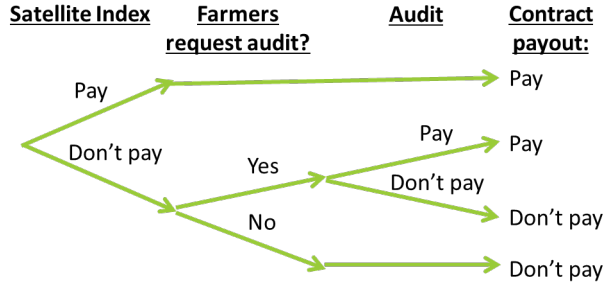
4 Minimizing design risk

In this section, we first define a set of index insurance contracts, including an area-yield contract, a satellite-based contract, and satellite-based conditional audit contract. We then estimate a response function mapping the satellite index values to the zone-level yield data and use these estimated yields as a basis for index insurance payouts. Finally, we estimate the design risk of each contract using the measures developed in section 2.2 and study how this design risk can be minimized by changing the parameters of the payout function.

4.1 Defining the contracts

As a baseline for comparison, consider first an ideal area-yield contract that compensates farmers with an amount equal to actual area-yield shortfalls below the zone mean. Assuming fixed prices, income is equal to yield, which implies that all insurance payouts and incomes can be are denoted in kg/acre. Under this contract, $I_{AY}(\bar{y}_{zt}^*) = max(-\bar{y}_{zt}^*, 0)$. Figure 4 plots the ideal area-yield indemnity function, I_{AY}, against normalized yields.

Next, consider the indemnity payment for an index insurance contract with design risk. In particular, we define the following indemnity payment for a non-perfect index insurance contract

13

Figure 3: Structure of a satellite-based conditional audit contract

Satellite Index	Farmers request audit?	Audit	Contract payout:

Pay

Pay

Don't pay

Pay → Pay

Yes

Don't pay

No → Don't pay

Don't pay

as:

$$I_{II}(f^*(S_{zt})) = max\left(min\left(-f^*(S_{zt}), 1\right), 0\right) \qquad (9)$$

where $-f^*(S_{zt})$ is the predicted loss based on the index.

Finally, consider a imperfect index insurance contract with conditional audits. Here, the primary insurance index would still be the satellite-based predicted losses, but it could in principle be any index. When the primary index imperfectly predicts actual area losses, a contract solely based on this index will still leave substantial basis risk in the hands of farmers. Hence, to partially combat this residual design risk, we add a second layer of protection by letting farmers request an audit whenever the satellite index does not accurately reflect area-yield losses. Figure 3 shows how such a contract might be structured under an all-or-nothing insurance contract. 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 has been piloted in Ethiopea (see Guush et al., 2013), 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.

For the continuous-payout contract considered here, a conditional audit contract pays farmers according to the results of the audit if the actual zone-level yield shortfall is larger than the predicted

indemnity payment by some margin. Specifically, we assume that farmers have perfect knowledge about zone-level yields and will request an audit if predicted indemnity payouts, $I_{II}(f^*(S_{zt}))$, are at least δ lower than actual zone-level yields, $\bar{y}_{zt}^*$:

$$I_{IA}(f^*(S_{zt})) = \begin{cases} I_{AY}(\bar{y}_{zt}^*) & if \ (I_{AY}(\bar{y}_{zt}^*) - I_{II}(f^*(S_{zt}))) > \delta \\ I_{II}(f^*(S_{zt})) & otherwise \end{cases} \tag{10}$$

The number of audits (expressed as a fraction), is given by:

$$a_z^{IA} = \frac{1}{T} \sum_{t=1}^{T} 1\left((I_{AY}(\bar{y}_{zt}^*) - I_{II}(f^*(S_{zt}))) > \delta\right) \tag{11}$$

Finally, the premium for each contract is given by:

$$\pi_z^C = \frac{1}{T} \sum_{t=1}^{T} I_C(X_{zt}) + q + \gamma a_z^C$$

where q is fixed loading cost that is independent of the contract (administrative costs, risk premium etc.) and γa_z^C is an audit cost that depends linearly on the number of audits performed. We assume $a_z^{II} = 0$ and $a_z^{AY} = 0$. We note that the premium for a conditional audit contract will be higher than or equal to that of a standard imperfect index insurance contract for two reasons. First, audits are expensive, which implies that $\gamma a_z^C \geq 0$, and second, the additional payouts due to audits will increase average indemnity payouts.

## 4.2 Creating a yield response function

In order to produce predicted yield deviations, $f^*(S_{zt})$ based on the available satellite index data, we estimate a simple response function which maps the two satellite indices, GPP and EVI, to actual zone-level yields. In particular, we estimate the following model:

$$\bar{y}_{zt}^* = \beta_0 + \beta_1 * S_{zt}^{GPP} + \beta_2 * S_{zt}^{EVI} + \beta_3 * S_{zt}^{GPP} * S_{zt}^{EVI} + \beta_4 * D_{zt}^{GPP} + \beta_5 * D_{zt}^{EVI} + \gamma_z + \epsilon_{zt}^* \tag{12}$$

where $\bar{y}_{zt}^*$ is the mean yield deviation in zone z and year t, S_{zt}^{IDX} the satellite index value for index IDX, D_{zt}^{IDX} is a dummy for whether the index is below its zone-level mean, and γ_z is a zone-level fixed effect.

