Agricultural Index Insurance: An Optimization Approach

José I. Velarde Morales   Linwei Xin

University of Chicago
Booth School of Business

October 22, 2024

Project Overview

- The goal of this project is to improve the design of index insurance contracts. I am particularly interested in the developing country setting.
- The original motivation was to develop a method to simultaneously design contracts for all insured zones, in order to better manage risk. However, upon learning more about the context I realized that the optimization based approach could be an improvement even in the single zone case.
- We develop a method that tries to maximize farmer utility, incorporates different kinds of constraints, and yields interpretable contracts.

Index Insurance Literature

- **Design of Index Insurance:** There has been relatively little research studying the design of index insurance. The method developed by Chantarat et al. (2013) is the most commonly used in academic publications (Jensen et al. (2019); Flatnes, Carter, and Mercovich (2018)). Recently, Chen et al 2023 developed a NN based method to design index insurance Chen et al. (2023).

Content

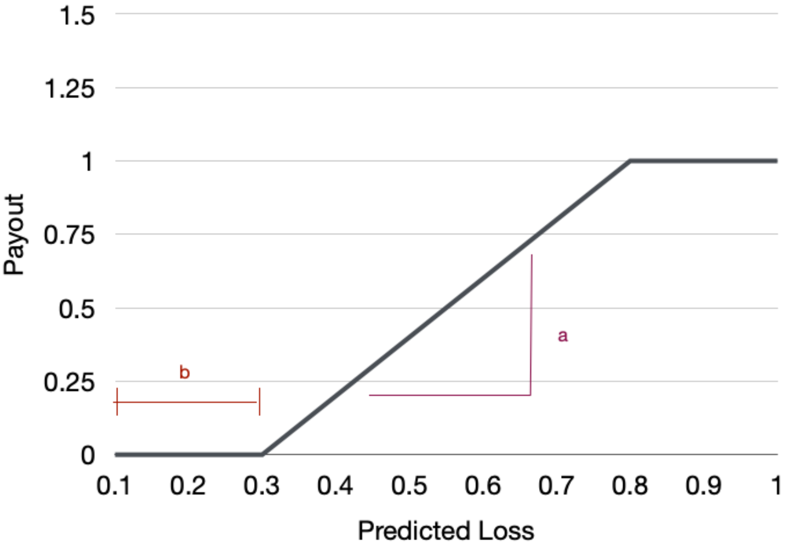
Some methods to design index insurance

- **Baseline Method:** Predict then design, choose contract that maximizes correlation between payouts and losses. Contracts map predicted losses to payouts. Chantarat et al. (2013)
- **NN Based Method:** End to end, use NN to design contracts. Contracts map complex weather data to payouts. Chen et al. (2023)
- **Our Method**: Predict then optimize, design contracts to maximize farmer utility. Contracts map predicted losses to payouts.

Index Insurance: Definition and Parameters

- Index insurance uses a signal, θ, to predict agricultural loss, $\hat{\ell}(\theta)$
- Contract form: $I(\theta) = \min\left\{\max\left\{a\hat{\ell}(\theta) - b, 0\right\}, 1\right\}$, a, b are the contract parameters.

Example of Index Insurance Contract



Content

Model

Farmers start with wealth, w_0 and experience loss ℓ. There is an index insurance contract I, that is determined by a p dimensional vector of indices $\theta = (\theta_1, ..., \theta_p)$. Premium for contract I is $\pi(I)$. Farmer wealth is:

$$w = w_0 - \ell + I(\theta) - \pi(I)$$
$$\pi(I) = \lambda \mathbb{E}[I(\theta)]$$

Optimization Problem

They solve the following optimization problem:

$$\max_{I \in \mathcal{I}} \quad \mathbb{E}\left[U\left(w_0 - \ell - \pi + I(\theta)\right)\right]$$

$$\text{s.t.} \quad \underline{\pi} \leq \pi(I) \leq \overline{\pi} \tag{1}$$

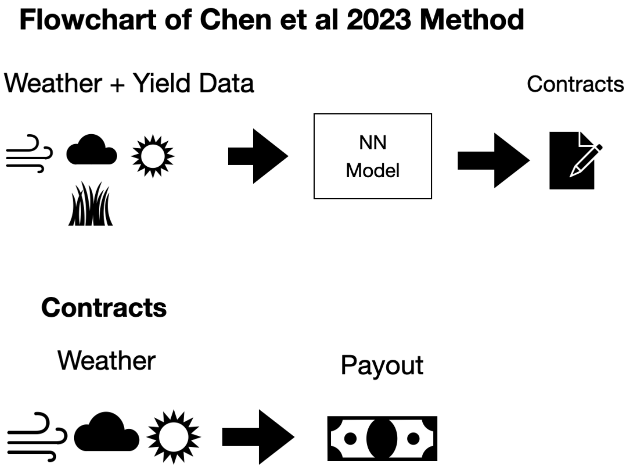
$$\pi = \lambda \mathbb{E}\left[I(\theta)\right] \tag{2}$$

