Managing Weather Risk with a Neural Network-Based Index Insurance

Zhanhui Chen

School of Business and Management, Hong Kong University of Science and Technology

Yang Lu

Concordia University

Jinggong Zhang

Nanyang Business School, Nanyang Technological University

Wenjun Zhu

Nanyang Business School, Nanyang Technological University

Nanyang Business School Research Paper No. 20-28

Nanyang Business School
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798
https://nbs.ntu.edu.sg/

Managing Weather Risk with a Neural Network-Based Index Insurance

Zhanhui Chen* Yang Lu† Jinggong Zhang‡ Wenjun Zhu§

Abstract

Weather risk affects economy, agricultural production in particular. Index insurance has been proposed to hedge against severe weather risk, but large basis risk and low demand accompany current piecewise-linear index insurance contracts. We propose embedding a neural network-based optimization scheme into an expected utility maximization problem to design the index insurance. Neural networks capture the highly nonlinear relationship between the high-dimensional weather variables and production losses. We endogenously solve for the optimal premium and demand of this index insurance. We test this index insurance using corn production loss and weather data in Illinois. We show that this index insurance reduces basis risk, increases farmers' utility, is less reliant on government subsidies, and improves social welfare.

Keywords: neural networks; weather risk; index insurance; basis risk; utility maximization

JEL code: C45, G11, G22, G52, O13, Q54

*Corresponding author: Zhanhui Chen, Department of Finance, School of Business and Management, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Tel.: +852-2358-7670; Fax: +852-2358-1749; E-mail: chenzhanhui@ust.hk.

†Concordia University, Montreal, QC, Canada. E-mail: yang.lu@concordia.ca.

‡Division of Banking & Finance, Nanyang Business School, Nanyang Technological University, Singapore 639798. E-mail: jgzhang@ntu.edu.sg.

§Division of Banking & Finance, Nanyang Business School, Nanyang Technological University, Singapore 639798. E-mail: wjzhu@ntu.edu.sg.

1 Introduction

Climate change and weather risk affect economy and livelihood over large scales (Nordhaus, 2019; Hong et al., 2020), especially for agricultural production which uses weather conditions as direct inputs (Schlenker and Roberts, 2009; Fisher et al., 2012; Lesk et al., 2016). For example, USDA (2014) estimates that 70%-90% of agricultural production loss can be attributed to adverse weather. In practice, weather insurance is often employed to hedge against weather risk.

Yet, conventional insurance contracts are often associated with high administrative costs, lengthy settlement processes, and severe adverse selection and moral hazard issues, all of which make conventional insurance less desirable for hedging against weather risk. To address these concerns, insurers have marketed index insurance to farmers who wish to hedge against extreme and systemic weather risk. The key difference between index insurance and conventional indemnity-based insurance is that the payoff of the latter solely depends on the actual losses incurred to the insureds. In contrast, the payoff of index insurance is exclusively based on some prespecified indices. For example, weather index insurance determines the claim payments based on future realizations of weather events determined from certain weather indices. While cost effective and promising, current index insurance faces low demand and sustainability issues (Cole et al., 2013; Clarke, 2016).[1] Governments often have to significantly subsidize the insurance premium to incentivize farmers, creating financial burdens to governments.[2] Low demand is largely due to basis risk (see, e.g., Cummins et al., 2004; Clarke, 2016; Jensen et al., 2016), the risk that the underlying indices and actual losses are mismatched. This calls for a better index insurance design to minimize basis risk, which is critical to improve the demand for index insurance.

There are two difficulties in reducing basis risk. First, to improve the performance of the index, we should include a sufficiently large number of weather variables when constructing the contract. However, this leads to the curse of dimensionality. Second, because

[1]For example, in Canada, participation rates for index insurance, such as forage index insurance, are as low as 10%.

[2]For example, the U.S. federal crop insurance program costs $74 billion in 2008-2017, 60% of which benefited farmers and 40% went to private insurance companies (Rosa, 2018a).

of biological reasons, crop production depends on weather conditions in a highly nonlinear way (Schlenker and Roberts, 2009; Roberts et al., 2013; Zhang et al., 2018; Rigden et al., 2020). Therefore, it is challenging to model the index insurance payoff functional form, which is highly nonlinear and nonmonotonic over high-dimensional weather variables. Because of these difficulties, most current index insurance products are based on low dimensional indices and use linear-type payoff functions, e.g., a single-index piecewise-linear contract (Mahul and Skees, 2007; Giné et al., 2007), which incurs a large basis risk. Panel (a) of Figure 1 plots the payoffs of a currently used linear index insurance which is based on a rainfall index, exhibiting a large basis risk.

In this paper, we address these difficulties by proposing a neural network-based (NN-based) index insurance design, tapping into recent advances in machine learning.[3] Neural networks help to capture high-dimensional, nonlinear, and complex interactions between weather indices and production loss. We embed an NN-based scheme into an expected utility maximization problem with budget constraints. We first use a penalty method to transform the constrained optimization problem into an unconstrained one and then solve the model with a backpropagation algorithm. A novel feature is that we endogenously determine the equilibrium insurance premium, using the historical insurance supply curve. This is important when we maximize the expected utility of farmers, because it incorporates the responses of insurers. This framework provides an optimal index insurance contract without any assumptions about (1) the dependence structure among variables (i.e., production loss variable and weather variables) and (2) the functional form of the insurance contract payoff. Our framework integrates NNs into expected utility maximization. In contrast, existing literature about index insurance design with machine learning

[3] Other machine learning techniques might be useful, e.g., reinforcement learning. We opt to the neural network approach for two reasons. First, NN provides satisfying performance and computing speeds (see, e.g., Gu et al. (2020a), Bianchi et al. (2020), and Feng et al. (2021)). Second, we assume a representative farmer in a fully rational expectation framework and the model is a static one, so there is not much benefits for farmers to learn the dynamics of weather conditions. In fact, current cutting-age atmospheric models can forecast weather only up to 10 to 15 days (Voosen, 2019; Zhang et al., 2019). This means weather conditions and hence production losses are impossible to be forecasted one-year ahead. Therefore, farmers cannot learn much about the intertemporal dependence of weather conditions and "time the index insurance market" at the annual frequency. Nevertheless, reinforcement learning could be helpful if we consider farmers with bounded rationality, e.g., farmers are short-sighted and only learn from the recent experiences. We avoid designing the insurance contract based on behavioral bias as this might introduce some litigation risk in practice.

techniques mostly uses restrictive piecewise-linear contracts and doesn't incorporate utility maximization analysis.

We explore various NN structures and select the optimal neural network with three hidden layers (64, 64, and 16 neurons in each layer). We apply this NN-based index insurance to corn production in Illinois. Results show that the NN-based index insurance policies are effective in improving farmers' utility because this contract reduces their exposure to wealth variations, especially the extreme downside risk. For example, using a constant absolute risk aversion utility (CARA), we see that the optimal NN-based index insurance improves farmers' utility by 12.33% for the test sample. Certainty equivalent wealth (CEW) improves by $19,345.20 for the test sample for an average farm. The optimal NN-based index insurance has a premium of $28.82/acre, which is 63% of the current corn premium ($45.50/acre) in Illinois, making the contract very attractive. The demand of the optimal NN-based index insurance is three times as that of the currently used contract. Panel (b) of Figure 1 plots the payoffs of the optimal NN-based index insurance policy. Comparing Panels (a) and (b), we see that the optimal NN-based contract effectively reduces basis risk. In fact, the payoff function of our proposed NN-based index insurance is very close to a conventional indemnity-based insurance with a deductible (i.e., a stop-loss insurance).

We dig deeply in two ways to understand the economic mechanism of the superior performances of the NN-based index insurance contract. First, following Cong et al. (2020), we conduct a polynomial sensitivity analysis to improve model transparency and interpretability. In particular, we explore the sensitivities of insurance payoffs to various weather indices by adopting the gradient-based characteristic importance method. Interestingly, we find that some weather indices outside growing seasons (the growing season in Illinois is normally from May to October for corns) are important, e.g., dew point temperature and maximum vapor pressure deficit in January, whereas these variables are overlooked by the linear insurance contract in current practice. This demonstrates the power of neural networks to extract important information from a large set of input factors. In addition, the sensitivity analyses confirm the nonlinear relationship between insurance payoffs and weather indices, including the interactions among weather indices.

Abundant evidence in the agronomy literature supports such nonlinear relationship (see, e.g., Schlenker and Roberts, 2009; Lobell et al., 2011; Rigden et al., 2020).

Second, we compare the performances of NN-based index insurance with other insurance contracts that are based on polynomial payoff functions, such as linear polynomials (the common practice nowadays), quadratic polynomials, and cubic polynomials. We also consider contracts using different numbers of weather indices as input factors. We find that including more weather variables and nonlinear terms improve farmers' utility. These results suggest the impressive performances of NN-based contract are indeed rooted in the contract's capability to extract complicated information from a large set of inputs, which are often nonlinear and nonmonotonic.

We further investigate the robustness and flexibility of NN-based index insurance in six ways. First, we perform an "out-of-state" tests, using three states adjacent to Illinois that have similar latitudes, namely, Indiana, Kentucky, and Missouri. That is, we use Illinois data as the training and validation samples and data from the three adjacent states as the test sample. We find our proposed NN-based index insurance is powerful enough to improve farmers' utility in these three contiguous states, even without using their data to train the model. Second, we consider farmers with different coverage demands and risk aversion. We also allow for time-varying risk aversion which depends on the loss experiences in the previous year. Third, we consider different insurance supply curves. Fourth, we consider the impacts of government subsidies. Fifth, we extend the contract to revenue insurance, that is, considering both production risk and corn price risk. Finally, we consider alternative utility functions, e.g., a power or logarithmic utility. We show that the NN-based index insurance performs robustly over these variations.

This paper contributes to the growing literature on the use of big data and machine learning in finance. Advances in machine learning enable us to process big and complex data and tackle high-dimensional problems in finance. Various machine learning algorithms have been examined to identify factors contributing to asset returns. For example, using NN-based language processing and generative statistical modeling, Cong et al. (2019) generate textual factors with enhanced interpretability that could be widely applied to different finance and economic problems. Cong et al. (2020) apply reinforcement learning to capture the high-dimensionality, nonlinearity, and dynamic nature of data to construct the optimal portfolio. Cong et al. (2021) assess the performances of different deep sequence models in asset pricing. Chinco et al. (2019), Freyberger et al. (2020), Gu et al. (2020b), Bianchi et al. (2020), Bryzgalova et al. (2020), Chen et al. (2020), Choi et al. (2020b), Gu et al. (2020a), Feng et al. (2020), and Feng et al. (2021) use various machine learning approaches (e.g., neural networks, decision trees, generative adversarial networks, autoencoder neural networks, least absolute shrinkage and selection operator, and principal component regressions) to link stock or bond returns with lots of predictors (e.g., pricing factors or firm characteristics). Hurlin et al. (2018) conduct a horse race between various machine learning methods for loss given default modeling in credit risk. Sirignano et al. (2018) propose to use deep neural networks to study the nonlinear relationship between the various economic factors and mortgage risk. Sirignano and Giesecke (2019) prove a law of large numbers and a central limit theorem for a broad class of statistical and machine learning models to analyze risk of large pools of loans. This paper highlights the technical benefits and economic gains of applying neural networks to the insurance contract design context.

This paper also belongs to the climate finance literature (see Hong et al. (2020) for a comprehensive overview). Prior literature investigates the pricing of climate risk and its impacts on investment (Hong et al., 2019; Baldauf et al., 2019; Murfin and Spiegel, 2019; Bernstein et al., 2019; Painter, 2020; Choi et al., 2020a; Krueger et al., 2020). Some papers explore financial solutions to manage weather risk. For example, Engle et al. (2019) consider climate risk hedging with mimicking portfolios. Pérez-González and Yun (2013) show that active risk management activities using weather derivatives increase the firm values of the weather-sensitive firms. Weagley (2019) shows the effects of financial sector stress on risk sharing, using weather derivative contracts. Our paper adds to the literature by proposing a new design of the weather-related insurance contract to manage weather risk.

This paper also contributes to the index insurance literature. Clarke (2016) shows theoretically that basis risk causes the low demand of index insurance. Bryan (2019) and Hartman-Glaser and Hébert (2020) explore alternative reasons for the failure of index insurance market, e.g., asymmetric information. Jensen et al. (2016) examine basis risk of livestock index insurance policies in northern Kenya. Assa and Wang (2020) design index insurance on agricultural goods which provides a high Sharpe ratio. Several studies conduct field experiments to investigate the demand and efficiency of index insurance (see, e.g., Cole et al., 2013; Karlan et al., 2014; Cole et al., 2017; Casaburi and Willis, 2018). We contribute to the literature by constructing an NN-based index insurance to reduce basis risk.

The rest of this paper proceeds as follows. Section 2 introduces the model and assumptions. Section 3 discusses the NN-based approach for solving the optimal index insurance contract. Section 4 presents the empirical performance of the NN-based insurance policy. Section 5 provides robustness checks for key policy parameters. Section 6 discusses several extensions of the model, e.g., adding government subsidies, considering corn price risk, and using alternative utility functions. Section 7 concludes.