Table (2) displays the results of estimating model 12 (column [4]) in addition to estimating the model for each index separately (columns [1] and [2]) and without the switching parameters (column [3]). We note that the coefficient on each of the satellite indices is positive and significant,

Figure 4: Area-yield indemnity schedule
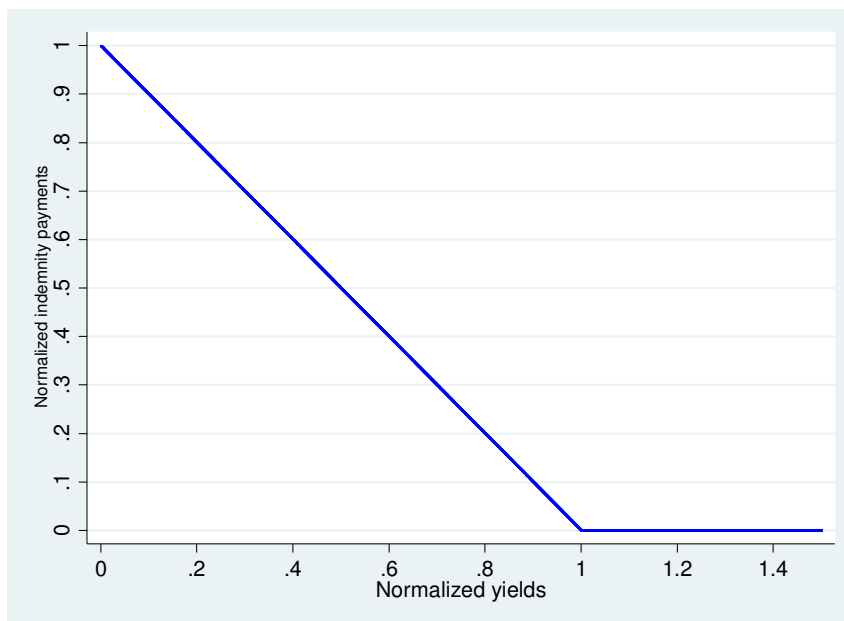
even when including both in the same model, demonstrating that the satellite data can explain inter-year variation in yields. Moreover, allowing for a different intercept for index values below the zone mean improves the fit; however, the individual parameters are no longer significant. Since we are concerned with predicting yields rather than the significance of each parameter, we use model [4] to predict yields in the analysis that follows, as this model has the best in-sample fit. Specifically, with an R^2 of 0.42, the model explains more than 40% of the variation in area yields. This implies that an index insurance based on this model will still have some design risk; however, this performance is better than that of most rainfall-based contracts (Smith and Watts, 2009).

4.3 The impact of design risk on contract performance

While the R^2 of the model is a commonly used measure of design risk, it does not fully capture the impact of design risk on the performance of index insurance contracts. For example, if the model accurately predicts bad years but inaccurately predicts yields during a good year, there may still be minimal design risk. In this section, we use the metrics developed in section 2.2 to quantity the design risk of the satellite-based index insurance contract and the satellite-based conditional audit contracts defined in section (4.1). Moreover, we show for the audit contract, design risk can be reduced by adjusting the audit trigger (δ) depending on the level of audit costs (γ).

First, to better understand the performance of our satellite-index yield model, we compare

16

Table 2: Regression results for select yield response functions

	[1]	[2]	[3]	[4]
EVI Index	0.080***		0.043*	0.047
	(0.01)		(0.02)	(0.03)
GPP Index		0.000089***	0.000071***	0.000019
		(0.000008)	(0.000008)	(0.00004)
EVI Index $\times$ GPP Index			-0.0000098	-0.0000063
			(0.000009)	(0.000009)
1[GPP < Mean(GPP)]				-0.12
				(0.08)
1[EVI < Mean(EVI)]				0.016
				(0.03)
Constant	0.85***	0.97***	0.91***	0.96***
	(0.02)	(0.003)	(0.03)	(0.05)
Zone effects	Yes	Yes	Yes	Yes
Observations	74	74	74	74
R^2	0.306	0.353	0.384	0.424

Standard errors in parenthesis, clustered at the zone level. Dependent variable is zone-level yield

*p < 0.10, ** p < 0.05, *** p < 0.01

predicted zone-level yields to actual yields. Figure (5) shows a scatter plot of actual (normalized) zone-level yields $(1 + \bar{y}_{zt}^*)$ versus predicted (normalized) yields $(1 + f^*(S_{zt}))$. It is clear from the plot that the model is quite effective at identifying bad years, but that it does a poor job at estimating the severity of losses. In particular, out of the 25 zone-years that suffered losses in excess of 0.05, only 3 were misclassified as having high yields. However, among the 4 zone-years that experienced severe yield losses (> 0.25 loss), none were predicted to have suffered a loss in excess of 0.2. While this result is difficult to explain, it appears that a low satellite index value is a rather noisy signal in that the yield variance is high, conditional on a low index value, while it is a less noisy signal for high index values. Moreover, 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.

Next, consider the payouts of an insurance contract based on these yield predictions. Figures 6a and 6b show a scatter plot of net indemnity payouts $(I_C(X_{zt}) - \pi_z^C)$ against actual normalized zone-level yields for the satellite-only contract and the satellite conditional audit contract, respectively[3]. Under the satellite-only contract, net payouts are mostly positive when zones experience losses; however, as the model fails to predict the severity of losses, indemnity payments under the satellite contract fall significantly short of those under an area-yield contract when losses are significant. In