$$U\left(w\right) = -(1/\alpha)e^{-\alpha w} \tag{3}$$

Solution and results

- They use a neural network to solve the optimization problem. Network they use has $\sim 10,000$ parameters.
- They test it on Illinois corn yield data and find that the contracts are utility improving on average. They also apply their Illinois model to other states in the US Midwest and find that it's utility improving.
- Their method is end-to-end, it goes directly from weather variables to payouts.

Chen Method Flowchart



Drawbacks

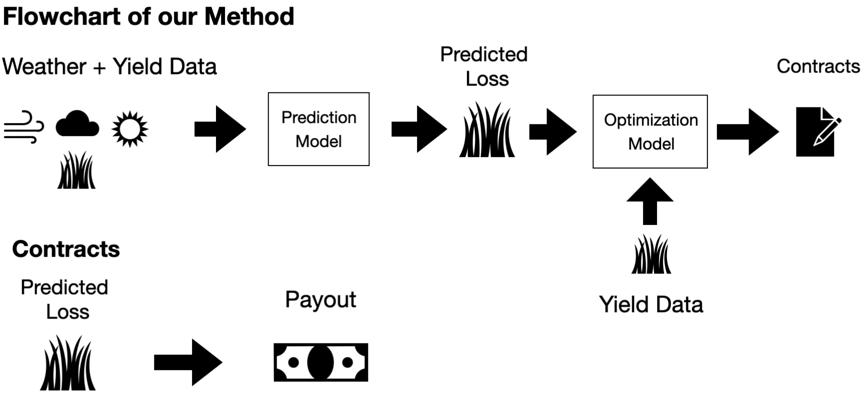
- Not interpretable, can't set constraints on variables that policy makers might care about (e.g. deductible). This also makes it hard to debug: is the performance due to poor prediction performance? Too high a deductible?
- Model requires a lot of data to train, the data they tested the model on went back to 1925, unrealistic for developing countries.
- Definition of premium used depends only on expected value of the payout and not on variance of the payouts, in practice the price would depend on riskiness of contract.

Content

Overview

- We opt for a "predict-then-optimize" approach.
- We use specialized time-series feature extraction algorithms for feature extraction and traditional ML algorithms (e.g. Random Forest, Gradient Boosting, Support Vector Machines)
- These algorithms require less data to train.

Our Method Flowchart



Premium Principles

Premium principles are the functionals used to map insurance contracts to premiums (prices). The exact pricing method used by insurers is considered to be proprietary information. However, there are several premium principles that are common in the actuarial literature, and our method is compatible with most of them:

Description	Definition
Expected value	$(1+\alpha)\mathbb{E}[I]$
Standard deviation	$\mathbb{E}[I] + \alpha\sigma(I)$
Variance	$\mathbb{E}[I] + \alpha\sigma^2(I)$
Exponential	$\frac{1}{\alpha}\log \mathbb{E}[e^{\alpha I}]$
Dutch	$\mathbb{E}[I] + \beta\mathbb{E}[(I - \alpha\mathbb{E}[I])_+]$

Model

We use the same model as Chen, just a different definition of the premium. We opt for a definition of the premium that captures the cost of having payouts with a big right tail.

- Chen Premium: $\pi(I(\theta)) = \lambda \mathbb{E}[I(\theta)]$
- Our Premium: $\pi(I(\theta)) = \mathbb{E}[I(\theta)] + c_\kappa K(I(\theta))$, where c_κ is the cost of capital, and K is required capital.
- $K(I(\theta)) = CVaR_{1-\epsilon_P}(I(\theta)) - \mathbb{E}[I(\theta)]$.

I got this definition from *Risk Modeling for Appraising Named Peril Index Insurance Products: A Guide for Practitioners* (Mapfumo, Groenendaal, and Dugger (2017)). The EU's Solvency requirements are similar.

