

Agricultural Index Insurance Design: An Optimization Approach

José I. Velarde Morales

August 24, 2022

1 Introduction

Lack of access to credit and insurance is often cited as a significant factor hindering agricultural productivity in developing countries. Nearly two thirds of the world’s poor are employed in agriculture, and addressing this problem could have significant welfare implications. Agricultural insurance is, even in the best circumstances, a hard problem. Many of the features one would want (independent units, uncorrelated risk, etc) are missing in this context. When considering insurance in developing countries, the problem becomes even harder because of verification costs. Traditionally, whenever an adverse event happens, the insured party contacts the insurer, and the insurer verifies the claim and issues a payout. However, agriculture in developing countries is often characterized by many small farmers spread out over hard to reach regions. This makes verification prohibitively costly.

Researchers developed index insurance as a less costly way to offer insurance in developing countries. In index insurance, an index (or statistic) is created using easily observable quantities, and it is used to determine whether the insured party suffered an adverse event. In the past, indices have been constructed using rainfall, weather, and satellite images. If the index falls below a pre-determined threshold, the insurance company automatically issues out payments to the insured. This allows the insurance company to circumvent the issue of verification, moral hazard, and adverse selection, since the actions of individual farmers can’t affect the index. Even though index insurance provided a less costly way of providing insurance for small farmers, it has been difficult to scale up. There are several problems with index insurance. One of the main problems is low take up: farmers are often unwilling to purchase the insurance at market prices. Another problem, as previously mentioned, is the cost. The purpose of this project is to make this insurance less costly by improving the design of insurance contracts. In the following section we describe the methods that are currently employed in the design of index insurance. In the last section we describe our proposed improvements.

2 Literature Review

There are many studies that evaluate how index insurance impacts the behavior of farmers (see, for example Karlan et al. (2014)). However, there has been relatively little research concerning the design of index insurance. In Chantarat et al. (2013), the authors describe the design of an index insurance for pastoralists in Northern Kenya. The authors use a satellite based index, the Normalized Difference Vegetation Index, to predict herd mortality. They use a statistical clustering method to assign each $8km^2$ block to a larger region. Then, for each region, they specify a piecewise linear contract based on the predicted herd mortality. In Jensen et al. (2019), the authors compare the welfare implications of using different satellite based indices for insuring pastoralists against drought. In this study, the regions are taken as given, and they also use piecewise linear contracts.

As to the setting of strike values, the authors offer the following: "Similar to IBLI (Index Based Livestock Insurance) policies being sold, the strike for each index product was set at each index's within-index-unit 20th percentile". Using these contracts, the authors then compared how the farmers would have fared under each of the different insurance contracts.

In Flatnes, Carter, and Mercovich (2018), the authors propose augmenting a traditional index insurance contract with the option for an audit. In this augmented contract, the insured farmer has the option to request an audit if they believe a payout should have been issued but wasn't. In this case, the insurance company would check the farmer's claim and issue a payment if the claim is correct. If the claim is correct, the insurance company would cover the cost of the audit, if it was incorrect, the farmer would have to cover the cost of the audit. In this study, the insurance zones are based on observed clustering of fields. "The main crops grown are paddy and maize, and paddy fields are clustered together... The paddy clusters are also used as the basis for the insurance zones". Here, the insurance contract pays out the losses predicted by the index. There are no trigger values set or piecewise linear contracts.

In general, there does not seem to be a standard methodology for developing index insurance products. As stated in Bank (2011), "The reader should be aware that there is no single methodology in this field ... [this paper] describes an approach that has been used in a number of index pilot activities undertaken by the World Bank and its partners."