2 The model

Consider a typical farmer who would like to hedge her exposure to weather risk, subject to certain budget constraints. Suppose the farmer has an initial wealth of w_0 and faces a random production loss of Y during a year. We assume there is an index insurance contract with a payoff determined by a p-dimensional random vector of indices, $\boldsymbol{X} = (X_1, X_2, ..., X_p)$. More specifically, the index insurance payoff is $I(\boldsymbol{X})$, where $I : \mathbb{R}^p \mapsto \mathbb{R}^+$ is the nonnegative payoff function. Denote the premium of the index insurance contract by $\pi(I)$, which is a functional of the indemnity function I. Assume that the index insurance could be subsidized by the government for a proportion of θ. Then the terminal wealth of the farmer at the presence of the index insurance is $w_0 - Y + I(\boldsymbol{X}) - (1-\theta)\pi(I)$. Our objective for the index insurance design is to select the optimal indemnity function I so that the policyholder's expected utility, $\mathrm{E}\,(U)$, is maximized under certain budget constraints. In this paper, we focus on designing index insurance against production losses, without taking into account crop price risk, that is, a yield insurance. We will extend this framework to revenue protection in Section 6, which considers crop price risk. Specifically, the index insurance design problem can be formulated as the following expected utility maximization problem:

$$\begin{cases} \sup\limits_{I \in \mathcal{I}} & \mathrm{E}\big(U[w_0 - Y + I(\boldsymbol{X}) - (1-\theta)\pi(I)]\big) \\ \text{s.t.} & P_L \leq \pi(I) \leq P_U, \end{cases} \tag{1}$$

where $\mathcal{I} := \{I : \mathbb{R}^p \mapsto \mathbb{R}^+ | I \text{ is measurable}\}$ defines the feasible set of indemnity function I. The budget constraint of the farmer is given by P_L and P_U, that is, the lower and upper bounds of the premium level. P_L and P_U are assumed to be given exogeneously.

The constraint for the premium level in the index insurance design problem (1) can be understood in the following ways. First, the interval $[P_L, P_U]$ represents the price range that the farmer is willing to accept. If the farmer has a specific coverage level in mind that corresponds to a premium level of $\tilde{P}$, then she is only willing to pay an amount of premium for the index insurance product within a neighborhood of $\tilde{P}$. On the contrary, if the farmer is mainly interested in utility maximization, regardless of the premium she pays, P_L and P_U might be chosen as zero and her highest affordable price, respectively. Second, a constraint on the premium level is also economically meaningful. Since an index insurance is often subsidized by the government, restrictions on how much premium is paid per unit of risk exposure is imperative, in order to prevent abusive and speculative use of this insurance contract in practice. Finally, imposing such a restriction guarantees the well-posedness of this problem; otherwise, an optimal contract may not necessarily exist.

In this paper, we employ the most commonly used premium principle both in the literature and in practice to determine the premium of the index insurance contract, the expectation premium principle.[4] That is, the index insurance premium $\pi(I)$ is proportional to the risk premium as follows:

$$\pi(I) = \lambda \mathrm{E}\left[I(\boldsymbol{X})\right], \tag{2}$$

where λ is the risk loading parameter and $\lambda \geq 1$. When $\lambda = 1$, $\pi(I)$ is called the actuarially fair premium. λ affects the insurance premium and insurer's profits, and also

[4] In the actuarial science literature, a premium principle is a functional quantifying an insurable risk. See Young (2014) for a comprehensive introduction of premium principles.

impacts farmers' index insurance demand. In this paper, we endogenously determine the equilibrium λ from the supply and demand curves of insurance contracts, which will be discussed in more details in Section 4.2.1. Such endogenous insurance premium takes into account the strategic interaction between farmers and insurers while the model solves for the utility maximization problem of farmers.

We need to compute the expected utility to solve problem (1). Instead of modelling the joint distribution of the loss Y and indices $\boldsymbol{X}$, we replace the expected utility with its empirical counterparts and directly search for the optimal index insurance.[5] Specifically, for a random sample of $(\boldsymbol{X}, Y)$: $\{(\boldsymbol{x}_j, y_j)\}_{j=1,\ldots,n}$, where $\boldsymbol{x}_j = (x_{j1}, x_{j2}, ..., x_{jp})$, after replacing quantities with sample statistics, problem (1) can be reformulated as the following problem:

$$\begin{cases} \min\limits_{I\in\mathcal{I}} & -\frac{1}{n}\sum\limits_{j=1}^{n} U\big(w_0 - y_j + I(\boldsymbol{x}_j) - (1-\theta)\pi_e(I)]\big), \\ \text{s.t.} & P_L \leq \pi_e(I) \leq P_U, \end{cases} \tag{3}$$

where $\pi_e(I)$ is the empirical counterpart of $\pi(I)$ and is given by $\pi_e(I) := \frac{\lambda}{n}\sum\limits_{j=1}^{n} I(\boldsymbol{x}_j)$. Note that we have rephrased problem (3) from a maximization problem to a minimization problem by adding a negative sign to the objective function.

3 A neural network-based solution

It is challenging to search for the optimal functional form of the insurance payout (I) in the feasible set $\mathcal{I}$ for the optimization problem (3). As a result, the existing literature has considered a certain restrictive feasible set $\tilde{\mathcal{I}}_0 \subset \mathcal{I}$. Step functions (where index insurance payment is triggered by some pre-defined events) and piecewise linear functions (an excess-of-loss-type contract where the triggered payment is a linear function of the index) are commonly used functional forms in the literature and in practice (see, e.g., Mahul and

[5]Modelling the joint distribution of the loss Y and indices $\boldsymbol{X}$ might introduce misspecifications and approximation errors, especially when the dimension of $\boldsymbol{X}$ is high, which leads to unstable estimation results. For example, Cong et al. (2020) discuss the numerical disadvantage of a two-step approach in portfolio optimization.

Skees, 2007; Giné et al., 2007).[6] However, a restrictive feasible set results in contracts with large basis risk, which hinders customer participation and generates market failure of index insurance (Clarke, 2016). Hence, allowing for a larger set of feasible functional forms that can reduce basis risk is necessary. However, the set of feasible functional forms cannot be too large, either. Otherwise, we will run into "overfitting" issues (see Appendix A for an example). Therefore, we need to choose a feasible set that balances flexibility and stability.

In this paper, we consider a feasible set which allows for nonlinear and nonmonotonic relationships. Specifically, we apply neural networks to search for the optimal contract in the expanded feasible set $\mathcal{I}_0 \subset \mathcal{I}$ (see Appendix B for comparisons of different feasible sets). NN is designed to capture high dimensions, nonlinearity, and complex interactions among input factors, which is particularly attractive for our research question. NN is able to "increase both the selectivity and the invariance" through nonlinear transformations in each layer (LeCun et al., 2015). In other words, NN is sensitive to critical details and insensitive to idiosyncratic outliers in the data, which is potentially useful in achieving good flexibility-stability balance in solving for the optimal contract.

3.1 Neural networks and feasible set

We consider a standard multilayer fully-connected NN (see Appendix C for an illustration of the NN structure). $\boldsymbol{X}_{Input}$ in the input layer includes all indices used to construct the index insurance contract, while I in the output layer represents the final payoff function solved by this system. This is a "deep learning" architecture. In addition to an input layer and an output layer, there are $H (H > 1)$ additional hidden layers. The h-th $(h = 1, 2, \ldots, H)$ hidden layer $\boldsymbol{Z}^{(h)}$ contains p_h neurons, which are obtained by transforming the linear combination of the neurons from the previous layer through an activation

[6]For example, the payoff based on the index of X, $I(X)$, of a piecewise linear contract can be defined as follows:

$$I(X) = \begin{cases} I_1 & 0 \leq X \leq \tau_1 \\ \frac{-I_2(X-\tau_2)}{\tau_2-\tau_1} & \tau_1 < X < \tau_2 \\ 0 & X \geq \tau_2, \end{cases}$$

where τ_1 and τ_2 are two trigger levels, and I_1 and I_2 are constants. This contract can be easily extended by adding more triggers and more pieces of payments, but it is not straightforward to extend this payment function to a high-dimensional $\boldsymbol{X}$.

function, f_h, elementwise. $\boldsymbol{\alpha}^{(h)}$ and $\boldsymbol{\omega}^{(h)}, h = 1, 2, \ldots, H$, are parameters of the linear combination. $\boldsymbol{\alpha}^{(h)}$ is a $(p_h \times 1)$-vector of bias units that captures the intercepts in the model; $\boldsymbol{\omega}^{(h)}$ is a $(p_h \times p_{h-1})$-dimensional weight matrix. $\boldsymbol{\alpha}^{(h)}$ and $\boldsymbol{\omega}^{(h)}$ $(h = 1, 2, \ldots, H)$ are learned through stochastic gradient descent (SGD), where their gradients are derived by backpropagation.[7] This network is fully connected, in the sense that neurons between two adjacent layers are fully pairwise connected, but neurons within a layer have no connections. In summary, an NN structure is defined by its number of hidden layers H, the number of neurons in each hidden layer p_h, activation functions f_h, and the parameters $\boldsymbol{\alpha}^{(h)}, \boldsymbol{\omega}^{(h)}$ $(h = 1, 2, \ldots, H)$. A specific NN structure defines the feasible set $\mathcal{I}_0 \subset \mathcal{I}$ in the optimization problem (3).

Note that this framework conveniently guarantees that I is nonnegative, which is necessary for insurance indemnity payoffs, e.g., by making the last layer activation function, f_H, nonnegative. Throughout this paper, we use the rectified linear unit (RELU), defined by $f(x) = \max(x, 0)$, as the nonlinear activation function in our empirical analysis.

One might be concerned with the overfitting issue of neural networks. To avoid overfitting, we follow the literature and impose complexity constraints on function I. We use early stopping (Prechelt, 1998; Caruana et al., 2001). We also try other complexity constraints, including regularization methods (e.g., L^1 and L^2 regularization and their combination), elastic net (Tibshirani, 1996; Zou and Hastie, 2005), and the dropout method (Srivastava et al., 2014), and we find similar results.

3.2 Solving for the optimal policy using a penalty method

Two issues remain before solving for the optimal index insurance policy. First, the objective function in problem (3) is a utility function, not a typical loss function in machine learning. Second, there is a budget constraint in problem (3), that is, $P_L \leq \pi_e(I) \leq P_U$. The first issue can be solved by defining a customized loss function when we formulate a NN program.[8] To maneuver the second issue, we propose a penalty method.

[7] See Bottou and Bousquet (2008) for details about SGD. See Hastie et al. (2009) for details about backpropagation.

[8] In a standard NN-based statistical learning problem, the objective function is usually a loss function defined as certain distance measure (e.g., L^1 norm for mean absolute error, L^2 norm for mean squared error, the Huber loss, or the cosine similarity, all of which are commonly used built-in functions in Keras).

Let's consider a sequence of unconstrained problems $\{\Phi_k\}_{k\geq 0}$:

$$\Phi_k = \min_{I\in\mathcal{I}_0}\left(-\frac{1}{n}\sum_{j=1}^{n} U\big(w - y_j + I(\boldsymbol{x}_j) - (1-\theta)\pi_e(I)\big) + \phi_k \cdot g(I)\right), \tag{4}$$

where $\{\phi_k\}_{k\geq 0}$ is a sequence of increasing nonnegative real numbers tending toward infinity (i.e., $\phi_k \geq 0, \phi_{k+1} \geq \phi_k, \lim_{k\to\infty}\phi_k = +\infty$), and $g(I)$ is the penalty function. The penalty function is defined as a sum of squared errors (Luenberger et al., 1984), as follows:

$$g(I) = [\max(\pi_e(I) - P_U, 0)]^2 + [\max(P_L - \pi_e(I), 0)]^2. \tag{5}$$

Our index insurance design problem (3) then can be connected to the unconstrained problem (4), using the following theorem.

Theorem 1 (Luenberger et al. (1984))**.** *Let I_k^* be a sequence of solutions solving the corresponding sequence of unconstrained problems $\Phi_k(I)$, as defined in problem* (4)*, where $\{\phi_k\}_{k\geq 0}$ is an increasing sequence tending toward infinity. Then any limit of $\{I_k^*\}_{k\geq 0}$ is a minimizer of the constrained problem* (3)*.*

Therefore, based on Theorem 1, we construct a sequence of increasing penalty coefficients, $\{\phi_k\}_{k\geq 0}$, and specify the penalty function, $g(I)$, according to Equation (5). The optimal contract for problem (3) can be solved by the following numerical procedures, which are summarized in Algorithm 1.

4 Empirical performances of the NN-based index insurance

In this section, we evaluate the empirical performances of the proposed NN-based design of a weather index insurance policy. To proceed, we consider a representative corn farmer with a negative exponential utility function $U(w) = -\frac{1}{\alpha}e^{-\alpha w}$, where $\alpha > 0$ is the absolute risk aversion coefficient.[9] Using a CARA utility provides some advantages. First, under this CARA utility, the optimal index insurance contact in problem (3) is independent of farmers' initial wealth level. This property is realistic from the practical

For our problem, we create a customized loss function (4) using the backend functions in Keras.

[9]Other preferences are feasible in designing insurance contract. For example, Li et al. (2020) consider habit formation.

Output: An optimal index insurance policy I

Input : Index data $\boldsymbol{X}$ and loss data Y

Build and initialize a neural network;

Initialization: $k = 0, \phi_0 = \epsilon_1$, obtain I by solving an unconstrained problem Φ_0;

while $|I - I_{last}| > \epsilon_3$ **or** $g(I) > \epsilon_2$ **or** $|\pi_e(I) - \pi_e(I_{last})| > \epsilon_4$ **do**

Set $I_{last} = I$, $\pi_e(I_{last}) = \pi_e(I)$;

Update $k \leftarrow k + 1$;

Train the neural network and obtain the optimal I for problem $\Phi_k(I)$:

$$\Phi_k(I) = -\frac{1}{n} \sum_{j=1}^{n} U\big(w - y_j + I(\boldsymbol{x}_j) - (1-\theta)\pi_e(I)\big) + \phi_k \cdot g(I),$$

where the loss function is customized according to $\Phi_k(I)$ and I_{last} is set to the initial value of optimization;

Update $g(I)$ and $\pi_e(I)$;

end

return (I)

Algorithm 1: Solve for the optimal index insurance policy.

viewpoint of designing insurance policies. It also makes results comparable among different policyholders. Second, the negative exponential utility function conveniently handles negative wealth (i.e., bankruptcy) cases, whereas other utility functions, such as power utility functions, may not be well defined when wealth is negative, which could occur under some catastrophic events. Therefore, we use a negative exponential utility function in our main results. As a robustness check, we consider other utility specifications, e.g., logarithmic and power utility functions, in Section 6.3.