Idealized Model

$$
\begin{aligned}
\max_{a,b,\pi,K} \quad & \mathbb{E}\left[U\left(w_0 - \ell - \pi + I(\theta)\right)\right] && (4)\\
\text{s.t.} \quad & I(\theta) = \min\left\{\max\left\{0, a\hat{\ell}(\theta) + b\right\}, 1\right\} && (5)\\
& \pi = \mathbb{E}\left[I(\theta)\right] + c_{\kappa}\left[\mathrm{CVaR}_{1-\epsilon_K}\left(I(\theta)\right) - \mathbb{E}[I(\theta)]\right] && (6)\\
& \pi \leq \overline{\pi}. && (7)
\end{aligned}
$$

Convex Relaxation

We use the following convex and concave relaxations of $I(\theta)$:

$$\overline{I(\theta)} \triangleq \max\left\{0, a\hat{\ell}(\theta) + b\right\}, \qquad \underline{I(\theta)} \triangleq \min\{a\hat{\ell}(\theta) + b, 1\}.$$

Note that $\overline{I(\theta)}$ is convex in $\hat{\ell}(\theta)$, $\underline{I(\theta)}$ is concave in $\hat{\ell}(\theta)$, and

$$\underline{I(\theta)} \leq I(\theta) \leq \overline{I(\theta)}.$$

We replace $I(\theta)$ in Problem (4) with either $\overline{I(\theta)}$ or $\underline{I(\theta)}$ where necessary to obtain a conservative and convex relaxation.

Convex Relaxation

$$
\begin{aligned}
\max_{a,b,K,\pi} \quad & \mathbb{E}\left[U\left(w_0 - \ell - \pi + \underline{I(\theta)}\right)\right] && (8)\\
\text{s.t.} \quad & \overline{I(\theta)} = \max\left\{0, a\hat{\ell}(\theta) + b\right\} \\
& \underline{I(\theta)} = \min\left\{a\hat{\ell}(\theta) + b, 1\right\} \\
& \pi = (1 - c_\kappa)\,\mathbb{E}\left[\overline{I(\theta)}\right] + c_\kappa \mathrm{CVaR}_{1-\epsilon_K}\left(\overline{I(\theta)}\right) && (9)\\
& \pi \leq \overline{\pi}.
\end{aligned}
$$

CP Reformulation

Using the results from Rockafellar and Uryasev (2002), we get:

$$
\begin{aligned}
\max_{a,b,\alpha,\omega,\gamma,t_K,K,\pi} \quad & \frac{1}{N}\sum_j \sum_z U\left(w_{0,z} - \ell_z^j - \pi_z + I_z(\theta_z^j)\right) \\
\text{s.t.} \pi_z &= \left(1 - \frac{c_\kappa}{\sum_z s_z}\right)\frac{1}{N}\sum_j \alpha_z^j + \frac{c_\kappa}{\sum_z s_z}\left(K - \frac{1}{N}\sum_j \sum_{z' \neq z} s_{z'}\omega_{z'}^j\right) \\
& t_K + \frac{1}{\epsilon_K}\sum_j p^j \gamma_K^j \leq K \\
& \gamma_K^j \geq \sum_z s_z \alpha_z^j - t_K, \forall j \\
& \gamma_K^j \geq 0, \forall j \\
& \alpha_z^j \geq a_z \hat{\ell}_z(\theta_z^j) + b_z, \forall j, \forall z \\
& \alpha_z^j \geq 0, \forall j, \forall z \\
& \omega_z^j \leq a_z \hat{\ell}_z(\theta_z^j) + b_z, \forall j, \forall z \\
& \omega_z^j \leq 1, \forall j, \forall z \\
& \pi_z \leq \overline{\pi_z}, \forall z.
\end{aligned}
$$

Content

Data

Used two main data sources

- Illinois annual corn yield data from the National Agricultural Statistics Service (NASS). Data is available at the county level from 1925-2022. 84 counties.
- Weather data from the PRISM climate group. Has monthly data on several weather variables (temperature, precipitation, etc). Available 1895-present.

- We use a 70/15/15 train/val/test split. Data is kept in chronological order. Training data has older years and test data has the newest years.

- We modified Chen's method to use the same definition of the premium as our method.

- We use the training and validation data to design the contracts using both methods, apply the contracts to farmers in the test set, and compute performance metrics.

- We used a data shortening exercise to evaluate how the performance of both methods changed as more data became available.

