An Optimization Approach to Index Insurance

The goal of this project is to improve the efficiency of index insurance through better contract design. In index insurance, easily observable quantities (e.g. rainfall) are used to predict agricultural loss. If the model predicts a large enough loss, the insurer automatically issues out payments. While index insurance is less costly to offer than traditional insurance, it can still be expensive and highly risky for insurers. We develop an optimization-based method to make the insurance less costly through improved risk management. The method we developed designs the contracts for all insured zones simultaneously, while taking into account the correlations between the zones. This reduces costs by reducing risk for the insurer. The Bank of Thailand has a uniquely rich data on agricultural losses. In this project, we hope to (1) evaluate our method using data from the Bank of Thailand, and (2) inform the design of their insurance program.

1 Introduction

Agricultural insurance is an important tool for risk managment, but it is usually too expensive to provide in developing countries due to high verification costs and moral hazard. Index insurance was developed as a less costly method to provide 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 a loss. If the index falls below a pre-determined threshold, the insurer automatically issues out payments to the insured. Today, tens of millions of farmers across the world are covered under such programs (Greatrex et al. (2015)). Despite this, take up remains low due to high costs, and providing this insurance can still be exceedingly risky for insurers (N. Jensen and C. Barrett (2017)). The goal of this project is to improve the efficiency of index isurance through better contract design. We develop a data-driven method to design interpretable and utility improving contracts. Our method can also design the contracts of multiple insured zones simultaneously, allowing for better risk management. This work will contribute to the literature on index insurance design. Previous work has focused on verfication mechanisms or signal selection (N. D. Jensen, C. B. Barrett, and Mude (2016); Elabed et al. (2013)). The

closest work to ours is from Estefania-Salazar et al. (2024), however it focuses on insurance zone boundaries.

2 Methodology and Research Approach

2.1 Contract Design Method

We consider a farmer that wishes to insure himself against weather risk. The farmer has initial wealth w_0 and experiences a random loss ℓ. The insurer observes an index, θ and makes insurance payouts to the farmer based on θ. Note, here θ can be a vector of raw weather indices (e.g. rainfall, temperature, etc) or it could be an estimate of loss based on weather indices, $\hat{\ell}(\theta)$. For the rest of this document, we will assume that we are using a loss prediction, but our method would also work if given the raw weather indices.

We consider index insurance contracts of the form:

$$I(\theta) = I(\hat{l}(\theta)) \triangleq \min\left\{\max\left\{0, a\hat{\ell}(\theta) + b\right\}, 1\right\}, \tag{1}$$

where a, b are the contract parameters. Without loss of generality we can normalize the the maximum payout to one. We also use $I(\theta)$ instead of $I(\hat{\ell}(\theta))$ for ease of notation. Piecewise-linear contracts are popular both in practice and in the academic literature due to their simplicity and explainability.

We assume that farmers pay a premium $\pi(I)$ for contract I.

$$\pi(I(\theta)) \triangleq \mathbb{E}[I(\theta)] + c_\kappa K(I(\theta)). \tag{2}$$

Here, the premium $\pi(I(\theta))$ is the expected payout plus costs of capital. $K(I(\theta))$ is a function of $I(\theta)$ that represents the amount of capital required to insure the contract, and it is usually set by regulators and meant to ensure that insurers have enough capital to fulfill their contractual obligations with high probability. The cost of capital $K(I(\theta))$ is based on the following commonly

used formula (Mapfumo, Groenendaal, and Dugger (2017):

$$K(I(\theta)) = \mathsf{CVaR}_{1-\epsilon_K}(I(\theta)) - \mathbb{E}[I(\theta)], \tag{3}$$

We are interested in designing contracts that will maximize farmer welfare subject to a price constraint on the insurance. For simplicity, we only include price constraints here, but we could also add constraints on the historical payout frequency of the insurance contract.

Including the premium and the realization of the loss, the farmer's wealth at the end of the season is $w = w_0 - \pi - \ell + I(\theta)$. Our method solves the following optimization problem:

$$\max_{a,b,K,\pi} \quad \mathbb{E}\left[U\left(w_0 - \ell - \pi + \underline{I(\theta)}\right)\right] \tag{4}$$

$$\text{s.t.} \quad \overline{I(\theta)} = \max\left\{0, a\hat{\ell}(\theta) + b\right\}$$

$$\underline{I(\theta)} = \min\left\{a\hat{\ell}(\theta) + b, 1\right\}$$

$$\pi = (1 - c_\kappa)\,\mathbb{E}\left[\overline{I(\theta)}\right] + c_\kappa \mathsf{CVaR}_{1-\epsilon_K}\left(\overline{I(\theta)}\right) \tag{5}$$

$$\pi \leq \overline{\pi}.$$

Our method can be extended to simultaneously design contracts for multiple zones at the same time. This allows us to account cor correlation between zones and make better trade-offs between cost and coverage. The multi-zone model can be found in the appendix.