---
[3]For these plots, we assume $\gamma = .05$, and $\delta = 0.15$
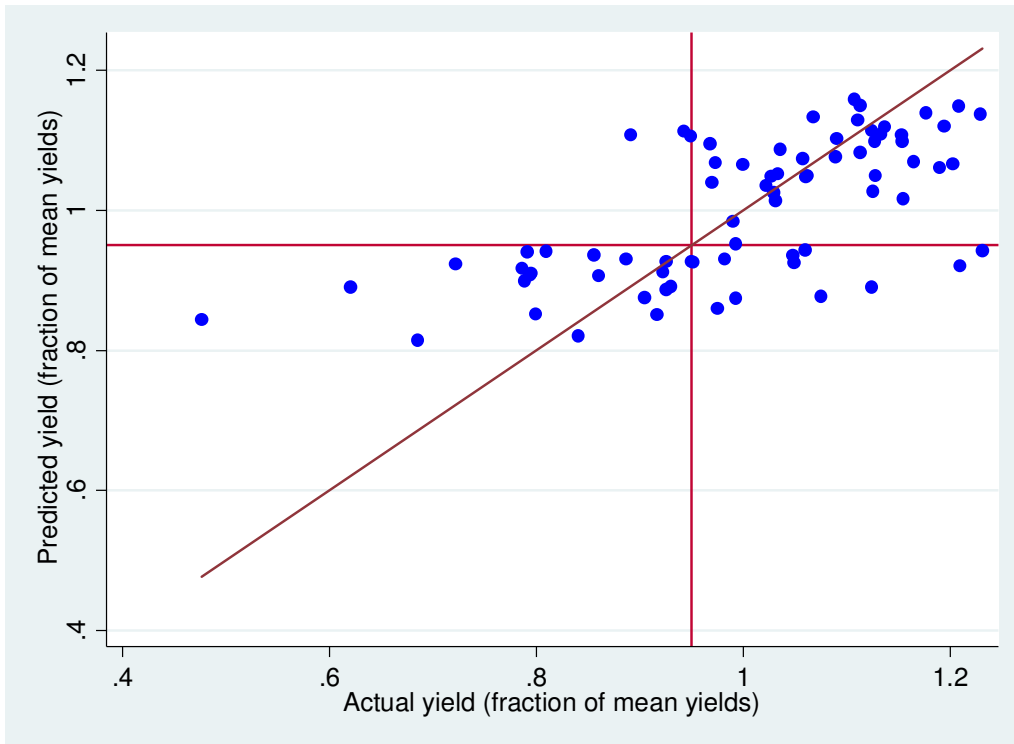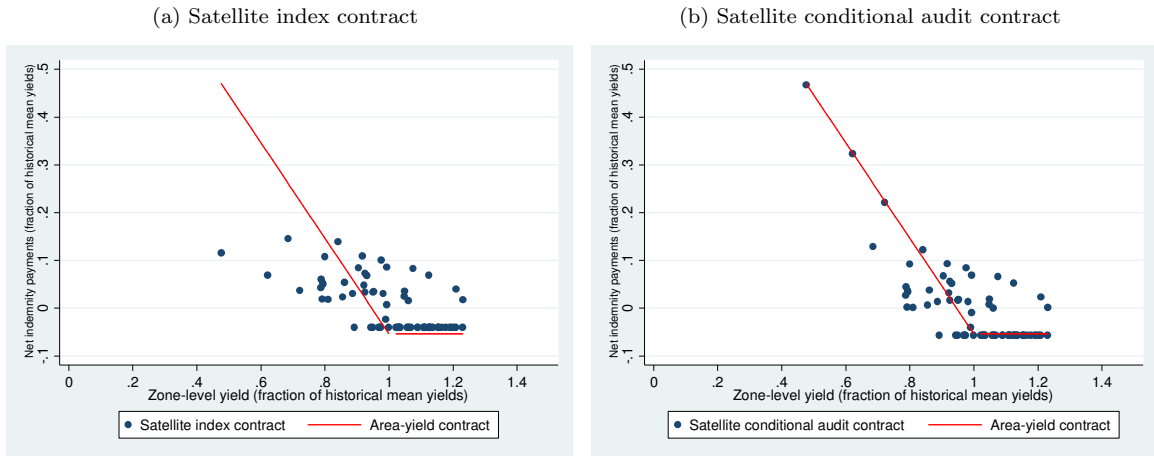
Figure 5: Actual versus predicted yields

18

Figure 6: Net indemnity payments versus zone-level yields

(a) Satellite index contract   (b) Satellite conditional audit contract

particular, the contract fails to fully compensate for a loss (within a 5% margin) 20% of the time (false negatives) and it pays out an amount higher than the actual loss (within a 5% margin) 15% of the time (false positives). Hence, the payout if consistent with actual losses 65% of the time.

To quantity the level of design risk for the satellite-only contract, we calculate D_2 and D_{31}[4]. In particular, we find that $D_2 = 0.79$ and $D_{31} = 0.53$. The first number implies that the mean gap between ideal net payouts under an area-yield contract and actual net payouts under the satellite-only contract is 79% of total net indemnity payments under the area-yield contract. The second number implies that the WTP to avoid design risk is 53% of the WTP (beyond the premium) for perfect insurance.

Next, consider the impact on design risk when farmers have the option to request an audit. Since audits are costly and do not address false positives, a conditional audit contract cannot completely eliminate design risk. However, it might significantly reduce it. Figure 6b shows that with an audit threshold of 0.15, the worst false negatives are eliminated. In particular, the two worst years on record are now fully indemnified due to the audit. Some loss years are still not fully indemnified; however, farmers will at most face of net loss of 0.15 in addition to the premium, which is still acceptable protection.

We now ask under which circumstances a conditional audit contract could reduce design risk relative to a standard no-audit contract. Specifically, given a fixed audit cost, what is the optimal audit threshold? If audits are free, then obviously a zero threshold would be optimal for any risk-averse farmer, since there is no net increase in mean income for the farmer. However, when audits are costly, there exists a trade-off between correcting false negatives and the premium increase due

---

[4]Our measure of D_{31} assumes a CRRA utility function with a coefficient of 1.5

to the cost of audits. We also note that a threshold equal to 1 is essentially the same as a no-audit contract. Hence, δ allows us to continuously adjust the level of audits. Figure 7 shows how our two measures of design risk (D_{21} and D_{31}) vary with δ (the audit threshold) for different levels of γ (audit cost). For comparison, the level of design risk under the no-audit contract is also included.

First, Figure 7a shows that the possibility of audits significantly reduces our payout-gap measure of design risk (D_{21}) for all levels of δ over a reasonable range of audit costs. In particular, for low audit costs, a zero-threshold can reduce D_{21} by as much as 0.60; however, as audit costs increase, the improvements are smaller and the optimal threshold increases. For example, if the audit cost is $\gamma = 0.10$ (which implies that each audit costs 10% of the sum insured), an audit threshold of 0.1 can reduce D_{21} to 0.43 (vs. 0.79 under a no-audit contract).