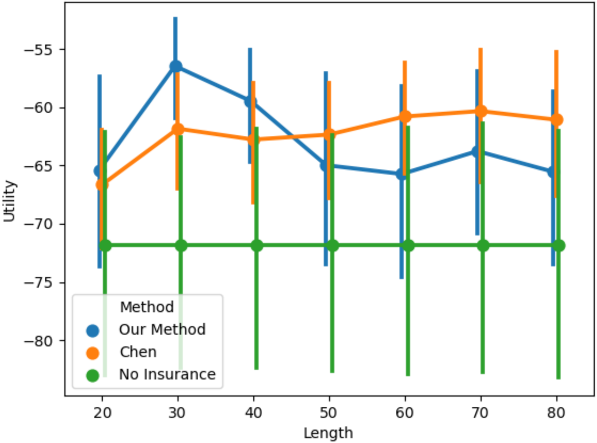
Overview

- Our method performs similarly or outperforms the Chen model when there is less than 50 years of data available.

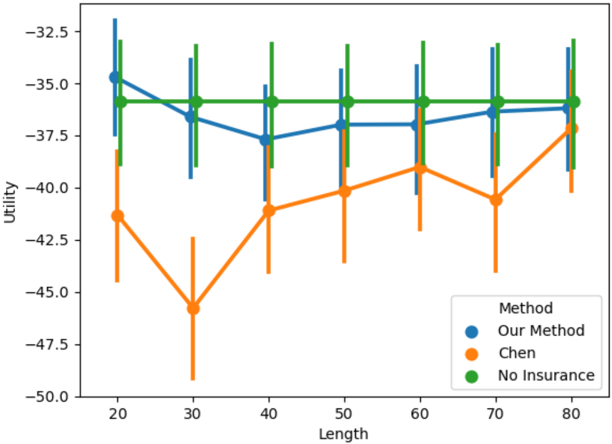
- In terms of farmer utility, our method tends to work better with realistic data lengths. Most satellite data starts at 1980 at the earliest.

- When using data from other states, our method consistently outperforms Chen's method, but is not always better than no insurance, at least at the full premium price. This might not be a huge problem, since agricultural insurance tends to be heavily subsidized, both in rich and poor countries. In the US, it averaged 62% of premiums in 2022.

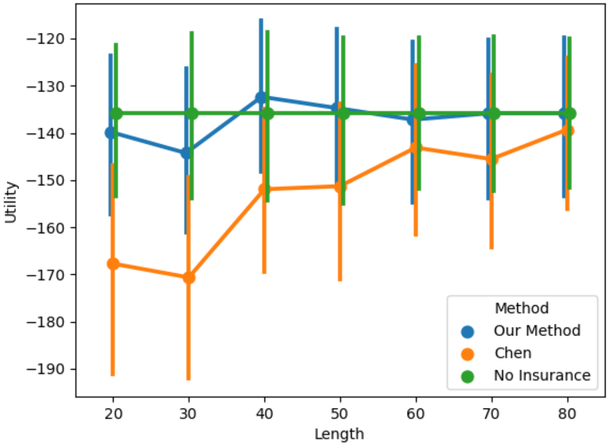
Illinois: Utility



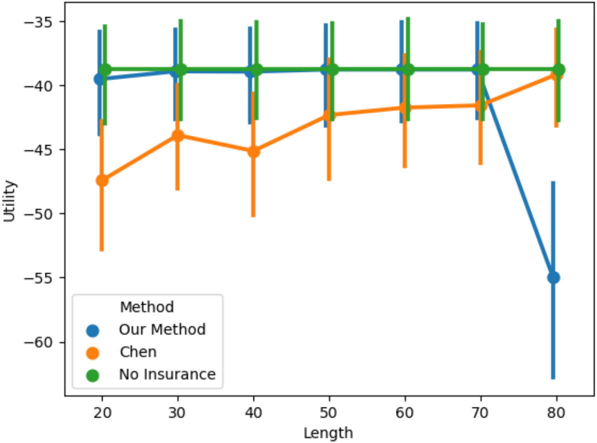
Iowa: Utility



Missouri: Utility



Indiana: Utility



Content

Multiple Zone Model

- Our method can be extended to design the contracts of multiple zones simultaneously.