2.2 Data

Thailand's Department of Agricultural Extension (DOAE) has an extensive dataset of plot level information for farmers in Thailand. This data includes each plot's geographical boundaries as well as information on rice planting activities. Loss data is available from the Government Disaster Relief Program, which provides administrative records of disaster relief payments for rice produc-

tion loss caused by natural disasters. This dataset includes the rice variety planted, the planting date, the harvest date, the total loss, and the date in which the disaster occurred. It has roughly 22 million observations spanning from 2015 until 2022. This dataset's coverage is nationwide. Due to the sensitive nature of the data, it can only be accessed from a secure computer located at the Bank of Thailand. One of the difficulties in conducting research in index insurance is the lack of farmer level yield data, especially in lower or middle income countries. It is worth highlighting that this is a highly unique dataset for the context of index insurance. Thus, this project can inform investment decisions for the governments of similar countries.

2.3 Evaluation

We have evaluated our method using county-level corn yield data from the US Midwest. Using this data, we found that our method outperforms existing methods. Our method leads to larger increases in farmer utility at lower risk for the insurer. However, this data is not representative of agriculture in developing countries. Thus, we hope to supplement this evaluation with more appropriate data.

We plan to evaluate our method using a utility-based framework in line with the previous literature (Clarke (2016); Flatnes, Carter, and Mercovich (2018)). We will also be using a cross-validation approach. We will iterate over the years in our dataset. In each iteration, the data from one year will be used as a test set, while the rest of the data will be used to design the contracts. This will allow us to evaluate contract performance out-of-sample. We will calculate all of the performance metrics of interest for each test year and report the average. We additionally plan to use the data to conduct policy simulations. More specifically, we want to compare the welfare effects of different subsidy policies. We will compare these to a baseline of no insurance, and to a policy alternative of cash transfers with an equivalent cost to the insurance. We are interested in the following metrics: certainty equivalent, insurance premium, demand, and costs for the insurer.

4

3   Budget Narrative and Timeline

I will be using this grant for a 3 month trip to Thailand to work with the Thai loss data. Due to the sensitive nature of the data, it can only be accessed from the Bank of Thailand's facilities in Bangkok. In the first month, my focus will on implementing the single-zone optimization model and the cross-validation evaluation. In the second month, I will focus on running policy simulations and on implementing the multi-zone model. During the last month, I will be focused on generating figures and tables from the results.

Month 1

- **Weeks 1-2:** Generate model cross-validation predictions and implement single-zone optimization model.

- **Weeks 3-4:** Implement insurance program evaluation

Month 2

- **Weeks 1-2:** Run single zone policy simluations

- **Weeks 3-4:** Implement multi-zone optimization model

Month 3

- **Weeks 1-2:** Implement multi-zone cross-validation simulations

- **Weeks 3-4:** Generate tables and figures from results

References

Rockafellar, R Tyrrell, Stanislav Uryasev, et al. (2000). "Optimization of conditional value-at-risk". In: *Journal of risk* 2, pp. 21–42.

Elabed, Ghada et al. (2013). "Managing basis risk with multiscale index insurance". In: *Agricultural Economics* 44(4-5), pp. 419–431.

Greatrex, Helen et al. (2015). "Scaling up index insurance for smallholder farmers: Recent evidence and insights". In: *CCAFS Report*.

Clarke, Daniel J (2016). "A theory of rational demand for index insurance". In: *American Economic Journal: Microeconomics* 8(1), pp. 283–306.

Jensen, Nathaniel D, Christopher B Barrett, and Andrew G Mude (2016). "Index insurance quality and basis risk: Evidence from northern Kenya". In: *American Journal of Agricultural Economics* 98(5), pp. 1450–1469.

Jensen, Nathaniel and Christopher Barrett (2017). "Agricultural index insurance for development". In: *Applied Economic Perspectives and Policy* 39(2), pp. 199–219.

Mapfumo, Shadreck, Huybert Groenendaal, and Chloe Dugger (2017). *Risk modeling for appraising named peril index insurance products: A Guide for practitioners*. World Bank Publications.

Flatnes, Jon Einar, Michael R Carter, and Ryan Mercovich (2018). "Improving the quality of index insurance with a satellite-based conditional audit contract". In: *Unpublished Working Paper*.