Moreover, as shown in Figure 7b there is also potential for significant reductions in our welfare-based measure of design risk (D_{31}) due to audits; however, the reductions are more limited, particularly when the cost of audits is higher. This is the case because D_{31} considers the premium increase due to audits for the entire distribution and not just for below-average yields. For an audit cost of 0.02 (roughly equivalent to 33% of the AFP), D_{31} can be reduced by 0.16 if the audit threshold is set to 0.1 (roughly equivalent to 166% of the premium). Moreover, we note that even at high audit costs, a conditional audit contract is still welfare-enhancing for farmers as long as δ is sufficiently high.

Overall, this analysis demonstrates that a conditional audit contract can reduce (or at least not increase) design risk relative to a standard no-audit index insurance contract assuming the audit threshold parameter is adjusted to maximize the benefit. If such an adjustment is not made (e.g. if audits are requested any time losses are marginally higher than estimated payouts), an audit contract may in fact be welfare-reducing.

5   Individual effects

By construction, index insurance cannot protect farmers against idiosyncratic shocks; however, the magnitude of idiosyncratic risk relative to covariant risk still affects the value of index insurance to farmers and thus the impact of design risk. For example, for a farmer who is not at all exposed to covariant risk, perhaps because she has invested in irrigation, even an 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. The following section explores the effect of the different insurance contracts on individual-level payouts and and welfare and quantifies the design risk of the satellite-based contract and the conditional audit contract when the assumption of zero idiosyncratic risk is relaxed.

Figure 7: Comparing design risk by audit threshold (δ) and audit cost (γ)

(a) Payout-gap design risk (D_{21})
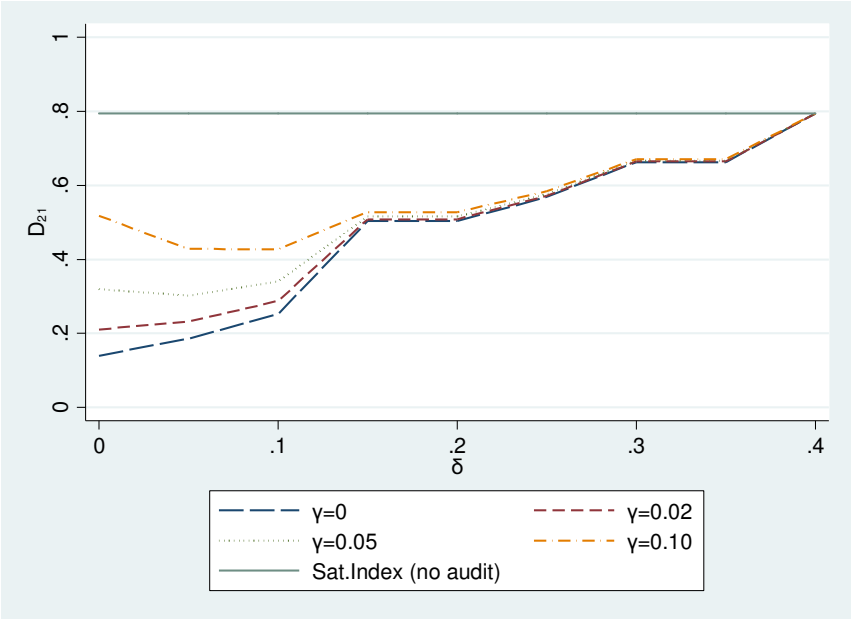
(b) Welfare-based design risk (D_{31})
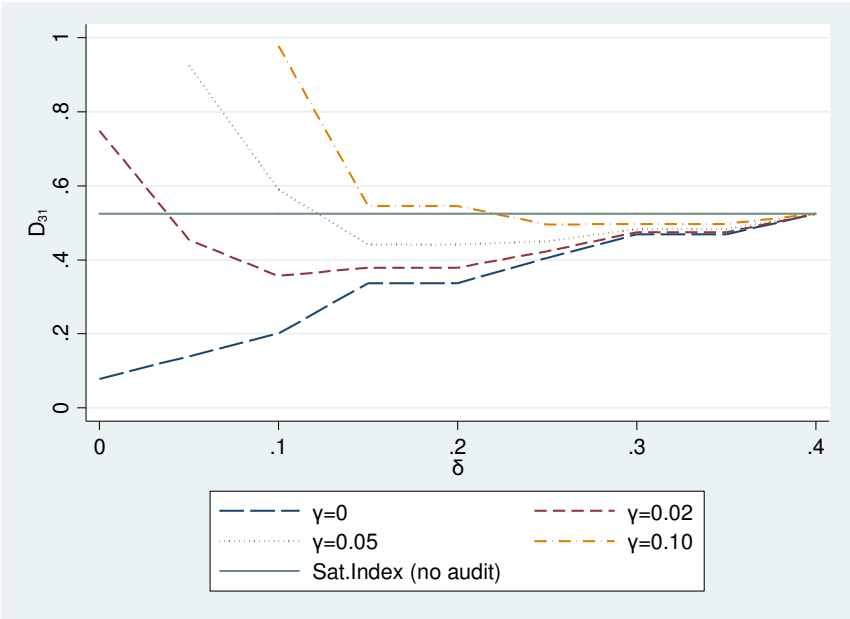
5.1 Insurance payouts

First, consider how net payouts (indemnity payments less premiums paid) vary with losses under the three contracts. Figure (8) shows the results of fitting a local polynomial regression to individual net payouts as a function of individual yields (expressed as a fraction of individual historic mean yields). First, note that the satellite-only contract performs the worst relative to the area-yield contract when losses are the highest. In particular, farmers who experience a total loss receive on average only a 7% net payment. This is compared to a 19% net payment under an area-yield contract. The satellite-based conditional audit contract corrects this shortcoming, and average net 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. 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.