Section 4.1 summarizes the data used in this paper. In Section 4.2, we focus on the baseline case of a representative farmer when budget constraints are barely binding.[10] In other words, the farmer is primarily interested in maximizing her expected utility and accepts the optimal coverage level and its corresponding insurance premium. We study the reduction in basis risk and the utility improvement of the optimal policy in the baseline case. Section 4.3 explores the interpretability of the optimal policy. Finally, we compare the NN-based optimal policy with some simpler contracts in Section 4.4.

[10] We will test the impact of budget constraints in the sensitivity analysis in Section 5.2.

4.1 Data

4.1.1 Production loss and farmer's wealth data

We use a data set consisting of county-level annual corn production experience for Illinois, obtained from the National Agricultural Statistics Service (NASS). Corn is the single most valuable agricultural commodity in the United States, and composes more than 45% of cropland by acreage in Illinois.[11] Another advantage of using Illinois corn data is its long history. We use a sample period from 1925 to 2018. To ensure stationarity of the loss experience over the long sample period, we detrend the crop yield data with a second-order polynomial, estimated with a robust regression technique (Yohai, 1987). We adjust heteroscedasticity according to the Harri et al. (2011) method. The detrending procedure also helps remove the impacts of long-run climate change or technology progress, such as improved genetics (cultivars), improved crop management, and other technological advances like the use of advanced farming equipment. In the main results of this paper, we focus on hedging fluctuations in crop yields (i.e., production losses) and assume that the corn price is a constant of \$3.5 per bushel (\$/bu), which is the 5-year average of the inflation-adjusted corn price from 2014 to 2018.[12] We consider the price risk in Section 6. The corn production loss data measured in bushels per acre (bu/acre) are then multiplied by the corn price to arrive at the monetary loss incurred by each farmer in dollars per acre.

The 2014-2018 USDA Agricultural Resource Management Survey and NASS show that net farm assets (i.e., farm assets minus farm debts) have a five-year average of \$456,977.[13] In addition, according to Illinois Farm Business Farm Management, the average size of Illinois grain farms is 1,176 acres.[14] To make the numerical results comparable and

[11] Source: USDA's National Agricultural Statistics Service, Economic Research Service.

[12] We use the most recent five-year price average to reflect the current corn price level. Specifically, we obtain the commodity prices for corn at harvest from the USDA Economic Research Service, which are estimated using data from USDA's Agricultural Resource Management Survey and other sources (see `https://www.ers.usda.gov/data-products/commodity-costs-and-returns/`). Inflation is adjusted for by the annual consumer price index (CPI) from the Bureau of Labor Statistics.

[13] Available at `https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Ag_Resource_Management/`.

[14] See "Farm Income and Production Costs for 2017: Advance Report", available at `http://www.fbfm.org/pdfs/AdvanceReport17.pdf`.

interpretable, we normalize all monetary quantities for the representative farmer by farm size. In other words, wealth levels, incomes, and other monetary quantities used in our analysis are all measured in dollars per acre, which is also consistent with the unit of our production loss data. Dividing the net farm assets by the farm size yields the initial wealth of $w_0 = \$389$/acre.

4.1.2 Climate and weather index data

We collect climate and weather data from the PRISM Climate Group.[15] PRISM publishes monthly meteorological information on six climate variables, including precipitation, max/min temperatures, max/min vapor pressure deficit, and dew points, from 1895 to present for the conterminous United States at a 4-km resolution. We use all of the six monthly climate variables over 1925-2018. This gives a 72-dimensional weather index matrix for our optimal insurance policy design. Table 1 describes the weather variables from the PRISM data set. Appendix D provides summary statistics for the weather variables.

Agriculture data, such as crop yield data, are often quite limited. For example, in the United States, for each particular location, only annual observations are available for several decades. This means if we investigate each area individually, there are not enough historical yield data, especially given the high-dimensional weather-related covariates we want to utilize.[16] We circumvent this difficulty by assuming that crop yield losses are both time and space homogeneous.[17] This simplification expands the size of our data to 7,869 county-years (84 counties $\times$ 94 years, with 27 missing data points removed from the sample). We then shuffle the data set and divide it into three subsamples: 70% of the sample as the training set, 15% of the sample as the validation set, and the remaining 15% as the test set, following the convention of machine learning.[18] NN parameters are

[15] PRISM is the USDA's official climatological data, available at `http://prism.oregonstate.edu/`.

[16] For example, in our empirical analysis, although we have 94 years of loss data (a longer time span than most agricultural insurance analysis) for each location, the data are still relatively scarce compared to the 72-dimensional weather index matrix.

[17] Doing so implies that we ignore spatial basis risk, and design a universal weather index insurance contract for all farmers in Illinois.

[18] In practice, insurance contract is renewed annually. Therefore, it is reasonable to treat insurance design as a static problem. That is, we treat our data as a random sample and train the index insurance contract. We also investigate the index insurance performance when preserving the temporal order of

trained on training set for each given set of hyperparameters, including the number of neurons and hidden layers. Then the validation set is used to choose the optimal set of hyperparameters. Finally, we use the test set to evaluate the performance of the selected optimal NN-based index insurance.

A straightforward measure of the relationship between losses and weather indices is correlation. Appendix F shows that during the growing season (normally from May to October for Illinois corn), especially in June, July, and August, indices tend to have higher correlations than those outside the growing season. Correlations only describe linear relationships. However, the impacts of weather indices on crop production losses can be nonmonotonic and highly nonlinear due to biological reasons.[19] Figure 2 illustrates the nonlinear patterns of four selected weather indices (see Appendix G for scatterplots of all 72 weather indices). Furthermore, different weather indices might interact with each other and jointly affect production losses. Panels (a) and (b) of Figure 3 show that the minimum vapor pressure deficit in December and precipitation in April do not individually affect production losses. However, Panel (c) of Figure 3 shows that they jointly affect production losses, in a highly nonlinear way. Rigden et al. (2020) discuss the importance of combining heat and moisture conditions (two factors that interact with one another) to better predict corn yields. These complexities cannot be captured by the linear models used by most current index insurance contracts, and, thus, more sophisticated contract design models are needed.

4.1.3 Insurance market data and the supply curve

We obtain insurance market data from the USDA National Summary of Business (SOB) Reports (1989-2017).[20] The variables that we collect include the total acres insured ($Acres$), premiums ($Prem$), liabilities ($Liab$), and indemnities ($Indem$). Liabilities stand for insurance guaranteed crop production levels. If the farmer's actual yields are lower than the liability level, the farmer will receive insurance payments. Indemnities correspond

the data. In particular, we split the sample into three disjoint time periods, with the earliest 70% data as training set, then the next 15% as validation set, and the last 15% as test set. We find similar results (see details in Appendix E).

[19]For example, Schlenker and Roberts (2009), Lobell et al. (2013), Lobell et al. (2014), and Rigden et al. (2020) show a nonlinear relationship between soil moisture, temperature, and maize yields.

[20]Available at: `https://www.rma.usda.gov/en/Information-Tools/Summary-of-Business`.

to realized insurance payments. The USDA Risk Management Agency computes the loss cost ratio (LCR) as the ratio of indemnity over liability: $LCR = \frac{Indem}{Liab}$. The premium is then calculated as the expected LCR times the liability, multiplied by the loading parameter λ, that is, $Prem = \lambda \cdot \mathrm{E}(LCR) \cdot Liab$. Therefore, we can use this relationship to infer the realized λ every year from the market data. Insurance coverage per acre is calculated as $Cov = \frac{Indem}{Acres}$ (adjusted by inflation), representing the unit acre insurance supply at market equilibrium each year. We use the historical data to fit the supply curve as a power function of λ, using nonlinear least squares method. The fitted supply curve is $Cov = f_S(\lambda) = 7.04\lambda^{2.92} + 12.83$. Due to the data limitation, one might worry about the simultaneity bias of fitting the supply curve from historical data. We explore the robustness of the results by using alternative supply curves in Section 5.4.

4.2 The optimal NN-based index insurance: The baseline case

We first consider the optimal index insurance whereby farmers are barely bound by any budget constraint. For example, P_L is set to zero and P_U is set to the farmer's total wealth. In other words, the farmer would like to pay any price that she could afford, to maximize her utility. We will discuss the effects of budget constraints in Section 5.2. We set the absolute risk aversion, α, to 0.008, which corresponds to a relative risk aversion of 3.1.[21] The subsidization proportion, θ, is set to zero in the baseline case. The impacts of θ on the optimal index insurance policy will be investigated in Section 6.1. In Section 4.2.1, we discuss how to endogenously determine the loading parameter λ at equilibrium. Then we discuss the performance of the optimal index insurance of the baseline case in Section 4.2.3.

4.2.1 Determining the loading parameter at market equilibrium

We determine the endogenous loading parameter at market equilibrium via a reduced-form approach, as follows. First, we use the USDA SOB Reports market data to fit the supply curve, $Cov = f_S(\lambda)$, as described in Section 4.1.3. Second, for a given value of

[21] This is in line with the literature. For example, Bontems and Thomas (2000) estimate an average of 3.7 relative risk aversion for U.S. crop farmers, and Chavas and Holt (1990) find that maize and soybean farmers have a relative risk aversion ranging from 1.41 to 7.62, using aggregated data.

$\lambda \in [1.02, 1.5]$,[22] we obtain the corresponding optimal coverage of index insurance policies, using the NN-based approach. Given the pairs of λ and the optimal coverage, we fit the demand curve, $Cov = f_D(\lambda)$, using the piecewise cubic hermite interpolating polynomial (PCHIP).[23] Finally, we equate the supply and demand functions, $f_S(\lambda) = f_D(\lambda)$, to solve for the loading parameter at market equilibrium, λ^*. Therefore, farmers and insurers jointly determine the insurance premium in the equilibrium. Figure 4 illustrates an example of finding λ^* at the intersection of the supply and demand curves. The demand curve is for an NN-based optimal index insurance with a 3-hidden-layer (64-64-16 neurons) structure, corresponding to the baseline model that we will discuss in Section 4.2.3. The equilibrium loading parameter is $\lambda^* = 1.1668$.

4.2.2 Selecting the optimal NN-based model

We consider and compare different neural network structures. For each neural network structure, we use the procedures described in Section 4.2.1 to compute the equilibrium loading parameter, λ^*. As we endogenously determine λ^*, it would be very computationally costly to endogenize the hyperparameters of the model. Instead, we examine a wide range of candidate models with different number of hidden layers and different number of neurons in each layer, and then select the optimal model based on the validation set. Panels B-E of Table 2 present the results from various NN-based index insurance contracts, while Panel A shows the results without using index insurance. We compute the insurance premium and compare expected utilities with and without index insurance. For ease of interpretation, we also compute the CEW with and without index insurance.

Panel B shows that a simple NN with only one hidden layer and two neurons significantly improves farmers' expected utility by 12.11% and CEW by \$16.14/acre in the validation sample. In addition, adding more neurons or adding more hidden layers generally improves the model performances, as shown in Panels B-D. However, increasing the number of hidden layers from 3 to 4 (moving from Panel D to Panel E) does not

[22] We start with $\lambda = 1.02$ and end with $\lambda = 1.5$, using an increment of 0.02. A loading parameter above 1.5 is generally infeasible for index insurance in practice.

[23] The advantage of PCHIP in curve fitting is its shape-preserving property, that is, the fitted curve preserves the monotonicity of the data (Fritsch and Carlson, 1980). This property is desirable given that we are fitting a decreasing demand curve.

necessarily improve the NN performance. We see that the three-hidden-layer NN with a 64-64-16 neurons structure yields the largest utility improvement and CEW improvement in the validation set. This NN-based contract provides a utility improvement of 19.30% in the training sample and 13.03% in the validation sample. Given these optimized hyperparameters, we combine the training set and the validation set as a larger training set and train the model again (see, e.g., Chollet and Allaire, 2018). We call this as the baseline model and evaluate its performances later.

4.2.3 Performances of the optimal index insurance

We now discuss the performance of the optimal index insurance in the test sample, reported in Table 3. The baseline contract (*NN72*, with a 64-64-16 neurons structure) provides a utility improvement of 12.33% in the test sample. With this insurance contract, the CEW improves by \$16.45/acre in the test set. Given an average farm size of 1,176 acres, the wealth improvement effect of the NN-based index insurance for a typical farm is economically significant, for example, \$19,345.2 in the test sample.

The optimal baseline index insurance policy has a premium of \$28.82 in the test sample. The corresponding equilibrium loading parameter is $\lambda^* = 1.1668$, implying a coverage level of \$24.70 in the test sample. Compared to the average insurance premium for Illinois corn farmers, which is around \$45.50/acre in 2017 (Smith, 2017), this optimal premium level is only 63% of the current premium level in practice. This lower premium for the proposed optimal index insurance would substantially increase the demand of index insurance in practice.

It is also shown that the proposed insurance achieves very similar expected utilities over training and test samples, suggesting that "overfitting" is of little concern. For example, the difference between policyholders' expected utilities in the training and test sample is 0.09, and the difference between CEW in the training and test sample is \$3.46, both of which are not sizable.

To further illustrate how the proposed NN-based insurance contract helps farmers improve their welfare, Figure 5 compares the wealth distributions between the "with insurance" case and the "without insurance" case. Without insurance, farmers' wealth distribution has a larger variation with a heavy left tail, i.e., the downside risk. In contrast, with the proposed index insurance, farmers' wealth distribution becomes less dispersed, and the left tail is significantly reduced.[24] We observe similar patterns in both training and test samples. Figure 5 demonstrates that our proposed index insurance policy effectively hedges the downside weather risk for farmers and substantially improves their utilities.

4.3 Interpreting the NN-based index insurance: Sensitivity analyses

Neural network results are often difficult to explain while economic interpretability is crucial for economics (Cong et al., 2019, 2020; Horel and Giesecke, 2020).[25] In this subsection, we perform a battery of sensitivity analyses to interpret the baseline case of NN-based index insurance contract.