Estefania-Salazar, Enrique et al. (2024). *Breaking Boundaries: Lower Tail Dependence Can Triple the Economic Value of Index Insurance for Rural Households*. Tech. rep. National Bureau of Economic Research.

4 Appendix

4.1 Multiple Zone Model

The multiple zone model is very similar to the single zone model. In the objective, we maximize the sum of farmers utilities across all zones. We can think of this as solving the social planner's

problem. The other change is that the budget constraint includes the payouts of all zones, and the required capital is determined using the sum of payouts across all zones.

Model Parameters:

- ϵ_K: This is the epsilon used in the formula for required capital. Recall that the required capital $K(I(\theta)) = \mathsf{CVaR}_{1-\epsilon_K}(I(\theta)) - E[I(\theta)]$.

- c_κ: cost of capital.

- s_z: total insured amount for zone z.

- $w_{0,z}$: this is initial wealth of farmers in zone z.

Model: In the model below, our objective is the average utility of farmers across all zones. The second constraint is the budget constraint, which now includes the sum of payouts across all zones. The formula for required capital was also changed to include the sum of payouts across all zones.

$$\max_{a,b,K,\pi} \quad \mathbb{E}\left[\sum_z U\left(w_{0,z} - \ell_z^j - \pi_z + I_z(\theta_z^j)\right)\right] \tag{6}$$

$$\text{s.t.} \quad \pi_z = \left(1 - \frac{c_\kappa}{\sum_z s_z}\right)\mathbb{E}\left[\overline{I_z(\theta_z)}\right] + \frac{c_\kappa}{\sum_z s_z}\left[\mathsf{CVaR}_{1-\epsilon_K}\left(\sum_z s_z\overline{I_z(\theta_z)}\right) - \mathbb{E}\left[\sum_{z'\neq z} s_{z'}\underline{I_{z'}(\theta_{z'})}\right]\right] \tag{7}$$

$$\overline{I_z(\theta_z)} = \max\left\{0, a_z\hat{\ell}_z(\theta_z) + b_z\right\}$$

$$\underline{I_z(\theta_z)} = \min\left\{a_z\hat{\ell}_z(\theta_z) + b_z, 1\right\}$$

$$\pi_z \leq \overline{\pi_z}.$$

Here, the required capital in (7) is an approximation of

$$K(I(\theta)) = \mathsf{CVaR}_{1-\epsilon_K}\left(\sum_z s_z I_z(\theta_z)\right) - \mathbb{E}\left[\sum_z s_z I_z(\theta_z)\right], \tag{8}$$

and the premium in (7) is an approximation of

$$\pi_z = \mathbb{E}\left[I_z(\theta_z)\right] + \frac{c_\kappa}{\sum_z s_z}\left[\mathsf{CVaR}_{1-\epsilon_K}\left(\sum_z s_z I_z(\theta_z)\right) - \mathbb{E}\left[\sum_z s_z I_z(\theta_z)\right]\right]. \tag{9}$$

We reformulated Problem (6) as a convex program using the results from Rockafellar, Uryasev, et al. (2000). In the model below, p^j is the probability of event j, and j indexes the possible realizations of θ, ℓ. N is the total number of samples.

$$\max_{a,b,\alpha,\omega,\gamma,t_K,K,\pi} \quad \frac{1}{N}\sum_j\sum_z U\left(w_{0,z} - \ell_z^j - \pi_z + I_z(\theta_z^j)\right)$$

$$\text{s.t.} \quad \pi_z = \left(1 - \frac{c_\kappa}{\sum_z s_z}\right)\frac{1}{N}\sum_j \alpha_z^j + \frac{c_\kappa}{\sum_z s_z}\left(K - \frac{1}{N}\sum_j\sum_{z'\neq z} s_{z'}\omega_{z'}^j\right)$$

$$t_K + \frac{1}{\epsilon_K}\sum_j p^j\gamma_K^j \leq K$$

$$\gamma_K^j \geq \sum_z s_z\alpha_z^j - t_K, \forall j$$

$$\gamma_K^j \geq 0, \forall j$$

$$\alpha_z^j \geq a_z\hat{\ell}_z(\theta_z^j) + b_z, \forall j, \forall z$$

$$\alpha_z^j \geq 0, \forall j, \forall z$$

$$\omega_z^j \leq a_z\hat{\ell}_z(\theta_z^j) + b_z, \forall j, \forall z$$

$$\omega_z^j \leq 1, \forall j, \forall z$$

$$\pi_z \leq \overline{\pi_z}, \forall z.$$