5.2 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 enhance the welfare of individual farmers. As shown in Clarke (2016), 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 and quantity the level of welfare-based design risk (D_{32}) for different model parameters[5].

Estimation of welfare effects is done using expected utility analysis, which is consistent with similar analysis of utility gains from index insurance (Woodard et al, 2012; Clarke, 2016). However, our panel dataset of individual historical yields is incomplete, as many farmers were only able to recall their yields for particular years and were specifically told to provide us with only the data they could remember with certainty. Since incomplete data records could have large impacts on expected utility, particularly when there are only a maximum of ten years per individual, we employ a simulation approach to create a complete yield history for individual farmers. In particular,

---

[5]Our analysis does not consider informal risk sharing among individuals within a zone. Hence, our estimates of WTP and demand are low bounds on the their true values. See Mobarak and Rosenzweig (2012) for an analysis of how informal risk sharing affects the impact of basis risk on index insurance uptake.

Figure 8: Net payouts (indemnity payment less premiums paid) by individual-level yields under each of the three contracts. A kernel density function of individual-level normalized yields is superimposed.
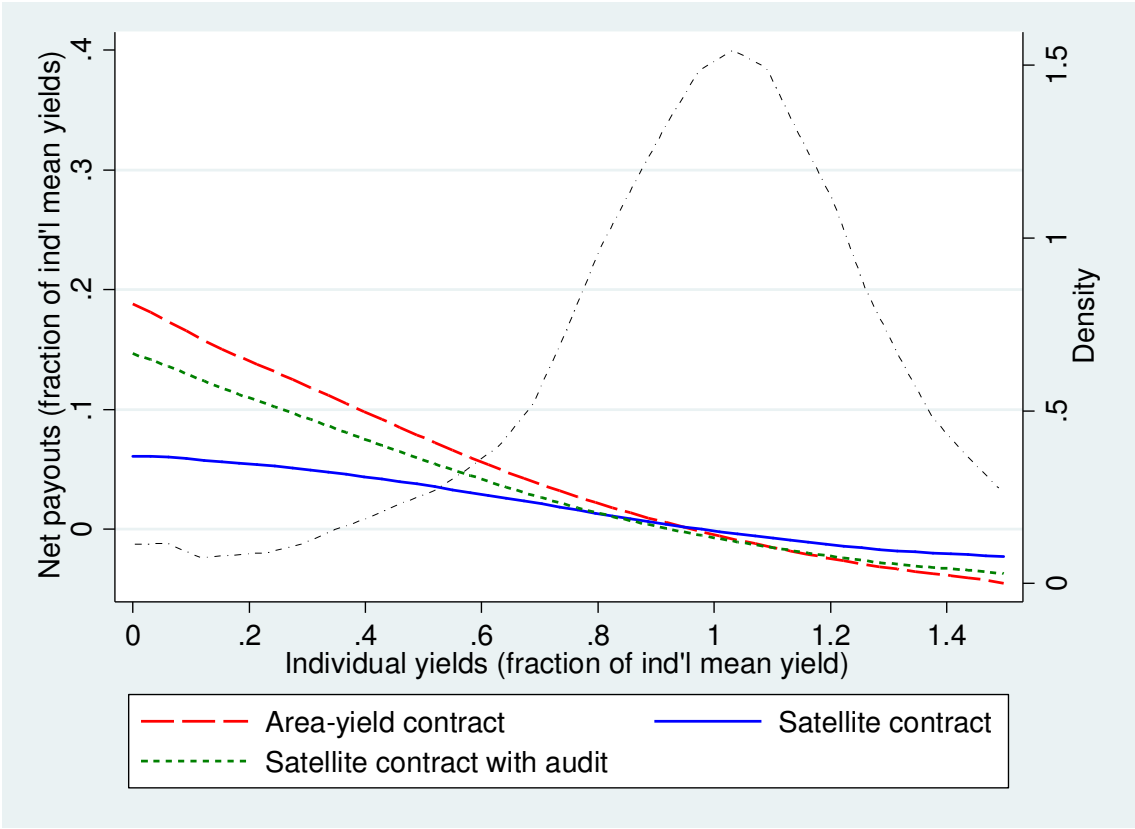
we first estimate a random intercepts and coefficients model of individual yields to obtain the distribution of the parameters and the error term:

$$\frac{y_{izt}}{y_z} = \alpha_{iz} + \beta_{iz}\bar{y}_{zt}^* + \epsilon_{izt}^* \tag{13}$$

We assume that α_{iz} and β_{iz} are distributed jointly normal and that ϵ_{izt}^* is independent and normally distributed. Estimating this model yields the following: $(\alpha, \beta) = (1,1)$, $\sum_{\alpha_{iz}, \beta_{iz}} = \begin{pmatrix} .113 & .0744 \\ .0744 & .413 \end{pmatrix}$, $\sigma_{\epsilon_{zt}^*}^2 = 0.574$, which we use to simulate individual (normalized) yields, $\tilde{y}_{izt}$, given zone-level normalized yield deviations, $\bar{y}_{zt}^*$. 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, respectively[6]:

$$EU_{iz,NI} = \sum_{t=1}^{T} U\left(w + \tilde{y}_{izt}\right) \tag{14}$$

$$EU_{iz,C} = \sum_{t=1}^{T} U\left(w + \tilde{y}_{izt} + I_C(X_{zt}) - \pi_z^C + P_{iz}^C\right) \tag{15}$$

where $U(z_{izt}) = \frac{1}{1-\alpha}(z_{izt})^{1-\alpha}$, α is the coefficient of relative risk aversion, and w is initial wealth[7]. P_{iz}^C is an additional premium, beyond the actuarially fair premium (AFP) (including the possible audit cost). Now, in order to estimate the welfare impact of index insurance, we calculate an individual's WTP for each insurance contract by solving for an individual's additional premium P_{iz}^C such that $EU_{iz,NI} = EU_{iz,C}$, and compute the mean WTP across all individuals.