Following Cong et al. (2020), we adopt a polynomial feature sensitivity analysis, which involves using the gradient-based characteristic importance method to examine contributions of feature inputs and extracts important features.[26] Hence, it provides "economic distillation" of the NN-based index insurance, and increases interpretation of the "black box" insurance policy. Given the optimal NN-based index insurance, $I(\boldsymbol{X})$, where $\boldsymbol{X}$ is the vector of weather indices, the sensitivity of $I(\boldsymbol{X})$ to the weather index X_k can be measured by its partial derivative with respect to X_k:

$$\delta_{X_k}(\boldsymbol{X}) = \frac{\partial I(\boldsymbol{X})}{\partial X_k} = \lim_{\Delta x_k \to 0} \frac{I(\boldsymbol{X}) - I(X_k + \Delta x_k, \boldsymbol{X}_{-k})}{\Delta x_k}, \tag{6}$$

where $\boldsymbol{X}_{-k}$ is the vector of weather indices with the kth index removed. Then the gradient-based sensitivity of insurance payoff with respect to X_k, $\mathcal{S}^g_{X_k}$, can be expressed as the

[24]The mean wealth is slightly lower with insurance, because of the insurance premium paid.

[25]Cong et al. (2019) provide economic interpretability from textual analyses. Cong et al. (2020) use polynomials sensitivity and textual analysis for "economic distillation". Horel and Giesecke (2020) propose a procedure that establishes the test statistics to rank input features. See Section 1.1 of Horel and Giesecke (2020) for a review of literature on feature importance analysis.

[26]When applying the gradient-based methods to neural networks, it is important to check whether the model suffers from gradient explosion/vanishing problems, especially in sequential data analysis. See Cong et al. (2020) for more discussions on this issue. In this paper, we are using a static model, and we do not find it to suffer from gradient explosion or gradient vanishing problems, neither in training set nor test set.

average influence of index X_k to the optimal index insurance payoff, calculated as:

$$\mathcal{S}^g_{X_k} = \mathrm{E}\left(\delta_{X_k}\right) = \int_\Omega \delta_{X_k}(\boldsymbol{X}) d\mathbb{P}(\boldsymbol{X}), \tag{7}$$

where $\mathbb{P}(\boldsymbol{X})$ is the joint probability distribution of $\boldsymbol{X}$, Ω is the sample space, and the superscript g indicates the gradient-based sensitivity analysis. Empirically, $\mathcal{S}^g_{X_k}$ can be estimated by

$$\bar{\mathcal{S}}^g_{X_k} = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_k}(\boldsymbol{x}_i), \tag{8}$$

where $\boldsymbol{x}_i, i = 1, \ldots, N$, are the sample data of $\boldsymbol{X}$ in the training set.[27]

We rank the importance of weather indices by the absolute value of $\bar{\mathcal{S}}^g_{X_k}$. Table 4 lists the 10 most important weather indices, as shown in the right panel. For comparison, we also rank weather indices based on the absolute correlations between insurance payoffs and the indices, and the top-ten indices are shown in the left panel.[28] We see that 7 of 10 indices are different in these two panels, and most of them have very different ranks under these two ranking methods. While the correlations capture linear dependence, sensitivities from NN-based insurance provide a more general nonlinear dependence measure to study the relationship between insurance payoffs and weather indices. From the perspective of designing effective index insurance contracts, those weather indices with a large absolute value of correlations are not necessarily the most important ones, whereas those variables with small correlations (dpt1, pcpn6, tmin2, tmax11, tmax10, and vpdmax1) may be very useful for the insurance. Finally, we notice that the 10 most important weather indices based on absolute correlations are all within the corn growing season, but the NN-based insurance contract picks up four weather indices outside the growing season and that are overlooked by existing contracts, i.e., dew point temperature in January (dpt1), minimum temperature in February (tmin2), maximum temperature in November (tmax11), and maximum vapor pressure deficit in January (vpdmax1). In fact, agronomic studies show that hydrological conditions over nongrowing seasons might be important

[27] Averaging over training data allows one to interpret the optimal index insurance in the training set, whereas averaging over the test set allows one to describe the policy behaviour on the test set. Cong et al. (2020) perform both distillation exercises. In our analysis, we find the behaviour of $\mathcal{S}^g_{X_k}$ on training set and test set are very similar.

[28] Appendix H.1 ranks all indices by the gradient-based sensitivity analysis.

for corn production for two reasons. First, water accumulation in winter will affect the soil moisture that becomes available for corn in the next growing season.[29] Second, soil water accumulation during the nongrowing season affects cover crops, which are grown to improve soil quality and influence corn growth later.[30] This further illustrates the ability of neural networks to extract important nongrowing season information from the data, which is often overlooked by the linear functions in current practice.

4.4 A comparison with other simple insurance contracts

Previously, we demonstrated the effectiveness of the NN-based index insurance design and assessed its interpretability. In this subsection, we further compare the NN-based insurance contract with other commonly used simple contracts, i.e., contracts with fewer indices and/or less nonlinear structures. This comparison highlights the benefits of using the NN-based insurance contract. In addition to the NN-based insurance contract, we consider some polynomial-based contracts, including (1) a linear insurance contract; (2) a quadratic insurance contract; and (3) a cubic insurance contract.[31] Depending on the contract structure, we use either all 72 weather indices if manageable or the most important one or five weather indices identified by gradient-based sensitivity analysis in the NN-based contract (including, precipitation in July (pcpn7), maximum VPD in July (vpdmax7), dew point temperature in January (dpt1), minimum temperature in February (tmin2), and maximum temperature in November (tmax11)). Specifically, we consider the following seven index insurance contracts:

- *Linear contract with one input (Linear1)*: a linear contract written on a single rainfall index, which is a currently used index insurance contract in practice;

[29]Suyker and Verma (2008) find that cumulative evapotranspiration during the "nongrowing" seasons contributes to 20%–25% of the annual evapotranspiration totals for corn. Li et al. (2019) show that water storage over nongrowing seasons affects corn yields.

[30]Cover crops, such as cereal rye, which help build and improve soil fertility and quality, control diseases and pests, and promote biodiversity, are commonly integrated into corn production in Illinois. In fact, in Illinois, farmers who adopt cover crops may be eligible to receive an insurance premium discount in the following year through the Illinois Department of Agriculture (IDOA) Cover Crop Premium Discount Program (Illinois Department of Agriculture, 2020).

[31]Polynomial terms might not be orthogonal, so using polynomials of a higher degree would introduce multicollinearity/robustness issues.

- *Linear contract with five inputs (Linear5)*: only linear polynomials are used, with the top-five weather indices;

- *Quadratic contract with five inputs (Quadratic5)*: quadratic polynomials are used, with the top-five weather indices;

- *Cubic contract with five inputs (Cubic5)*: cubic polynomials are used, with the top-five weather indices;

- *NN with five inputs (NN5)*: NN with a 64-64-16 neuron structure, using the top-five weather indices;

- *Linear contract with 72 inputs (Linear72)*: only linear polynomials are used, with all 72 weather indices;

- *Baseline model (NN72)*: NN with a 64-64-16 neuron structure, using all 72 weather indices. This is the baseline optimal contract discussed in Sections 4.2 and 4.3.

For each contract, the loading parameter is endogenously determined via the reduced-form approach described in Section 4.2.1. Table 3 presents the performances of these seven contracts in the test set. We compare the contract performances mainly by examining their utility improvements, CEW improvements, and risk reductions. We measure risk reductions by the standard deviation, or the 5% and 1% value-at-risk (VaR) of policyholders' wealth.

First, adding more weather indices as inputs significantly improves the contract performance of *Linear72*, relative to *Linear1* and *Linear5*. Second, comparing *Linear5*, *Quadratic5*, and *Cubic5*, we see that utility improves a little when adding nonlinear terms to the insurance contract in Panel A. We see similar improvements in CEW in Panel B, standard deviation of wealth in Panel D, and tail risk in Panel E. Third, using the same five weather indices, we see *NN5* outperforms *Linear5*, *Quadratic5*, and *Cubic5*. Fourth, the baseline model (*NN72*) delivers the best performance. *NN72* provides the highest expected utility and CEW to farmers, and it also achieves the highest risk reduction measured by standard deviation reduction and tail risk reduction. Furthermore, from the insurers' perspective, *NN72* is also the best contract because it provides the largest profits. Quantitatively, *Linear1* improves farmers' utility by 3.11%; *Linear71* improve their utility by 9.75%; *NN72* improves their utility by 12.33%. In terms of CEW, *Linear1*, *Linear72*, and *NN72* improve it by \$3.96, \$12.82, and \$16.45, respectively. Insurers profits are \$1.31/acre, \$2.70/acre, and \$4.12/acre, for the *Linear1*, *Linear72*, and *NN72*, respectively. This is due to higher market demand and hence a higher equilibrium loading, λ^*. Finally, inspecting the insurance profits from *Linear5*, *Cubic5*, and *NN5*, it is also interesting to note that *Linear5* has slightly higher profits than *Cubic5* and *NN5*. This indicates that when the index insurance policy has high basis risk and the demand is low, more complicated contracts will hurt insurers' profits. This result partially explains why in the current index insurance practice, where the underlying dimension of indices is low, insurers only provide index insurance contracts with simple structures. Our results demonstrate that, by solving the curse of dimensionality and reducing basis risk, the NN-based solution can significantly improve the demand of index insurance and increase the profit margin of insurers.

Basis risk is a critical issue that dissuades farmers from buying index insurance, as policyholders are most concerned with a scenario in which they actually suffer a huge loss in crop yields, but only receive a small or even zero payments from the index insurance. The opposite case is also unappealing: although it does not hurt when farmers get a good harvest and receive some insurance payments at the same time, insurance companies will suffer losses in this situation. Moreover, it means that the insurance premiums are not effectively allocated to the worst scenarios when insurance payoffs are more needed, and, thus, the insurance contract is not well designed to stabilize farmers' financial positions. Panel (a) of Figure 1 replicates the large basis risk observed in current practice, which is a single-index piecewise-linear insurance contract (*Linear1*). Figures 6 and 7 illustrate how well insurance payoffs match the real losses incurred for the other six index insurance contracts discussed above, using the training sample and test sample, respectively. Across all contracts, except *NN72*, we observe a notably large mismatch between losses and insurance payoffs, especially for the test set. In contrast, *NN72* has a payoff function that is similar to the stop-loss payoff function of a conventional indemnity-based insurance, indicating its dramatic accuracy in mimicking the actual losses by utilizing complex information conveyed in the weather variables. Therefore, the baseline model achieves low basis risk, which is similar to a conventional indemnity-based insurance, and concurrently enjoys other advantages of an index-based insurance, i.e., zero moral hazard and low administrative costs. These results illustrate the importance of using nonlinear, high-dimensional inputs when designing the index insurance contracts.[32]

5 Robustness of NN-based index insurance

In this section, we investigate the robustness of the NN-based index insurance contract in several ways, using a NN structure of 64-64-16 neurons. First, we use data from three states adjacent to Illinois to perform the out-of-state tests. Second, we examine the impacts of farmers' characteristics, such as different coverage levels and risk aversion, on the NN-based index insurance contract. Third, we examine the impacts of insurers' supply curves on the insurance contract.

5.1 Out-of-state tests

As a more stringent out-of-sample test, we run some out-of-state tests for states adjacent to Illinois. Specifically, we use Illinois data as the training set to estimate the parameters in the NN-based index insurance policy and use data from the adjacent states as the test set to evaluate the model performance. We select three states adjacent to Illinois and that have similar latitudes: Indiana, Kentucky, and Missouri. For example, when we evaluate the index insurance performances in Indiana, the insurance contract is designed with Illinois data and does not employ any information from Indiana. Table 5 summarizes the results. We see that the NN-based index insurance improves the expected utility by 8.63% in Indiana, 6.82% in Kentucky, and 8.69% in Missouri. This is about half of the improvement in the training sample of Illinois. As for CEW, improvements are $11.28/acre, $8.84/acre, and $11.36/acre in Indiana, Kentucky, and Missouri, respectively, which are also approximately half of the improvement in Illinois. These results demonstrate the power of our proposed NN-based index insurance.

[32]Instead of using 72 weather indices, one might selectively use few weather indices and still achieve reasonable insurance contracts. Appendix H.2 compares models with different number of weather indices.

Of course, in practice, insurance companies would not ignore any available data from the region of interest and solely rely on data from other regions to design an insurance product. Nevertheless, these "out-of-state" test results provide useful practice guideline for index insurance design when data is limited. For example, insurers can employ a transfer learning procedure, in which they first train the insurance policy in contiguous region with rich data and then apply the model to another region with limited data. This is particularly useful in the case when data are completely unavailable, for example, in developing regions where data are sparse or less reliable or when insurance companies plan to explore a new business line for a certain crop type, but historical data are not available yet.

5.2 Farmers with different coverage demands

In the baseline case, we focus on farmers who are solely interested in maximizing their utility, regardless of the coverage they purchase and premiums they pay. In practice, however, farmers often have a predetermined level of coverage in mind, because of either a better understanding of their financial position and insurance demand, or a relatively tight budget constraint. As a result, these farmers may be interested in more customized index insurance contracts. The NN-based index insurance design proposed in this paper is convenient to create customized contracts to meet their demands.

For illustration purposes, we consider a set of index insurance plans with a coverage of \$10, \$20, \$30, and \$40. Table 6 summarizes the results of these four contracts. For comparison purposes, we also list the baseline model which has the optimal coverage of \$27.16. We see that the amount of utility improvement first increases with coverage, peaking in the baseline case, and then decreasing with the coverage level. Overall, the NN-based insurance contract provides reasonable utility improvement for various coverage levels.

5.3 Farmers with different levels of risk aversion

Farmers' risk aversion varies with their age, education, farming experience, wealth, etc. One might wonder how different risk appetites lead to different demands for insurance. In this subsection, we consider policyholders with various levels of risk aversion. In addition to the baseline case in which $\alpha = 0.008$ (corresponding to a relative risk aversion of 3.1), we consider farmers with relative risk aversions of $2, 4$, and 5, which correspond to absolute risk aversion coefficients of $\alpha = 0.0051, 0.0103$, and 0.0129, respectively. Table 7 summarizes the results. Table 7 shows that the optimal index insurance design achieves greater utility and CEW improvements for farmers with higher risk aversion. This is because more risk-averse policyholders are more concerned about volatilities in their wealth. As such, insurers can charge these farmers higher premiums (i.e., imposing a higher loading parameter, λ^*).