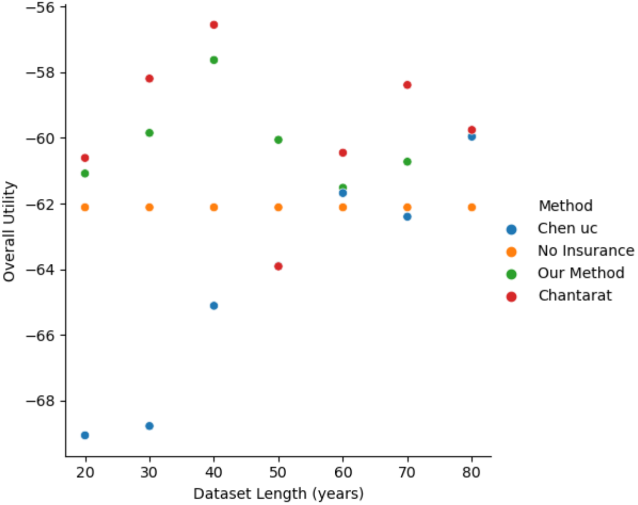
- This allows it to take into account the correlation between the insured zones, allowing it to manage risk better.

$$
\begin{aligned}
\max_{a,b,K,\pi} \quad & \mathbb{E}\left[\sum_z U\left(w_{0,z} - \ell_z^j - \pi_z + I_z(\theta_z^j)\right)\right] && (10)\\
\text{s.t.} \quad & \pi_z = \mathbb{E}\left[\overline{I_z(\theta_z)}\right] + \frac{c_\kappa}{\sum_z s_z} K && (11)\\
& K = \mathrm{CVaR}_{1-\epsilon_K}\left(\sum_z s_z \overline{I_z(\theta_z)}\right) - \mathbb{E}\left[\sum_z s_z \underline{I_z(\theta_{z'})}\right] && (12)\\
& \overline{I_z(\theta_z)} = \max\left\{0, a_z \hat{\ell}_z(\theta_z) + b_z\right\}\\
& \underline{I_z(\theta_z)} = \min\left\{a_z \hat{\ell}_z(\theta_z) + b_z, 1\right\}\\
& \pi_z \leq \overline{\pi_z}.
\end{aligned}
$$

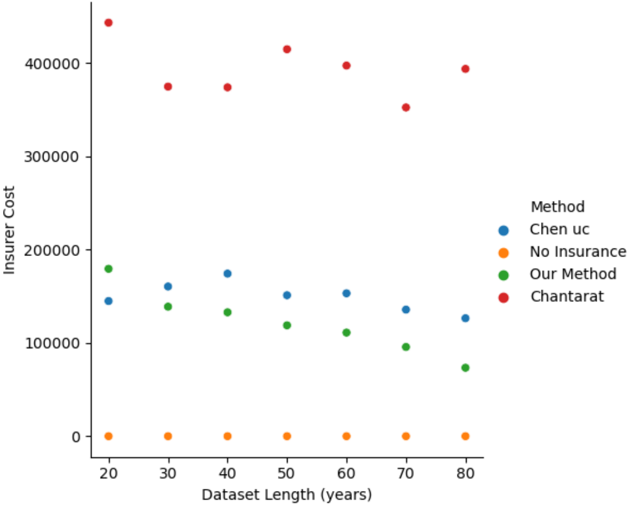
Overview

- Our multiple zone model adjusts contracts based on the correlation between the insured zones. In this case, it leads to contracts that pay out more frequently, but at a lower rate. This reduces the tail risk for the insurer and reduces the amount of capital needed.
- It outperforms Chen's model and the no insurance case consistently, and has lower costs and required capital than the Chen model.
- We also wanted to compare it to using our single zone model. In the following figures, Our Method: SZ refers to using our single zone method to design the contract of each state individually, but then calculating the premium as if it was a part of the portfolio.