First, consider our welfare-based measure of design risk when allowing for idiosyncratic basis risk, D_{32}. Figure **??** shows how D_{32} varies with the audit trigger (δ) under various audit cost assumptions (γ) compared to a no-audit contract (assuming $\alpha = 1.5$). These results demonstrate that a conditional audit contract can significantly reduce design risk also when accounting for the idiosyncratic risk inherent in index insurance. In fact, the potential for a conditional audit to reduce design risk is even higher when accounting for individual heterogeneity. For example, with an audit cost of 0.02, the conditional audit contract has a design risk of only 0.15 (with $\delta = 0.15$) relative to a design risk of 0.53 under a the satellite-only contract.

---

[6]We simulate yield data for 1000 individuals per insurance zone

[7]We assume w 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

Figure 9: Comparing design risk by audit threshold (δ) and audit cost (γ) under idiosyncratic risk
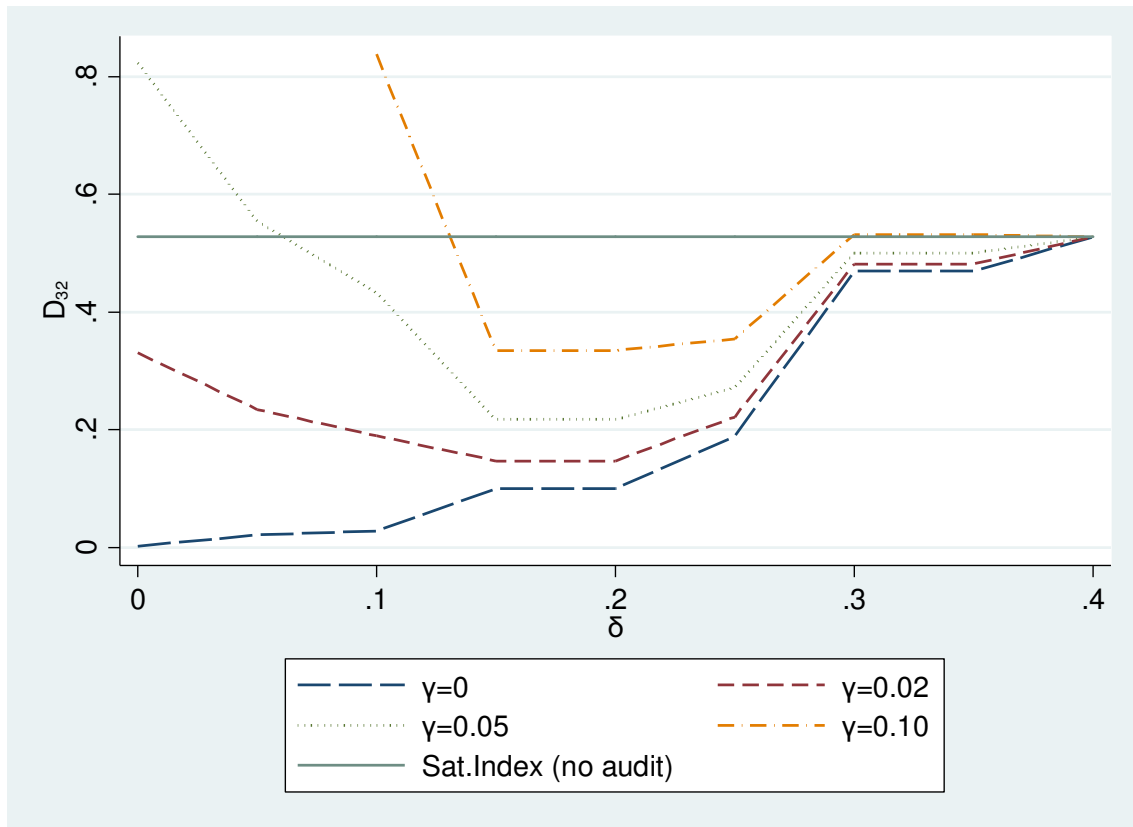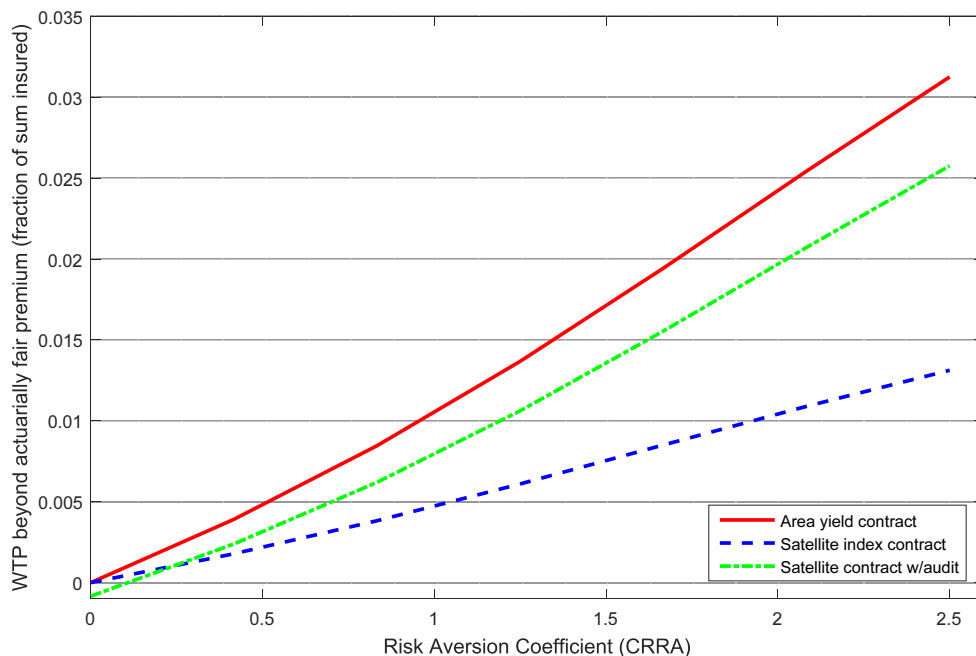
Figure 10: WTP for index insurance contracts