Next, we consider farmers with time-varying risk aversion which depends on losses in the previous year. For example, farmers might become more risk averse after large losses, especially for less educated farmers without long-term learning skills (Cai et al., 2020). Such time-varying risk aversion might capture time inconsistency as well. Specifically, we first compute the 75^{th} and 25^{th} percentiles of yield loss. Suppose the farmer's average relative risk aversion is $RRA = 3.1$ (absolute risk aversion is $\alpha = 0.008$). If the farmer experiences a loss larger than the 75^{th} percentile in year $t-1$, her risk aversion in year t becomes $3.1 \times (1 + x)$; on the contrary, if the farmer experiences a loss lower than the 25^{th} percentile in year $t-1$, her risk aversion in year t is $3.1 \times (1 - x)$. That is, the farmer's risk aversion is $3.1 \times (1 - x), 3.1$, or $3.1 \times (1 + x)$, depending on the previous loss experience. We consider different levels of risk aversion variations, i.e., $x = 0.1, 0.2$, and 0.3. The results are summarized in Table 8. Generally, we see that the NN-based index insurance consistently improves farmers' utility and CEW, and reduces basis risk across the specifications. Nevertheless, time-varying risk aversion impedes the performance of the designed index insurance, with larger variations of risk aversion hindering the insurance performance more.

5.4 Different insurers' supply curves

In our previous analysis, the equilibrium loading parameter, λ^*, was determined via a reduced-form approach, where the supply curve is estimated using market data from the USDA SOB Reports, which might have estimation errors. In this subsection, we further investigate the impacts of insurers' supply curve. We use the upper and lower bounds of [10%, 90%] and [25%, 75%] confidence intervals of the supply curve estimates to determine the equilibrium loading parameter. Figure 8 displays our estimated supply curve with its confidence intervals.

Table 9 summarizes the results. When insurers' supply curve shifts to the upper bounds of its confidence intervals, equilibrium loading parameter, λ^*, decreases, and the equilibrium insurance demand increases. As a consequence, farmers buy more index insurance, and achieve more utility improvements and higher CEW. While insurance demand increases, we observe that insurers do not gain higher profits as the insurance is priced lower. When the insurance supply curve moves toward the lower bounds of its confidence intervals, λ^* increases and insurance demand decreases. Correspondingly, farmers' utilities and CEW improvements are reduced. Overall, under various supply curves, the NN-based index insurance contract provides robust results for utility and CEW improvements, risk reduction, and insurers' profits.

6 Extensions

In this section, we explore several extensions to the baseline model. In Section 6.1, we consider the impacts of government subsidies. In Section 6.2, we extend the proposed framework to protect both production and corn price risks. Section 6.3 evaluates the NN-based index insurance contract using alternative utility functions, i.e., log utility and power utility.

6.1 Impacts of government subsidies

Previously, we discussed the optimal index insurance design from the farmers' and the insurers' viewpoints, evaluating the farmers' utility improvement and the insurers' profits. Agricultural insurance often involves a public-private partnership (PPP). Government transfer payments are indispensable for the current crop insurance market. The current agricultural insurance market heavily relies on government subsidies. For example, in the United States, the Federal Crop Insurance Program (FCIP) cost the federal government about \$6.3 billion in 2017, with a subsidy ratio of about 60% (Smith, 2017). In this subsection, we study the impacts of government subsidies.

We consider various subsidy rates, θ. The results are summarized in Table 10, where θ is selected from $\{0, 5\%, 7.5\%,\ 10\%, 14.29\%\}$. Note that $\theta = 14.29\%$ is the subsidy level such that farmers pay the actuarially fair premium and the government pays related costs and loadings, which is much lower than the rate of 60% in crop insurance practice (Rosa, 2018b). With subsidies, the supply curve stays the same, while the demand curve of policyholders moves upward, resulting in a higher loading parameter at equilibrium. Producer's utility, CEW, and coverages increase slightly. The insurer's profits increase significantly. This is because the insurer has a small price elasticity, while farmers are more sensitive to prices. Therefore, subsidies cause larger changes to insurance demand curve. A higher profit margin will encourage more insurance companies to participate in the weather index insurance market. Therefore, the NN-based insurance proposed in this paper not only reduces basis risk, but also efficiently utilizes the government budget, increases insurance participation and improves social welfare.

6.2 Protecting corn price risk

Previously we focus on discussing index insurance for production losses, as the production loss is heavily affected by weather risk and natural disasters, and yield-based policies are primary policies with the longest history in the U.S. FCIP. However, as corn prices fluctuate, one might consider simultaneously providing corn price protection to farmers. In this subsection, we apply the same NN-based framework and design an index insurance contract to protect both the production and price risks, that is, the revenue protection.

We use the average price of the Chicago Mercantile Exchange (CME) Group December futures contracts during the month of February as the projected corn price.[33] The futures price contains the market's expectation of the corn commodity demand and supply within the same calendar year. In addition, the average of December CME Group futures contract price during February is also used as the projected price of the revenue protection in the FCIP in the U.S. Therefore, it is an appropriate measure of price risk. The sample period for the futures prices is from 1980 to 2017. We compare two contracts: *NN72* (the baseline

[33] See Rouwenhorst and Tang (2012) and Kang et al. (2020) for discussions of commodity pricing.

model), and *Linear72* (a linear contract with all 72 weather indices).

Table 11 summarizes the index insurance performances. We see that after considering price risk, the *NN72* index insurance remains effective in improving farmers' utilities and CEW, stabilizing their wealth distributions, and reducing downside tail risks. The *NN72* contract achieves a CEW improvement of \$240.67/acre in the training sample and \$206.75/acre in the test sample, improving CEW by 67.43% and 53.02% in the training sample and test sample, respectively. This is a larger improvement compared to the case that considers production risk only. This is because producer's CEW without insurance is lower, whereas her CEW with insurance is higher, after considering price risk. For example, in the test set, the producer's CEW without insurance becomes \$390/acre, which is \$42/acre lower than the case without price risk (see the last column of Table 3). After using the *NN72* index insurance, the farmer's CEW becomes \$597/acre, which is \$149/acre higher than the case without price risk. Farmers face more significant tail risk after considering the price risk; therefore, the *NN72* index insurance achieves more significant risk reductions, as measured by both standard deviation and value-at-risk. Correspondingly, it also has a higher premium compared to the case without price risk. Comparing *NN72* and *Linear72* contracts, we see that *Linear72* has about 10% worse performance, even with a slightly higher premium. Figure 9 presents the basis risk of the two index insurance contracts. *NN72* inarguably achieves a more effective basis risk reduction than does *Linear72*, as the insurance payoffs better match the real incurred losses.

6.3 Alternative utility functions

In this subsection, we evaluate the performance of the proposed NN-based index insurance using constant relative risk aversion (CRRA) utility functions. We consider the power utility with various levels of risk aversion, that is, relative risk aversion (RRA) of 2, 3, 4, and 5, and log utility (RRA = 1). Again, we use the 3-hidden-layer (64-64-16 neurons) structure, as in the baseline model. Table 12 summarizes the results.[34] The performance is similar to the baseline case with negative exponential utility. Using log

[34] Logarithm utility and power utility functions are not defined for negative wealth. Therefore, to avoid the negative wealth cases, we winsorize the loss data at 99% percentile.

utility, we find that the farmer's utility and CEW improvements are marginal because the risk aversion of her log utility is low (RRA = 1). Evaluating the results of power utility, we see that as policyholders become more risk averse, they purchase more coverage, and insurers also make higher profits. Index insurance performance also significantly increases with risk aversion. For example, the CEW improvement for RRA = 5 is about 6 times larger than that for RRA = 2. Comparing the results with the negative exponential utility case in Table 7, we observe that the insurance has higher utility and CEW improvements, given the same RRA. This is because power utility functions penalize extremely low wealth cases more severely.

7 Conclusion

Index insurance could effectively manage systemic weather risk. However, basis risk, the discrepancy between the actual losses and the insurance payoffs dictated, hinders market participation. In this paper, we formulate a neural network-based design of an index insurance contract to capture the high-dimensional, nonlinear, and interactive nature of weather indices on production losses. We solve the problem by a penalty method, together with backpropagation.

We illustrate its superior performance by applying it to corn farmers in Illinois. Results show that the NN-based optimal index insurance contract greatly outperforms other contracts (e.g., piecewise linear contract, quadratic contract, or cubic contract). In fact, the payoff function of our proposed NN-based index insurance is very close to a conventional indemnity-based insurance with deductible (i.e., a stop-loss insurance), indicating that the NN-based optimal index insurance effectively reduces basis risk. This NN-based insurance contract significantly improves farmers' utility, increases market participation, and mitigates downside risk. We verify this using out-of-state tests and examine the impacts of different coverage demands, risk aversion levels, utility specifications, insurer's supply curves, and government subsidies. Our results suggest that such index insurance contract can improve market demand and is viable with little or even zero subsidies from the government and, hence, improves social welfare.

References

Assa, H., Wang, M., 2020. Price index insurances in the agriculture markets. North American Actuarial Journal forthcoming.

Baldauf, M., Garlappi, L., Yannelis, C., 2019. Does climate change affect real estate prices? Only if you believe in it. Review of Financial Studies 33, 1256–1295.

Bernstein, A., Gustafson, M. T., Lewis, R., 2019. Disaster on the horizon: The price effect of sea level rise. Journal of Financial Economics 134, 253–272.

Bianchi, D., Büchner, M., Tamoni, A., 2020. Bond risk premia with machine learning. Review of Financial Studies forthcoming.

Bontems, P., Thomas, A., 2000. Information value and risk premium in agricultural production: The case of split nitrogen application for corn. American Journal of Agricultural Economics 82, 59–70.

Bottou, L., Bousquet, O., 2008. The tradeoffs of large scale learning. In: *Advances in Neural Information Processing Systems, eds. M. Jordan, Y. LeCun and S. Solla*, pp. 161–168.

Bryan, G., 2019. Ambiguity aversion decreases the impact of partial insurance: Evidence from african farmers. Journal of the European Economic Association 17, 1428–1469.

Bryzgalova, S., Pelger, M., Zhu, J., 2020. Forest through the trees: Building cross-sections of stock returns, Working Paper, Stanford University.

Cai, J., De Janvry, A., Sadoulet, E., 2020. Subsidy policies and insurance demand. American Economic Review 110, 2422–53.

Caruana, R., Lawrence, S., Giles, C. L., 2001. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In: *Advances in Neural Information Processing Systems*, pp. 402–408.

Casaburi, L., Willis, J., 2018. Time versus state in insurance: Experimental evidence from contract farming in Kenya. American Economic Review 108, 3778–3813.

Chavas, J.-P., Holt, M. T., 1990. Acreage decisions under risk: The case of corn and soybeans. American Journal of Agricultural Economics 72, 529–538.

Chen, L., Pelger, M., Zhu, J., 2020. Deep learning in asset pricing, Working Paper, Stanford University.

Chinco, A., Clark-Joseph, A. D., Ye, M., 2019. Sparse signals in the cross-section of returns. Journal of Finance 74, 449–492.

Choi, D., Gao, Z., Jiang, W., 2020a. Attention to global warming. Review of Financial Studies 33, 1112–1145.

Choi, D., Jiang, W., Zhang, C., 2020b. Alpha Go everywhere: Machine learning and international stock returns, Working Paper, Chinese University of Hong Kong.

Chollet, F., Allaire, J. J., 2018. Deep Learning with R. Manning Publications, New York, United States.

Clarke, D. J., 2016. A theory of rational demand for index insurance. American Economic Journal: Microeconomics 8, 283–306.

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, 104–35.

Cole, S., Giné, X., Vickery, J., 2017. How does risk management influence production decisions? Evidence from a field experiment. Review of Financial Studies 30, 1935–1970.

Cong, L. W., Liang, T., Zhang, X., 2019. Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information, Working Paper, Cornell University.

Cong, L. W., Tang, K., Wang, J., Zhang, Y., 2020. AlphaPortfolio for investment and economically interpretable AI, Working Paper, Cornell University.

Cong, L. W., Tang, K., Wang, J., Zhang, Y., 2021. Deep sequence modeling: Development and applications in asset pricing. Journal of Financial Data Science 3, 28–42.

Cummins, J. D., Lalonde, D., Phillips, R. D., 2004. The basis risk of catastrophic-loss index securities. Journal of Financial Economics 71, 77–111.

Engle, R. F., Giglio, S., Kelly, B. T., Lee, H., Stroebel, J., 2019. Hedging climate change news. Review of Financial Studies 33, 1184–1216.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo: A test of new factors. Journal of Finance 75, 1327–1370.

Feng, G., Polson, N., Xu, J., 2021. Deep learning in characteristics-sorted factor models, Working Paper, University of Chicago.

Fisher, A. C., Hanemann, W. M., Roberts, M. J., Schlenker, W., 2012. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment. American Economic Review 102, 3749–3760.

Freyberger, J., Neuhierl, A., Weber, M., 2020. Dissecting characteristics nonparametrically. Review of Financial Studies 33, 2326–2377.

Fritsch, F. N., Carlson, R. E., 1980. Monotone piecewise cubic interpolation. SIAM Journal on Numerical Analysis 17, 238–246.

Giné, X., Townsend, R., Vickery, J., 2007. Statistical analysis of rainfall insurance payouts in southern India. American Journal of Agricultural Economics 89, 1248–1254.

Gu, S., Kelly, B., Xiu, D., 2020a. Empirical asset pricing via machine learning. Review of Financial Studies 33, 2223–2273.

Gu, S., Kelly, B. T., Xiu, D., 2020b. Autoencoder asset pricing models. Journal of Econometrics forthcoming.

Harri, A., Coble, K. H., Ker, A. P., Goodwin, B. J., 2011. Relaxing heteroscedasticity assumptions in area-yield crop insurance rating. American Journal of Agricultural Economics 93, 707–717.