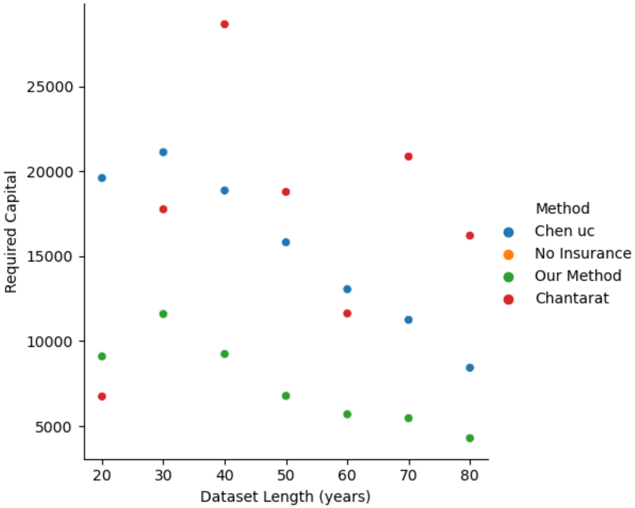
Midwest: Overall Utility



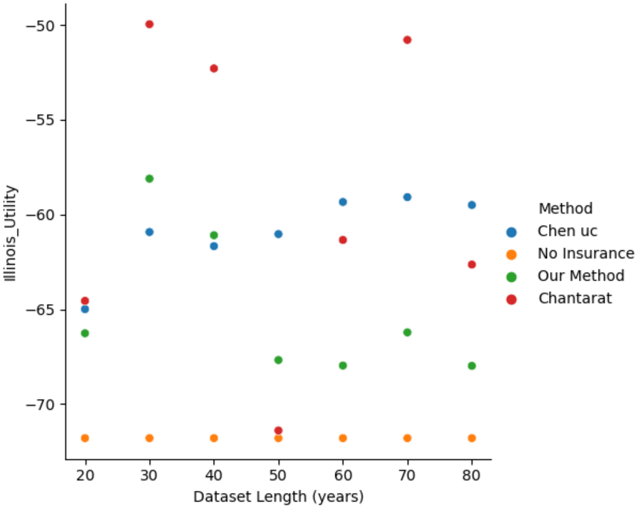
Midwest: Insurer Cost



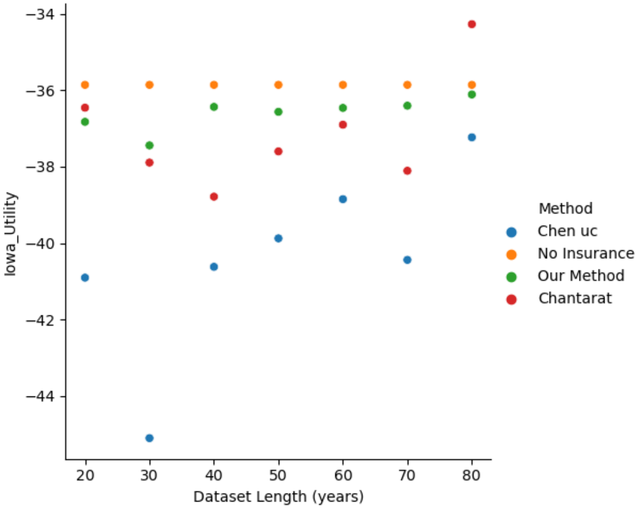
Midwest: Required Capital



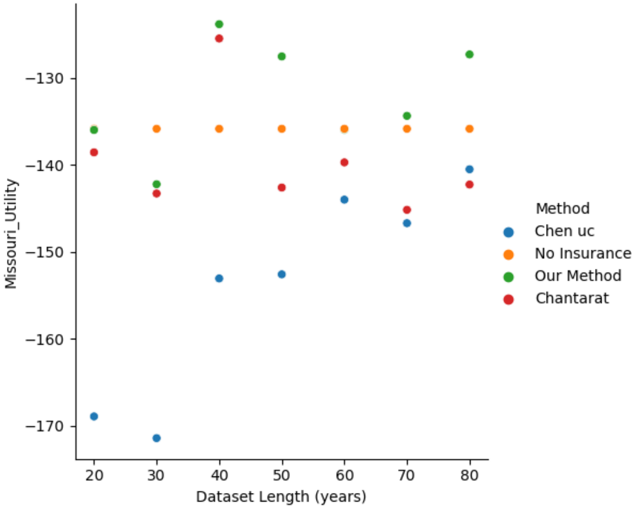
Illinois: Utility



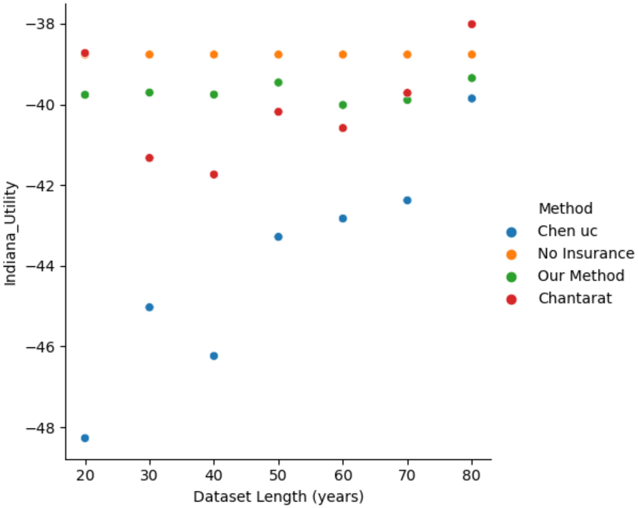
Iowa: Utility



Missouri Utility



Indiana Utility



Things I was thinking about adding

- Deep dive as to why my method does better, and when it does better?
- Run the same evaluation but with shuffled data instead of having it orderd.
- I wanted to do more of a deep dive on the benefits of taking a portfolio approach, but not sure how beyond just looking at capital requirements. Maybe using an expected return and variance approach?
- A robustness check on the stability of the solutions at different dataset lengths.