Next, consider how WTP compares across the three different contracts. Figure (10) shows the average WTP beyond the AFP (including possible audit costs) among farmers for the three different contracts over a reasonable range of risk preferences ($\alpha \in [0, 2.5]$), assuming an audit cost of $\gamma = 0.02$ (again, roughly equal to 33% of the AFP) and an optimal audit trigger. Note that this WTP is expressed as a fraction of the total sum insured rather than as a fraction of the AFP. As expected, when individuals are risk neutral, the WTP equals zero for the area-yield contract and the satellite contract, and slightly negative for the conditional audit contract (due to the audit costs). 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. In particular, for an individual with a CRRA of 1.5, the WTP for a satellite-only contract is 0.007 above the AFP (12% of the premium given an AFP of 6% of the sum insured). Moreover, farmers' WTP for the satellite-based conditional audit contract is only marginally less than that of the ideal area-yield contract. In particular, for an individual with a CRRA of 1.5, the WTP for the ideal area-yield contract is 0.017 (28%) above the AFP, versus a 0.013 premium (22%) for the satellite-based conditional audit contract.

The above analysis only considers the average WTP across all farmers. However, given that

26

Figure 11: Demand for index insurance contracts assuming a loading cost of 1% of total sum insured
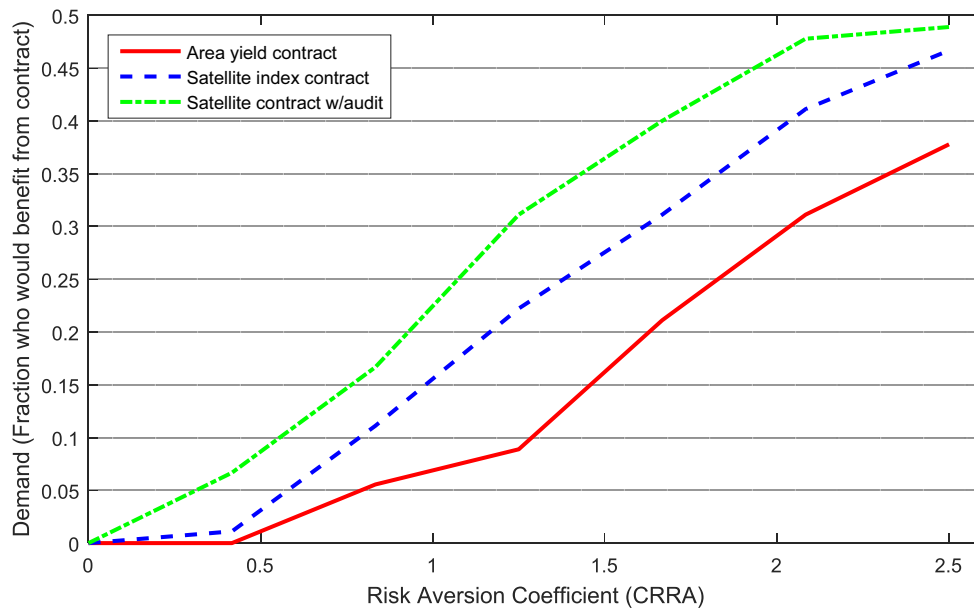
there is significant heterogeneity among farmers in how closely their yields are correlated with area yields (as measured by β_{iz} in equation 13), there will also be heterogeneity in farmers' WTP for index insurance. Moreover, the ideal area-yield contract is generally not feasible in data-scarce environments, such as much of sub-Saharan Africa. Hence, to compare the hypothetical demand for each contract, we assume the area-yield contract fetches an additional premium equal to conducting an audit every season (thus approximating the cost of collecting annual yield data). To estimate the demand for each of the three index insurance contracts, we calculate the percentage of farmers in the simulated population that would benefit from the insurance, assuming all three contracts have an additional loading cost of 1% of the total sum insured (~17% of the average premium). We can then calculate the proportion of farmers benefiting from insurance as:

$$\frac{1}{N} \sum_{i=1}^{N} 1\left(EU_{iz,C} > EU_{iz,NI}\right) \tag{16}$$

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 (11) 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. We now see that demand for a feasible area-yield contract is quite low, due to the high cost of collecting annual yield data. In fact, a satellite-only contract, which has a higher basis risk but a lower cost has a higher demand than a feasible area-yield contract. Furthermore, when the audit provision is added to the contract, demand increases further despite the increase in cost due to audits. This result is driven by the fact that under the conditional audit contract, audits are only conducted when they result in a net welfare gain for farmers, which is when the severity of a basis risk event is sufficiently large to justify the cost of the audit. In particular, for farmers with a CRRA of 1.5, demand would be 16%, 27%, and 36%, under the area-yield contract, the satellite-only 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.

6 Conclusion and Discussion

While index insurance in theory offers a promising solution to the problem of covariant risk in smallholder agriculture, its impact has so far been tempered 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 and resulting in the perverse finding that index insurance may in fact increase the risk faced by the farmer. This paper addresses the basis risk problem contractually by proposing and analyzing an index insurance contract that specifically targets design risk. In particular, we design a satellite-based conditional audit contract that minimizes cost and simultaneously addresses negative basis risk events. However, since audits are costly, there is a tradeoff between mitigating uncompensated losses and the costs due to additional audits. We therefore develop a welfare-based measure of design risk allowing us to characterize the optimal audit trigger given the audit cost and to compare the design risk of various contracts. Using a unique panel dataset of plot-level rice yields in Tanzania to study, we analyze the quality, welfare implications, and demand of a satellite-only contract and the satellite-based conditional audit contract.