2.1 Papers on the design of Index Insurance

- Dalhaus, Barnett, and Finger (2020)

3 Optimization Approach

3.1 Objective and Constraints

We spoke to practitioners to determine what a good objective would be for the model. Minimizing income variance is not necessarily a good objective, because it is symmetric, and in this setting, we are only worried about losses. Minimizing variance also has the drawback that it leads to extreme solutions. In interviews we conducted, practitioners expressed that the ideal objective would be to minimize the probability that farmers' wealth would drop below a certain threshold. They also expressed that the most relevant constraints in practice are price constraints, and payout frequency constraints. Through interviews, we also learned that piecewise payout functions are always used in index insurance because they are simple to explain and work well. This lead us to the following, idealized, model.

3.2 Original Simplified Model

The first model we proposed aimed to minimize the probability that a farmer's net loss would exceed an exogenously given threshold. This was subject to a maximum premium constraint and a constraint specifying a piecewise linear structure for the payout function. In the model below, l is the farmer's loss, π is the premium he pays, $I(\theta)$ is the insurance payout, c_k is the cost of capital, and \bar{l} is the loss threshold we don't want the farmer to exceed. In other words, we don't want the farmer to have net losses higher than \bar{l} with high probability. The premium contains two terms, the first is for the expected payout of the contract, and the second is for the costs of holding the capital necessary to insure the contract. K^P is the amount of capital needed by the insurer to be able to uphold their side of the contract with high probability. For exposition, we first show the

optimization model for insuring a single zone, and we then show the model for insuring multiple zones.

Single Zone

$$\min_{a,b,\pi} P(l + \pi - I(\theta) \geq \bar{l}) \quad (1)$$

$$\text{s.t. } I(\theta) = (a\theta - b)^+ \quad (2)$$

$$\pi = E[I(\theta)] + c_k K^P \quad (3)$$

$$K^P = CVaR_{1-\epsilon}(I_z(\theta_z)) - \pi \quad (4)$$

$$b \geq 0 \quad (5)$$

$$\pi \leq \bar{\pi} \quad (6)$$

Multiple Zone In the multiple zone case, there is an additional cost coming from the capital requirements to insure the entire portfolio. In the model below, ρ_z is a measure of the relative riskiness of zone z

$$\min_{t,a,b,e,\pi} \max_z P(l_z - \pi_z + I_z(\theta_z) \geq \bar{l}) \quad (7)$$

$$\text{s.t. } I_z(\theta_z) = (a_z \theta_z + b_z)^+, \forall z \quad (8)$$

$$K^P = CVaR_{1-\epsilon} \left(\sum_z I_z(\theta_z) \right) - \sum_z \bar{\pi}_z \quad (9)$$

$$\pi_z = E[I_z(\theta_z)] + c_k K^P \rho_z, \forall z \quad (10)$$

$$\pi_z \leq \bar{\pi}, \forall z \quad (11)$$

3.3 Minimum CVaR Model

Since probabilistic constraints are generally non-convex, we decided to minimize the conditional value at risk (CV@R) of the farmer's net loss instead. There are several advantages to using the CV@R as an objective. Conceptually, minimizing the CV@R of the losses makes sure that farmers still do well even when losses are in the right tail of the distribution. In other words, it aims to make sure that farmers are well off even in the case of catastrophic events, which is the purpose of insurance. It also has the advantage of having been well studied in the literature, and there are tractable reformulations for having $CV@R$ in the objective and constraints for general loss distributions. This model minimizes the $CV@R$ of the farmer's net loss subject to a constraint on the premium. The premium constraints are expressed as a fraction of the full insured amount. So, if $\bar{\pi}$ is the maximum premium share and K is the full insured amount, then the constraint on the premium would be: $\pi \leq K\bar{\pi}$.

Single Zone Model

Model Parameters

- ϵ : This defines the CV@R objective. $\epsilon = 0.1$ means that our objective is on the expected value of the loss given that it is above the 90th percentile.
- $\bar{\pi}$: This is the maximum value of the premium.

- K : maximum insured amount

For made some simplifications to the model to ensure tractability.