Hartman-Glaser, B., Hébert, B., 2020. The insurance is the lemon: Failing to index contracts. Journal of Finance 75, 463–506.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: Data mining, inference and prediction, vol. 27. Springer.

Hong, H., Li, F. W., Xu, J., 2019. Climate risks and market efficiency. Journal of Econometrics 208, 265–281.

Hong, H. G., Karolyi, G. A., Scheinkman, J., 2020. Climate finance. Review of Financial Studies 33, 1011–1023.

Horel, E., Giesecke, K., 2020. Significance tests for neural networks. Journal of Machine Learning Research 21, 1–29.

Hurlin, C., Leymarie, J., Patin, A., 2018. Loss functions for loss given default model comparison. European Journal of Operational Research 268, 348–360.

Illinois Department of Agriculture, 2020. Cover crops premium discount program.

Jensen, N. D., Barrett, C. B., Mude, A. G., 2016. Index insurance quality and basis risk: Evidence from northern Kenya. American Journal of Agricultural Economics 98, 1450–1469.

Kang, W., Rouwenhorst, K. G., Tang, K., 2020. A tale of two premiums: The role of hedgers and speculators in commodity futures markets. Journal of Finance 75, 377–417.

Karlan, D., Osei, R., Osei-Akoto, I., Udry, C., 2014. Agricultural decisions after relaxing credit and risk constraints. Quarterly Journal of Economics 129, 597–652.

Krueger, P., Sautner, Z., Starks, L. T., 2020. The importance of climate risks for institutional investors. Review of Financial Studies 33, 1067–1111.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Lesk, C., Rowhani, P., Ramankutty, N., 2016. Influence of extreme weather disasters on global crop production. Nature 529, 84–88.

Li, W., Tan, K. S., Wei, P., 2020. Demand for non-life insurance under habit formation. Insurance: Mathematics and Economics forthcoming.

Li, Y., Guan, K., Schnitkey, G. D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Global Change Biology 25, 2325–2337.

Lobell, D. B., Bänziger, M., Magorokosho, C., Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nature Climate Change 1, 42–45.

Lobell, D. B., Hammer, G. L., McLean, G., Messina, C., Roberts, M. J., Schlenker, W., 2013. The critical role of extreme heat for maize production in the United States. Nature Climate Change 3, 497–501.

Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., Hammer, G. L., 2014. Greater sensitivity to drought accompanies maize yield increase in the US Midwest. Science 344, 516–519.

Luenberger, D. G., Ye, Y., et al., 1984. Linear and Nonlinear Programming, vol. 2. Springer.

Mahul, O., Skees, J., 2007. Managing Agricultural Risk at the Country Level: The Case of Index-based Livestock Insurance in Mongolia. The World Bank.

Murfin, J., Spiegel, M., 2019. Is the risk of sea level rise capitalized in residential real estate? Review of Financial Studies 33, 1217–1255.

Nordhaus, W., 2019. Climate change: The ultimate challenge for economics. American Economic Review 109, 1991–2014.

Painter, M., 2020. An inconvenient cost: The effects of climate change on municipal bonds. Journal of Financial Economics 135, 468–482.

Pérez-González, F., Yun, H., 2013. Risk management and firm value: Evidence from weather derivatives. Journal of Finance 68, 2143–2176.

Prechelt, L., 1998. Automatic early stopping using cross validation: Quantifying the criteria. Neural Networks 11, 761–767.

Rigden, A., Mueller, N., Holbrook, N., Pillai, N., Huybers, P., 2020. Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting US maize yields. Nature Food 1, 127–133.

Roberts, M. J., Schlenker, W., Eyer, J., 2013. Agronomic weather measures in econometric models of crop yield with implications for climate change. American Journal of Agricultural Economics 95, 236–243.

Rosa, I., 2018a. Farm bill primer: Federal crop insurance. Congressional Research Service.

Rosa, I., 2018b. Federal crop insurance: Program overview for the 115th congress. Congressional Research Service.

Rouwenhorst, K. G., Tang, K., 2012. Commodity investing. Annual Review of Financial Economics 4, 447–467.

Schlenker, W., Roberts, M. J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proceedings of the National Academy of Sciences 106, 15594–15598.

Sirignano, J., Giesecke, K., 2019. Risk analysis for large pools of loans. Management Science 65, 107–121.

Sirignano, J. A., Sadhwani, A., Giesecke, K., 2018. Deep learning for mortgage risk. Journal of Financial Econometrics forthcoming.

Smith, V. H., 2017. Achieving rational farm subsidy rates. R street policy study.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15, 1929–1958.

Suyker, A. E., Verma, S. B., 2008. Interannual water vapor and energy exchange in an irrigated maize-based agroecosystem. Agricultural and Forest Meteorology 148, 417–427.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.

USDA, 2014. World agricultural supply and demand estimates report (WASDE). Office of the chief economist (OCE), United States Department of Agriculture .

Voosen, P., 2019. A 2-week weather forecast may be as good as it gets. Science 363, 801–801.

Weagley, D., 2019. Financial sector stress and risk sharing: Evidence from the weather derivatives market. Review of Financial Studies 32, 2456–2497.

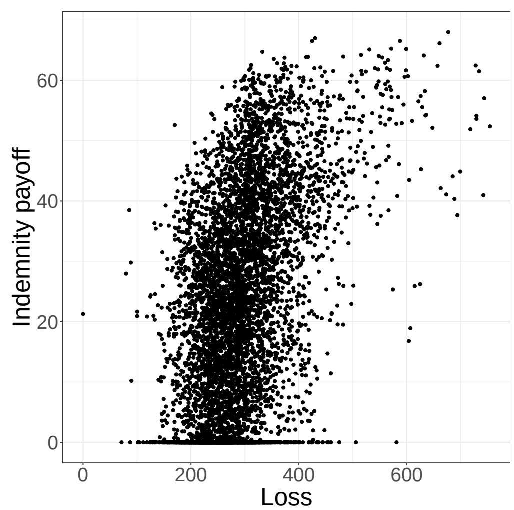
Yohai, V. J., 1987. High breakdown-point and high efficiency robust estimates for regression. Annals of Statistics 15, 642–656.

Young, V. R., 2014. Premium principles. Wiley StatsRef: Statistics Reference Online.

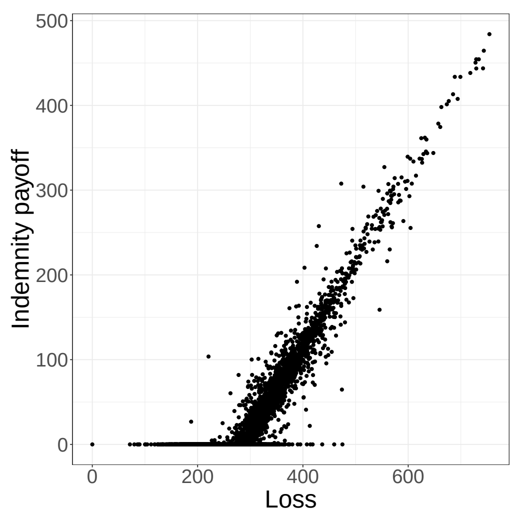
Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S.-J., Chen, J.-H., Emanuel, K., 2019. What is the predictability limit of midlatitude weather? Journal of the Atmospheric Sciences 76, 1077–1091.

Zhang, J., Tan, K. S., Weng, C., 2018. Index insurance design. ASTIN Bulletin 49, 491–523.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67, 301–320.

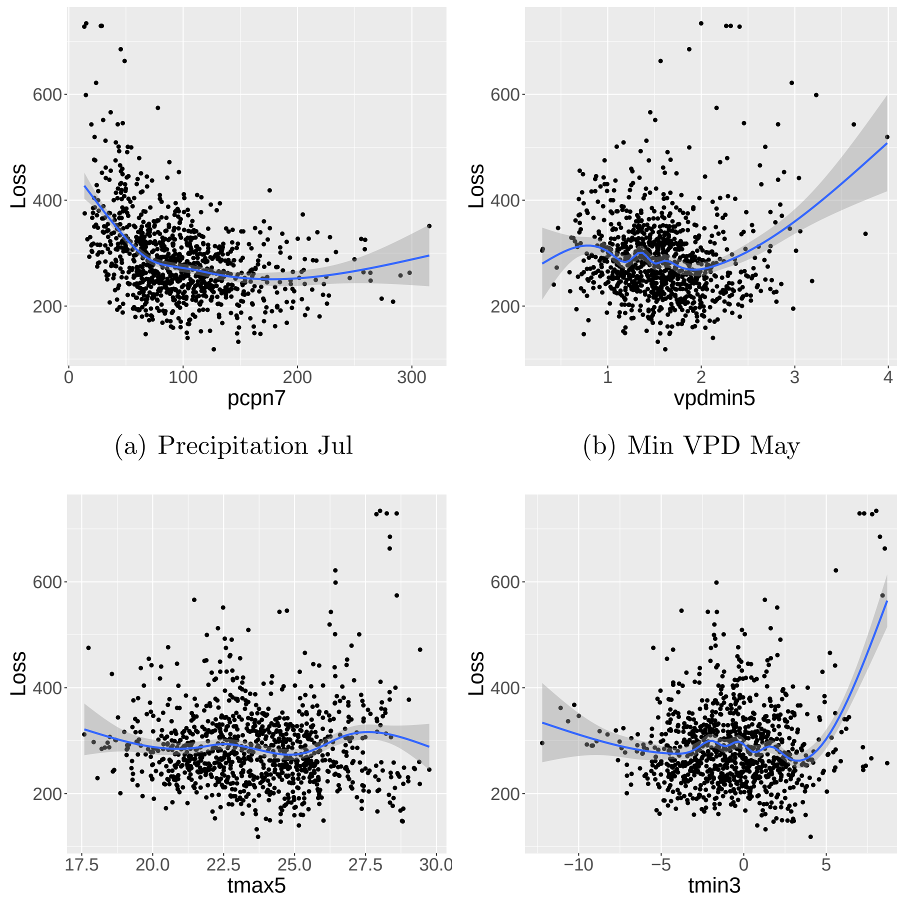


(a) A piecewise linear contract



(b) An optimal NN-based contract

Figure 1: **Payoffs of index insurance and actual losses.** These panels plot the insurance payoffs against actual losses to illustrate the basis risk. Panel (a) plots a piecewise-linear insurance contract based on a rainfall index, which is commonly used in practice. Panel (b) plots an NN-based index insurance contract.



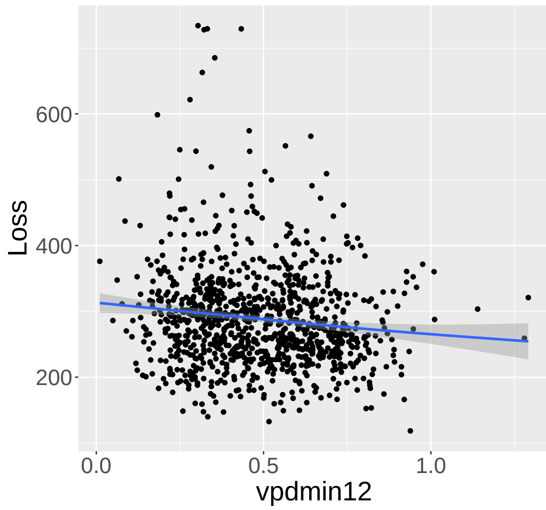
(a) Precipitation Jul

(b) Min VPD May

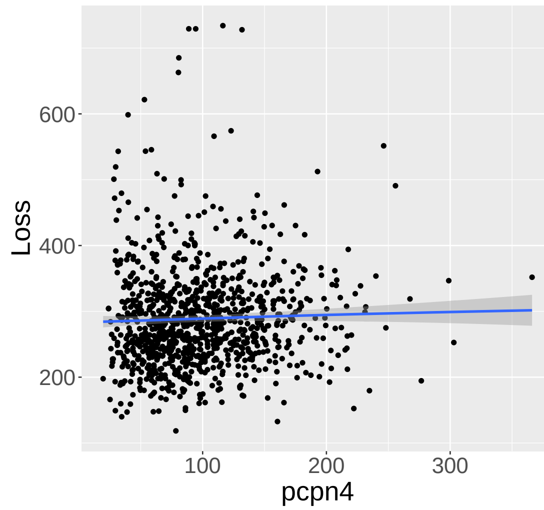
(c) Max Temperature May

(d) Min Temperature March

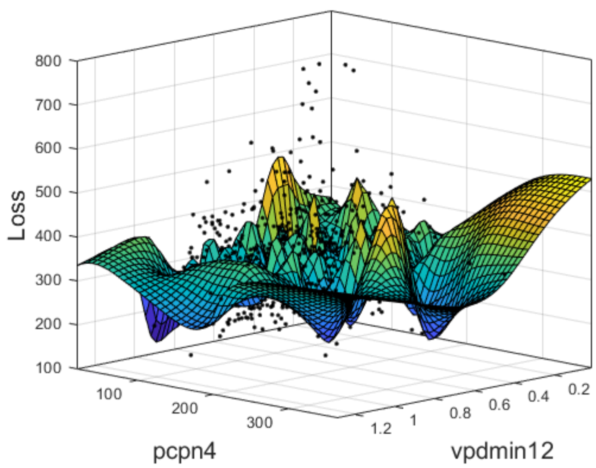
Figure 2: **Nonlinear relationships between crop losses and weather indices.** This figure visualizes crop losses with four selected weather indices, based on 1,000 random draws from the sample. The blue curve is fitted by a generalized additive model. The shadow area indicates a 95% confidence interval. Appendix G provides scatterplots for all 72 weather indices.



(a) Minimum Vapor Pressure Deficit Dec



(b) Precipitation Apr



(c) Minimum Vapor Pressure Deficit Dec and Precipitation Apr

Figure 3: **Individual and joint effects of weather indices on crop losses.** These panels plot production losses against the minimum vapor pressure deficit in December and precipitation in April, using 1,000 random draws from the sample. Panels (a) and (b) show the individual effect of each weather index on production losses. The blue curve is fitted by a generalized additive model. The shadow area indicates a 95% confidence interval. Panel (c) shows the joint effects of these two weather indices on production losses. The surface is fitted by a thin-plate smoothing spline model.