Questions

- What else should I report?
- How can I strengthen this evaluation?
- Any robustness checks I should add?
- I feel like I should do a deep dive as to why the method does

Thai Data

- Plot-level losses caused by natural disasters between 2015-2022.
- I can access the raw data, and they can send me aggregate versions as well. There are around 7000 counties in Thailand, and around 80,000 villages, the data can be aggregated at both of those levels.

Loss Definition

From what I can tell from the replication files, they seem to define loss in every year as:

$$\ell_{st} = R_s^* - R_{st}$$

where R_s^* is the maximum revenue observed in state s across all time periods, and R_{st} is the revenue in state s at time t.

Initial wealth

According to the paper, they set $w_0 = 389$. However, in the replication files, they set it to be $w_0 = 813 - 504 + 389$. According to the comments, 504 is the fixed cost of operating a farm, and there are no comments regarding the 813, I'm assuming it corresponds to R^*.

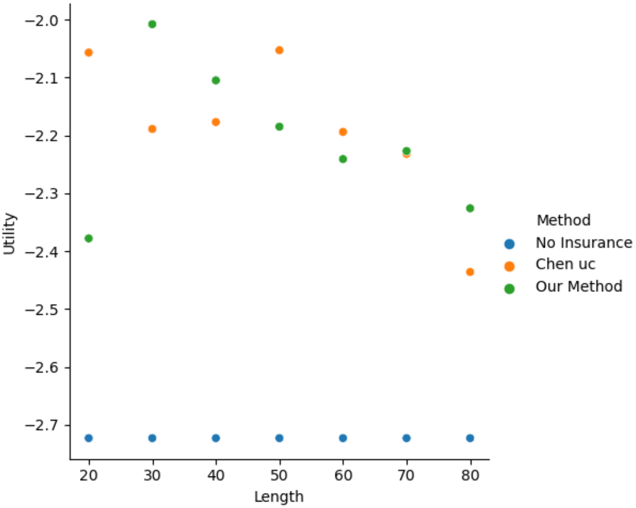
Detrending

- According to the paper, they detrend the county level yield data using a 2nd order polynomial fit with a "robust" regression method, but they don't specify what they use, and it's not in the replication files. They also don't specify if they remove the trend using additive or multiplicative decomposition model. Using an additive decomposition model yielded the most similar losses to what they provide in the replication files.
- There are a couple of papers showing that using locally weighted regression to detrend works better.

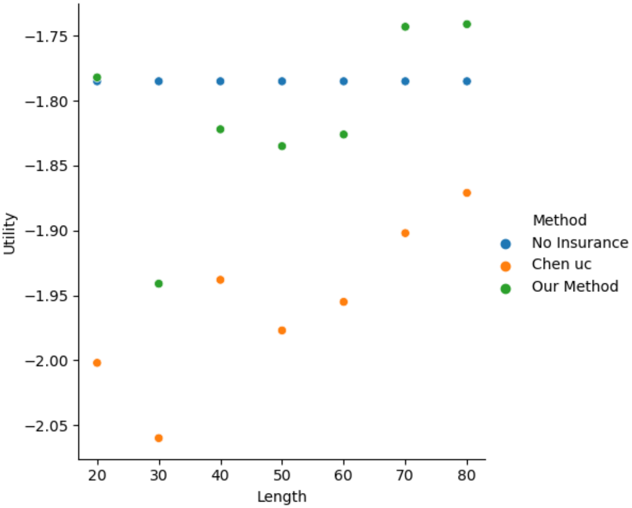
Questions

- Would it make sense to define loss as deviation from historical average? Allowing it to be positive in some years? In other words, we would first adjust all of the yield data to 2020 levels and then calculate the historical average. The loss in each year would be the deviation from this historical average.
- Should I simply follow their lead on detrending? Or should I try to improve on it?
- Do you think it's necessary to show results with both definitions of the premium?
- Do you think subsidy vs lump sum results would be interesting?