Model

$$\min_{a,b \geq 0} CV@R_{1-\epsilon}(\ell - \min\{(a\theta + b), K\}) \quad (12)$$

$$\text{s.t. } a\mathbb{E}[\theta] + b \leq K\bar{\pi} \quad (13)$$

$$a\theta + b \geq 0 \quad (14)$$

We reformulated the problem in the following way using the results from Rockafellar, Uryasev, et al. (2000). In the model below, p_k is the probability of event k , and k indexes the possible realizations of θ, l .

$$\min_{a,b,\gamma,t} t + \frac{1}{\epsilon} \sum_k p_k \gamma_k \quad (15)$$

$$\text{s.t. } \gamma_k \geq l^k - \min\{(a\hat{l}(\theta^k) + b), K\} - t, \forall k \quad (16)$$

$$\gamma_k \geq 0, \forall k \quad (17)$$

$$0 \leq a\hat{l}(\theta^k) + b, \forall k \quad (18)$$

$$K\bar{\pi} \geq a\mathbb{E}[\hat{l}(\theta)] + b \quad (19)$$

Multiple Zone Model

Model Parameters

- ϵ : This defines the CV@R objective. $\epsilon = 0.1$ means that our objective is on the expected value of the loss given that it is above the 90th percentile.
- $\bar{\pi}$: This is the maximum value of the premium.
- $\underline{\pi}$: This is the minimum value of the premium.
- β_z : This is a measure of the relative riskiness of zone z .
- ϵ_P : This is the epsilon corresponding to the CV@R of the entire portfolio. This is used to determine the required capital for the portfolio. Values I've seen used are $\epsilon_P = 0.01$ and $\epsilon_P = 0.05$.
- Z : number of insured zones.
- c_k : cost of capital
- K_z : maximum insured amount of zone z .

Model

$$\min_{a,b,K^P} \max_z CV@R_{1-\epsilon}(\ell_z - \min \left\{ (a_z \hat{\ell}_z(\theta_z) + b_z), K_z \right\}) \quad (20)$$

$$\text{s.t. } K\bar{\pi} \geq a_z \mathbb{E}[\hat{\ell}_z(\theta_z^k)] + b_z + c_k \beta_z K^P, \forall k, \forall z \quad (21)$$

$$0 \leq a_z \hat{\ell}_z(\theta_z^k) + b_z, \forall k, \forall z \quad (22)$$

$$K^P + Z\bar{\pi} \geq CV@R_{1-\epsilon_P} \left(\sum_z a_z \hat{\ell}_z(\theta_z) + b_z \right) \quad (23)$$

The reformulation, again derived using the results from Rockafellar, Uryasev, et al. (2000) and Rockafellar and Uryasev (2002) is:

$$\min_{a,b,\gamma,t,m,K^P} m \quad (24)$$

$$\text{s.t. } t_z + \frac{1}{\epsilon} \sum_k p_k \gamma_z^k \leq m, \forall z \quad (25)$$

$$\gamma_z^k \geq \ell^k - \min \left\{ (a_z \hat{\ell}_z(\theta_z^k) + b_z), K_z \right\} - t_z, \forall k, \forall z \quad (26)$$

$$\gamma_z^k \geq 0, \forall k, \forall z \quad (27)$$

$$t_p + \frac{1}{\epsilon_P} \sum_k p_k \gamma_P^k \leq K^P + Z\bar{\pi} \quad (28)$$

$$\gamma_P^k \geq \sum_z a_z \hat{\ell}_z(\theta_z^k) + b_z - t_p, \forall k \quad (29)$$

$$\gamma_P^k \geq 0, \forall k \quad (30)$$

$$K_z \bar{\pi} \geq a_z \mathbb{E}[\hat{\ell}_z(\theta_z)] + b_z + c_k \beta_z K^P, \forall z \quad (31)$$

$$0 \leq a_z \hat{\ell}_z(\theta_z^k) + b_z, \forall k, \forall z \quad (32)$$

4 Evaluation

4.1 Toy examples

- Describe data generating process and general setup
- Discuss evaluation metrics

4.2 Data-Driven Evaluation

- Data description
- Simulation description
- Evaluation metrics