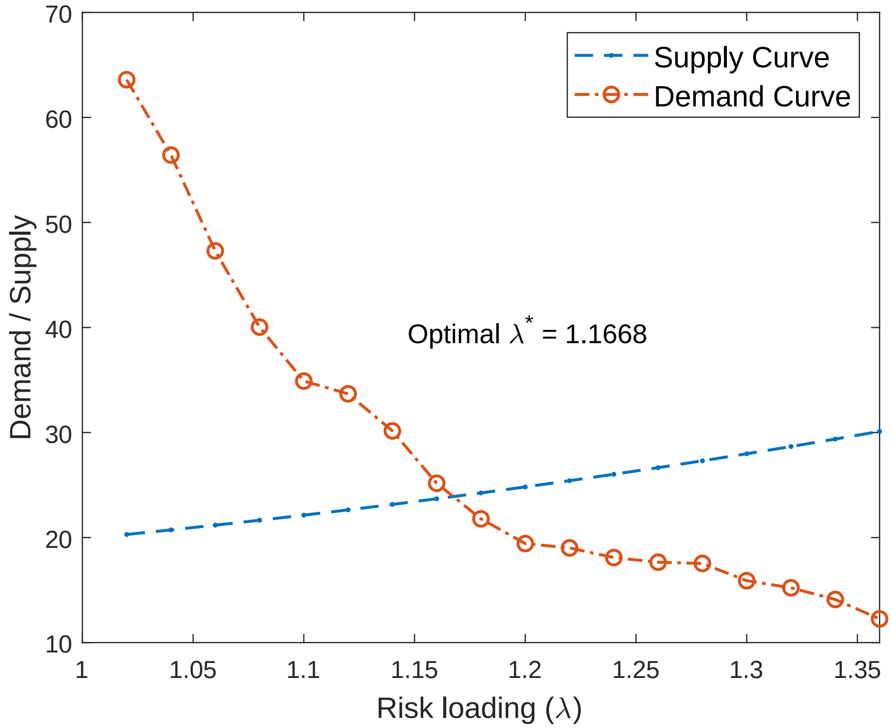
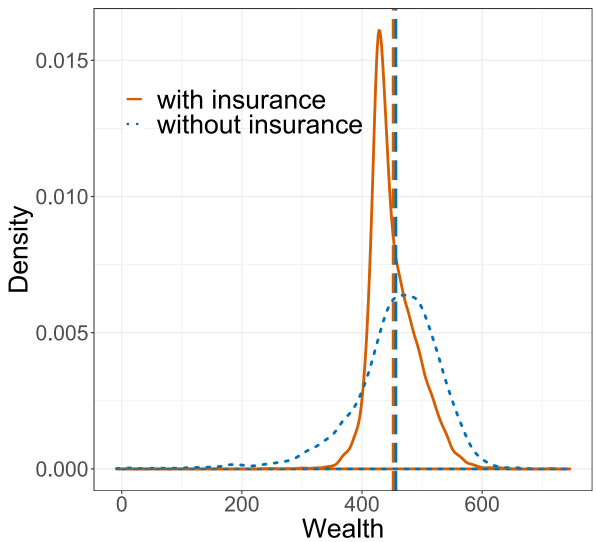
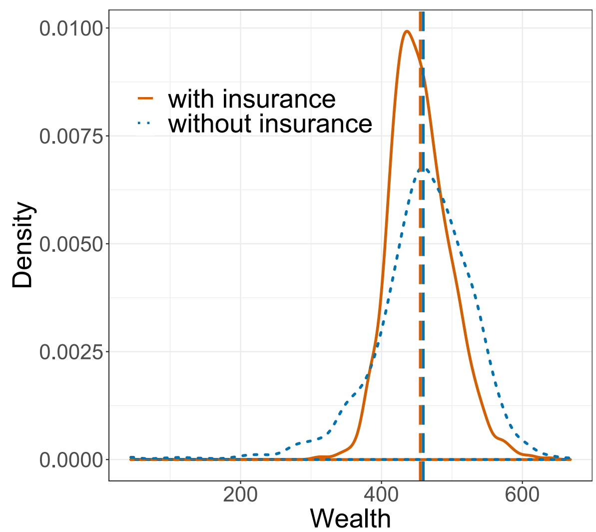


Figure 4: **The supply and demand curves of the index insurance.** This figure displays the supply and demand curves of the index insurance. Their intersection gives the loading parameter λ^* at market equilibrium. The insurance supply curve is fitted from the USDA SOB Reports data with a power function using the nonlinear least squares method. The demand curve is for an NN-based optimal index insurance with a 3-hidden-layer (64-64-16 neurons) structure and is fitted with a piecewise cubic hermite interpolating polynomial. The equilibrium loading parameter is $\lambda^* = 1.1668$.

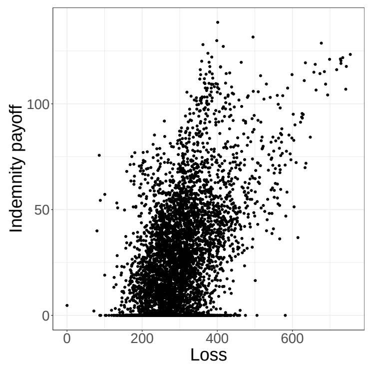


(a) Training sample

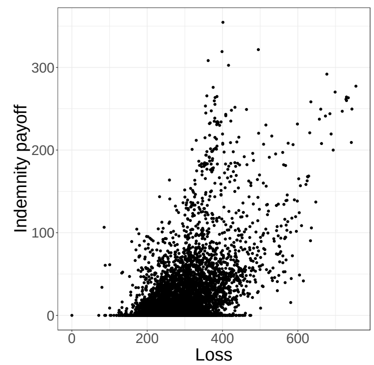


(b) Test sample

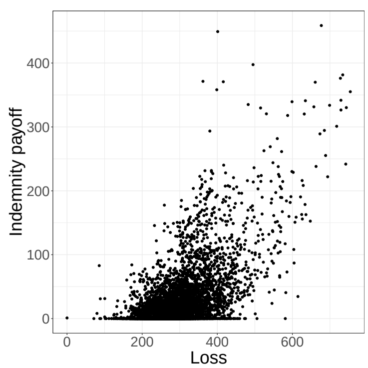
Figure 5: **Farmers' wealth distribution: With index insurance versus without insurance.** The blue-dashed curve represents the probability density of wealth without using insurance, and the red-solid curve represents the probability density of wealth using the optimal NN-based index insurance. Panel (a) plots the training sample, and Panel (b) plots the test sample.



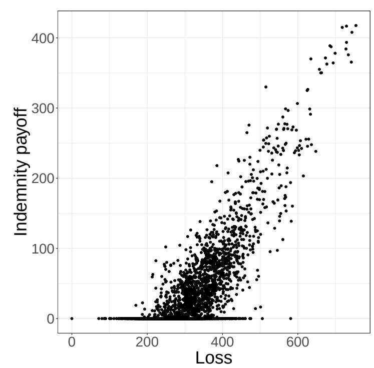
(a) Linear5



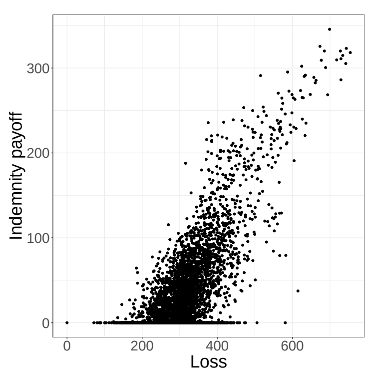
(b) Quadratic5



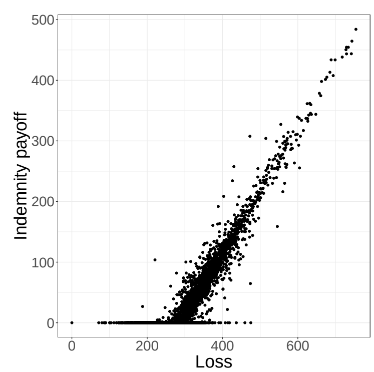
(c) Cubic5



(d) NN5

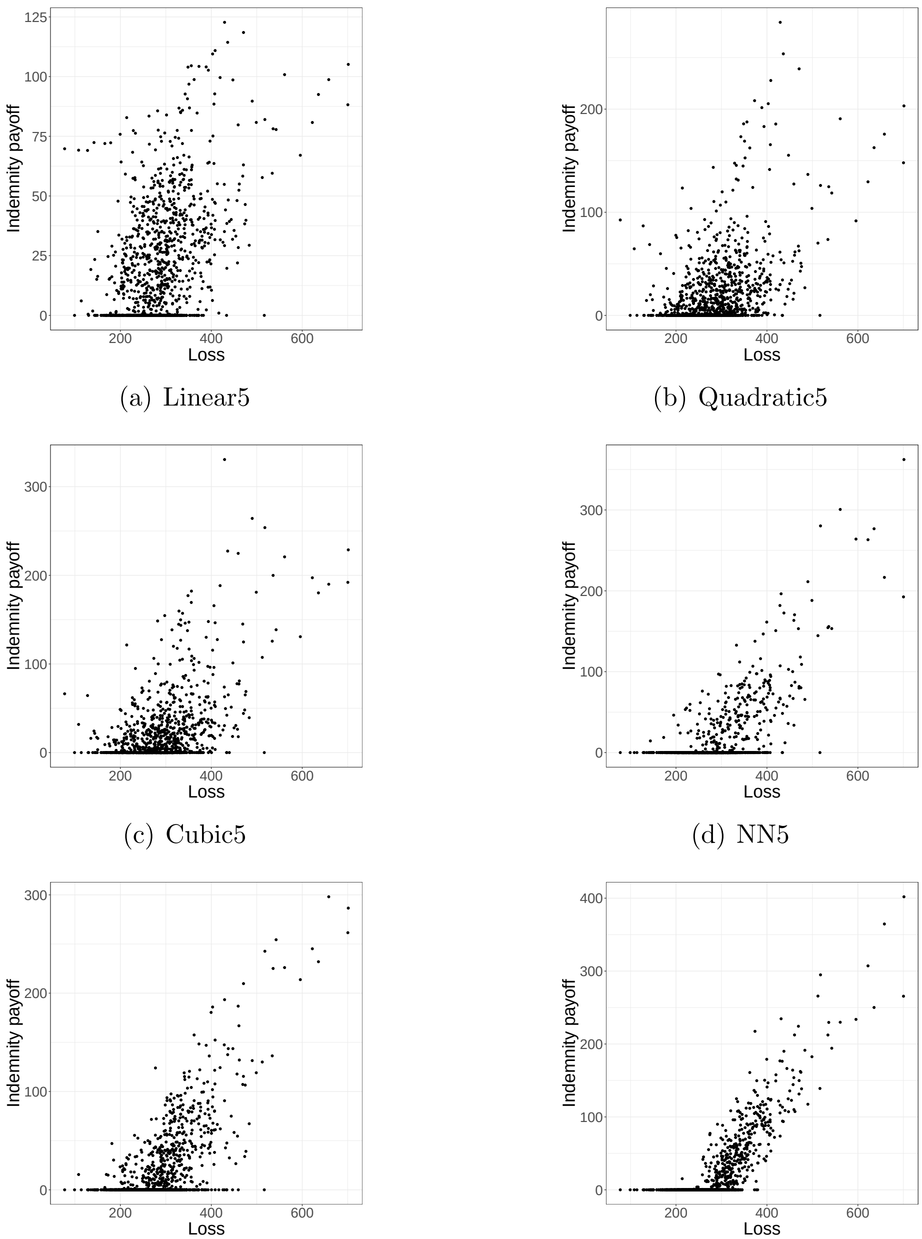


(e) Linear72



(f) NN72

Figure 6: **Basis risk of various index insurance contracts, using the training sample.** These panels plot the insurance payoffs against actual loss, using the training sample. Six insurance contracts are presented, including (a) a linear insurance contract with five weather indices (Linear5); (b) a quadratic insurance contract with five weather indices (Quadratic5); (c) a cubic insurance contract with five weather indices (Cubic5); (d) an NN-based contract with five weather indices (NN5); (e) a linear insurance contract with 72 weather indices (Linear72); and (f) the baseline model (NN72, an NN-based contract with 72 weather indices).



(a) Linear5 (b) Quadratic5

(c) Cubic5 (d) NN5

(e) Linear72 (f) NN72

Figure 7: **Basis risk of various index insurance contracts, using the test sample.** These panels plot the insurance payoffs against actual loss, using the test sample. Six insurance contracts are presented, including (a) a linear insurance contract with five weather indices (Linear5); (b) a quadratic insurance contract with five weather indices (Quadratic5); (c) a cubic insurance contract with five weather indices (Cubic5); (d) an NN-based contract with five weather indices (NN5); (e) a linear insurance contract with 72 weather indices (Linear72); and (f) the baseline model (NN72, an NN-based contract with 72 weather indices).