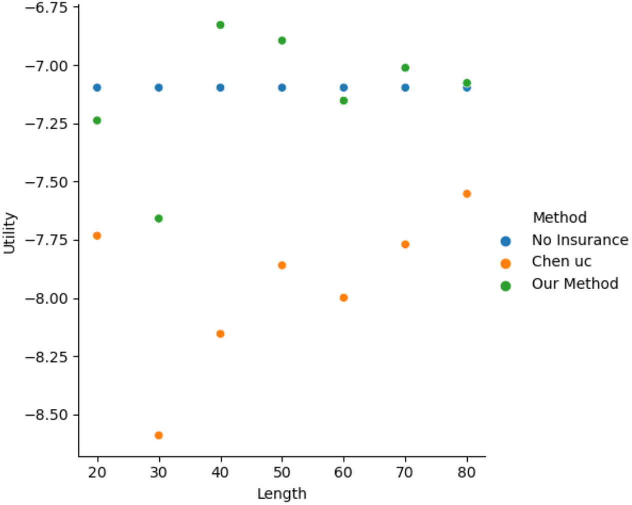
Their Defn of Premium: Illinois Utility



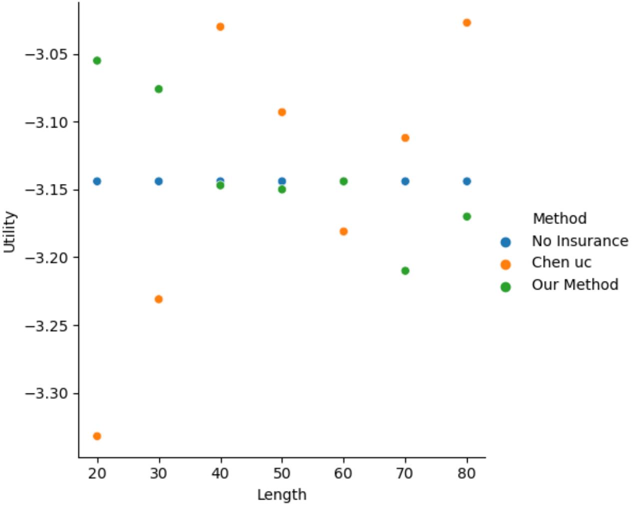
Their Defn of Premium: Iowa Utility



Their Defn of Premium: Missouri Utility



Their Defn of Premium: Indiana Utility



References

Rockafellar, R Tyrrell and Stanislav Uryasev (2002). "Conditional value-at-risk for general loss distributions". In: *Journal of banking & finance* 26(7), pp. 1443–1471.

Chantarat, Sommarat et al. (2013). "Designing index-based livestock insurance for managing asset risk in northern Kenya". In: *Journal of Risk and Insurance* 80(1), pp. 205–237.

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*.

Jensen, Nathaniel et al. (2019). "Does the design matter? Comparing satellite-based indices for insuring pastoralists against drought". In: *Ecological economics* 162, pp. 59–73.

Chen, Zhanhui et al. (2023). "Managing weather risk with a neural network-based index insurance". In: *Management Science*.

Idealized CVaR Model

- **Objective:** conditional value at risk of the farmers' loss net of insurance.
- **Constraint 1:** piecewise linear structure of the contract.
- **Constraint 2:** budget constraint.
- **Constraint 3:** definition of required capital.

$$\min_{a,b,\pi,K} \quad CVaR_{1-\epsilon}\left(\ell - I(\theta)\right)$$

$$\text{s.t. } I(\theta) = \min\{(a\hat{\ell}(\theta) + b)^{+}, P\} \tag{13}$$

$$\mathbb{E}\left[I(\theta)\right] + c_{\kappa} K \leq B \tag{14}$$

$$K = \left(CVaR_{1-\epsilon}\left(I(\theta)\right) - \mathbb{E}[I(\theta)]\right) \tag{15}$$

The problem is non-convex, so we need convex approximations

We use the following approximations of $I(\theta)$ to make the problem convex:

$$\overline{I(\theta)} \triangleq \max\left\{0, a\hat{\ell}(\theta) + b\right\}$$

$$\underline{I(\theta)} \triangleq \min\{a\hat{\ell}(\theta) + b, K\}$$

- Note that $\overline{I(\theta)} \geq I(\theta)$ and $\overline{I(\theta)}$ is convex. Conversely, $\underline{I(\theta)} \leq I(\theta)$ and $\underline{I(\theta)}$ is concave.
- We replace $I(\theta)$ with either $\overline{I(\theta)}$ or $\underline{I(\theta)}$ where necessary to obtain conservative and convex approximations.
- We also need approximations or proxies for $E[I(\theta)]$ in constraint . We use $\pi_{SQ} = E[I_{SQ}(\theta)]$, where I_{SQ} is the contract designed using the status quo method, as a proxy for $E[I(\theta)]$ in constraint .