Our results indicate that demand for both the satellite-only contract and the satellite-based conditional audit contract would be substantial under reasonable loading cost assumption and assuming typical levels of risk aversion. At the same time, demand for a feasible area-yield contract in data-sparse environments would be low, due to high data collection costs. In particular, assuming a CRRA of 1.5, demand would be 16%, 27% and 36%, under the feasible area-yield contract, the satellite-only contract, and the satellite-based conditional audit contract, respectively. Moreover, we show that a positive audit trigger is optimal whenever audits costs are positive. Specifically, for our sample, audits should only be allowed if actual losses are at least 15-20% higher than predicted

yields, assuming an audit equivalent of approximately 30-80% of the premium.

These findings have several important implications for the future of index insurance in developing countries. First, 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. Second, when implementing a conditional audit contract, the audit trigger should be positive and ideally calibrated to minimize design risk. Finally, the welfare-based design risk measures developed in this paper could be used to evaluate the performance of any existing index insurance contacts, thus allowing insurance practitioners and researchers to objectively compare index insurance quality and to assess whether a contract may even be welfare-enhancing for insurance-holders.

Finally, it is important to consider some of the limitations of these results. In particular, the estimates presented in this paper are based on hypothetical simulations rather than actual demand measurements. Hence, further research is needed to test the viability of these contracts in a real-world setting, such as conducting an an RCT, in which different contracts are randomly offered to people. Moreover, 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] Bastiaanssen, W. G., & Ali, S. (2003). A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. Agriculture, ecosystems & environment, 94(3), 321-340.

[2] Bhattacharya, B. K., Mallick, K., Nigam, R., Dakore, K., & Shekh, A. M. (2011). Efficiency based wheat yield prediction in a semi-arid climate using surface energy budgeting with satellite observations. Agricultural and forest meteorology, 151(10), 1394-1408.

[3] 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.

[4] Carter, M., de Janvry, A., Sadoulet, E., & Sarris, A. (2014). Index-based weather insurance for developing countries: A review of evidence and a set of propositions for up-scaling. Development Policies Working Paper, 111.

[5] Clarke, D. J. (2016). A theory of rational demand for index insurance. American Economic Journal: Microeconomics, 8(1), 283-306.

[6] Clarke, Daniel, Olivier Mahul, Kolli Rao, and Niraj Verma. 2012. "Weather Based Crop Insurance in India." World Bank Policy Research Working Paper No. 5985, The World Bank.

[7] 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), 104-35.

[8] Dercon, S., & Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. Journal of development economics, 96(2), 159-173.

[9] 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)

[10] Elabed, G., & Carter, M. R. (2015). Compound-risk aversion, ambiguity and the willingness to pay for microinsurance. Journal of Economic Behavior & Organization, 118, 150-166.

[11] Gitelson, A. A., Viña, A., Verma, S. B., Rundquist, D. C., Arkebauer, T. J., Keydan, G., & Suyker, A. E. (2006). Relationship between gross primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. Journal of Geophysical Research: Atmospheres, 111(D8) Guush, B., Clarke, D., Vargas, H., & Alemayehu, S. (2013). Insuring Against the Weather: Addressing the Challenges of Basis Risk in Index Insurance Using Gap Insurance in Ethiopia. Unpublished, Index insurance Innovation Initiative.

[12] Jensen, N. D., Mude, A. G., & Barrett, C. B. (2018). How basis risk and spatiotemporal adverse selection influence demand for index insurance: Evidence from northern Kenya. Food Policy, 74, 172-198.

[13] Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. The Quarterly Journal of Economics, 129(2), 597-652.

[14] Kazianga, H., & Udry, C. (2006). Consumption smoothing, Livestock, insurance and drought in rural Burkina Faso. Journal of Development economics, 79(2), 413-446.

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

[16] Mobarak, Ahmed Mushfiq, and Mark Rosenzweig. 2012. "Selling Formal Insurance to the Informally Insured." http://www.econ.yale.edu/growth_pdf/cdp1007.pdf.

[17] Prince, S. D. (1991). A model of regional primary production for use with coarse resolution satellite data. International Journal of Remote Sensing, 12(6), 1313-1330.

[18] Rosema, A. (1993). Using METEOSAT for operational evapotranspiration and biomass monitoring in the Sahel region. Remote Sensing of Environment, 46(1), 27-44.

[19] Running, S., Mu, Q., Zhao, M. (2015). MOD17A2H MODIS/Terra Gross Primary Productivity 8-Day L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MOD17A2H.006

[20] Running, S. W., Nemani, R., Glassy, J. M., & Thornton, P. E. (1999). MODIS daily photosynthesis (PSN) and annual net primary production (NPP) product (MOD17) Algorithm Theoretical Basis Document. University of Montana, SCF At-Launch Algorithm ATBD Documents (available online at: www.ntsg.umt.edu/modis/ATBD/ATBD_MOD17_v21. pdf).

[21] Schaaf, C., Wang, Z. (2015). MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global - 500m V006 [Data set]. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MCD43A4.006

[22] Skees, J. R., & Barnett, B. J. (1999). Conceptual and practical considerations for sharing catastrophic/systemic risks. Review of Agricultural Economics, 21(2), 424-441.

[23] Smith, V., & Watts, M. (2009). Index based agricultural insurance in developing countries: Feasibility, scalability and sustainability. https://www.agriskmanagementforum.org/sites/agriskmanagementforum.org/files/Documents/vsmith-index-insurance.pdf

[24] Stiglitz, J. E. (1981). The theory of commodity price stabilization: a study in the economics of risk. Clarendon Press.

[25] Woodard, J. D., Pavlista, A. D., Schnitkey, G. D., Burgener, P. A., & Ward, K. A. (2012). Government insurance program design, incentive effects, and technology adoption: the case of skip-row crop insurance. American Journal of Agricultural Economics, 94(4), 823-837.

[26] Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J. Y., Salas, W., & Moore III, B. (2006). Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sensing of Environment, 100(1), 95-113.