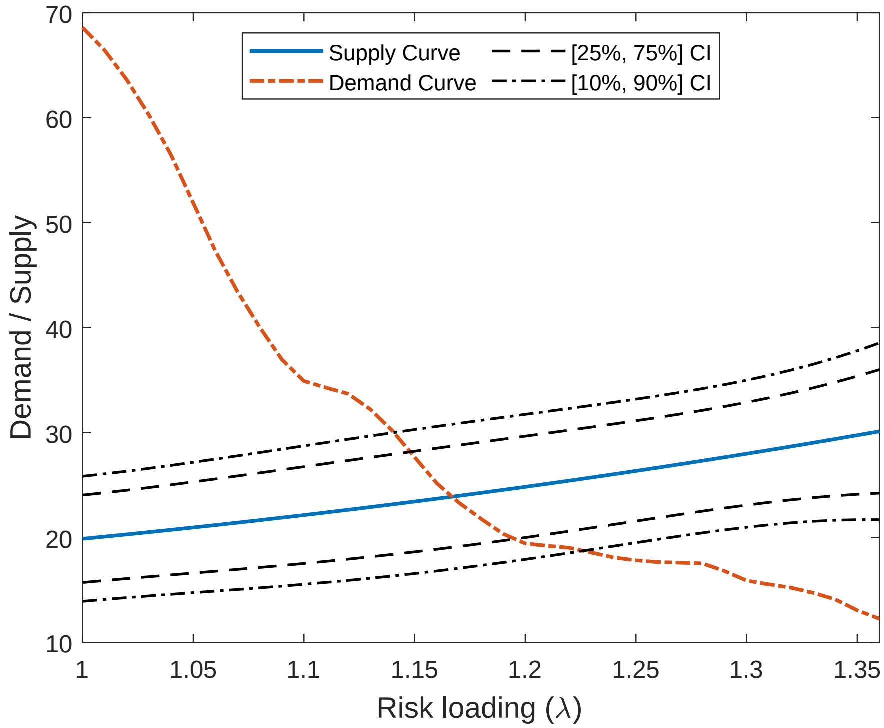
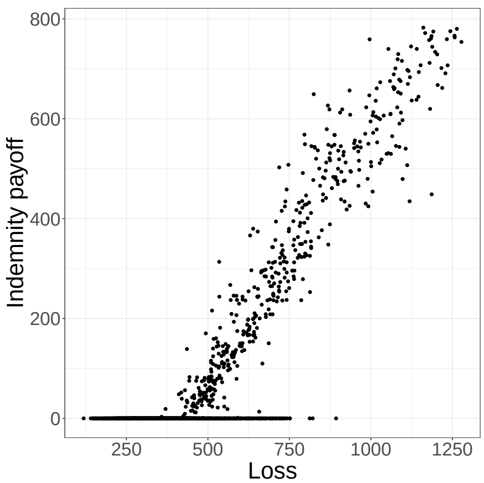
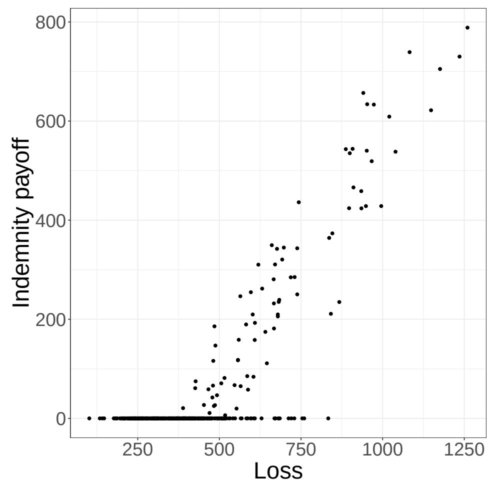


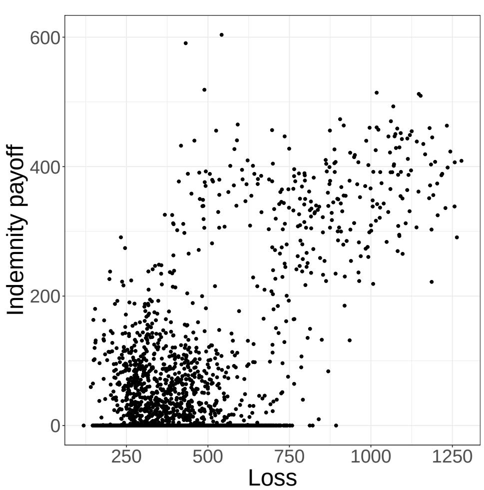
Figure 8: **Confidence intervals for the insurance supply curve.** This figure displays the upper and lower bounds of the [10%, 90%] and [25%, 75%] confidence intervals for the estimated supply curve of the index insurance. The insurance supply curve is fitted from the USDA SOB Reports data with a power function using the nonlinear least squares method. The demand curve is for the NN-based optimal index insurance with a 3-hidden-layer (64-64-16 neurons) structure, and is fitted with a piecewise cubic hermite interpolating polynomial.



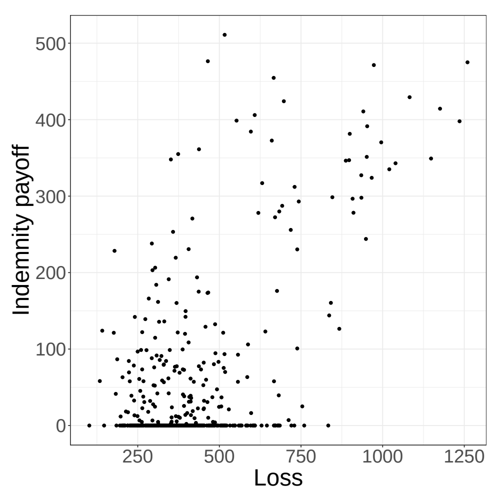
(a) NN72, training set



(b) NN72, test set



(c) Linear72, training set



(d) Linear72, test set

Figure 9: **Basis risk of index insurance protecting both production and price risks.** These figures plot the insurance payoffs against actual loss, over the training or test set. The index insurance is designed to protect both production and price risks. The *NN72* contract has a 3-hidden-layer (64-64-16 neurons) structure, as in the baseline model. The *Linear72* is a linear contract with 72 weather indices.

Table 1: **Weather indices.** This table summarizes the weather variables available from the PRISM data set. The sample period is 1925-2018.

Variable	Description
pcpn_k	Total precipitation (rain+melted snow) for month k (mm)
tmax_k	Daily maximum temperature averaged over all days in month k (°C)
tmin_k	Daily minimum temperature averaged over all days in month k (°C)
dpt_k	Daily mean dew point temperature averaged over all days in month k (°C)
vpdmax_k	Daily maximum vapor pressure deficit averaged over all days in month k (hPa)
vpdmin_k	Daily minimum vapor pressure deficit averaged over all days in month k (hPa)
k	Calender month, i.e., Jan - Dec

Table 2: **Selecting an NN model**. This table summarizes the performances of several NN-based index insurance policies. Panel A shows expected utility and certainty equivalent wealth (CEW) without index insurance, over the training and validation samples. Panels B - E present results with different NN structures. For example, "16-8" in Panel C indicates 16 and 8 neurons in the first and second hidden layers, respectively. Each panel reports the expected utility, percentage utility improvement, CEW, CEW improvement in dollars, CEW improvement in percentage, and insurance premium for each optimal insurance policy, for both training and validation samples. The risk loading parameter at equilibrium (λ^*) is reported in parentheses. The insurance policy with the largest utility improvement is in boldface.

Panel A: Without insurance						
	Training	Validation				
U w/o insurance	-4.41	-4.16				
CEW w/o insurance	417.92	425.22				
Panel B: 1-Hidden-Layer NN						
	2 ($\lambda^* = 1.1411$)		8 ($\lambda^* = 1.1549$)		64 ($\lambda^* = 1.1779$)	
	Training	Validation	Training	Validation	Training	Validation
U with insurance	-3.63	-3.66	-3.59	-3.64	-3.57	-3.63
U improvement	17.81%	12.11%	18.62%	12.61%	19.18%	12.90%
CEW with insurance	442.44	441.36	443.68	442.06	444.53	442.47
CEW improvement	24.52	16.14	25.76	16.85	26.61	17.26
CEW improvement (%)	5.87%	3.80%	6.16%	3.96%	6.37%	4.06%
Premium	25.08	23.62	30.02	29.53	27.61	27.39
Panel C: 2-Hidden-Layers NN						
	16-8 ($\lambda^* = 1.1686$)		64-8 ($\lambda^* = 1.1667$)		64-16 ($\lambda^* = 1.1607$)	
	Training	Validation	Training	Validation	Training	Validation
U with insurance	-3.57	-3.64	-3.56	-3.65	-3.57	-3.63
U improvement	19.10%	12.61%	19.26%	12.40%	19.12%	12.77%
CEW with insurance	444.41	442.07	444.66	441.77	444.45	442.30
CEW improvement	26.50	16.85	26.74	16.55	26.53	17.08
CEW improvement (%)	6.34%	3.96%	6.40%	3.89%	6.35%	4.02%
Premium	27.35	26.88	26.33	25.02	26.72	26.48
Panel D: 3-Hidden-Layers NN						
	64-16-8 ($\lambda^* = 1.1667$)		64-16-16 ($\lambda^* = 1.1755$)		**64-64-16 ($\boldsymbol{\lambda^* = 1.1668}$)**	
	Training	Validation	Training	Validation	Training	Validation
U with insurance	-3.56	-3.63	-3.57	-3.63	**-3.56**	**-3.62**
U improvement	19.45%	12.81%	19.04%	12.76%	**19.30%**	**13.03%**
CEW with insurance	444.96	442.35	444.32	442.29	**444.72**	**442.66**
CEW improvement	27.04	17.13	26.40	17.07	**26.80**	**17.45**
CEW improvement (%)	6.47%	4.03%	6.32%	4.01%	**6.41%**	**4.10%**
Premium	29.25	28.83	25.61	24.90	**27.56**	**27.23**
Panel E: 4-Hidden-Layers NN						
	64-16-8-8 ($\lambda^* = 1.1609$)		64-16-8-8 ($\lambda^* = 1.1637$)		64-16-8-8 ($\lambda^* = 1.1756$)	
	Training	Validation	Training	Validation	Training	Validation
U with insurance	-3.56	-3.64	-3.55	-3.63	-3.55	-3.63
U improvement	19.33%	12.67%	19.62%	12.89%	19.68%	12.76%
CEW with insurance	444.77	442.16	445.23	442.46	445.32	442.28
CEW improvement	26.85	16.94	27.31	17.24	27.40	17.07
CEW improvement (%)	6.43%	3.98%	6.53%	4.06%	6.56%	4.01%
Premium	26.91	26.66	32.03	31.39	30.87	29.98

Table 3: **Performances of the baseline insurance contract and alternative contracts** This table summarizes the performances of seven insurance contracts in the test sample, including (1) a linear insurance contract with one weather index (*Linear1*), which corresponds to a currently used contract; (2) a linear insurance contract with five weather indices (*Linear5*); (3) a quadratic insurance contract with five weather indices (*Quadratic5*); (4) a cubic insurance contract with five weather indices (*Cubic5*); (5) an NN-based contract with five weather indices (*NN5*); (6) a linear insurance contract with 72 weather indices (*Linear72*); and (7) the baseline model (*NN72*, an NN-based contract with 72 weather indices). Panel A summarizes expected utilities with and without (w/o) index insurance and the percentage of utility improvement. Panel B reports certainty equivalent wealth (CEW) with and without (w/o) index insurance and the CEW improvement in dollars and as a percentage. Panel C summarizes policy characteristics including policy premium, coverage, and profits of the insurer. Panel D summarizes the risk reduction effect of an index insurance policy, measured by the standard deviation of wealth. Panel E summarizes the tail risk reduction, measured by the 5%- and 1%-level value-at-risk (VaR). The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

Contract	Linear1 ($\lambda^* = 1.0628$)	Linear5 ($\lambda^* = 1.0829$)	Quadradic5 ($\lambda^* = 1.0855$)	Cubic5 ($\lambda^* = 1.0747$)	NN5 ($\lambda^* = 1.0640$)	Linear72 ($\lambda^* = 1.1337$)	NN72 (BL, $\lambda^* = 1.1668$))
Panel A: Utility improvement							
U with insurance	-3.83	-3.76	-3.69	-3.62	-3.54	-3.57	-3.47
U w/o insurance	-3.95	-3.95	-3.95	-3.95	-3.95	-3.95	-3.95
U improvement (%)	3.11%	4.97%	6.76%	8.34%	10.45%	9.75%	12.33%
Panel B: CEW improvement							
CEW with insurance	435.68	438.10	440.48	442.61	445.52	444.55	448.18
CEW w/o insurance	431.73	431.73	431.73	431.73	431.73	431.73	431.73
CEW improvement	3.96	6.37	8.75	10.88	13.79	12.82	16.45
CEW improvement (%)	0.92%	1.48%	2.03%	2.52%	3.19%	2.97%	3.81%
Panel C: Policy characteristics							
Premium	22.10	22.57	21.75	21.07	14.91	22.87	28.82
Coverage	20.80	20.84	20.03	19.60	13.46	20.17	24.70
Insurer Profit	1.31	1.73	1.71	1.46	1.44	2.70	4.12
Panel D: Risk reduction measured by standard deviation							
Std	64.96	63.69	62.53	59.87	54.66	53.07	42.60
Std w/o insurance	70.78	70.78	70.78	70.78	70.78	70.78	70.78
Std reduction	8.21%	10.01%	11.65%	15.41%	22.77%	25.02%	39.80%
Panel E: Risk reduction measured by Value-at-Risk (VaR)							
$\text{VaR}_{5\%}$	349.60	353.48	351.28	353.68	365.70	364.91	391.53
$\text{VaR}_{5\%}$ w/o insurance	341.99	341.99	341.99	341.99	341.99	341.99	341.99
$\text{VaR}_{5\%}$ improvement	7.61	11.49	9.29	11.69	23.71	22.92	49.54
$\text{VaR}_{1\%}$	268.78	268.87	288.91	295.43	331.28	306.43	370.54
$\text{VaR}_{1\%}$ w/o insurance	232.56	232.56	232.56	232.56	232.56	232.56	232.56
$\text{VaR}_{1\%}$ improvement	36.23	36.32	56.36	62.87	98.72	73.87	137.99

Table 4: **Identifying important weather indices.** This table shows the 10 most important weather indices ranked by the absolute correlations between insurance payoffs and weather indices (left panel) or the gradient-based sensitivities of an NN-based insurance payoffs to the weather indices (right panel). Indices that only appear in one top-10 list are in boldface. Column "Rank Diff" shows the difference between two ranks for a weather index. See Table 1 for the index variable descriptions.

$Rank^{corr}$	Index	Absolute correlations	Rank diff $Rank_g^{NN} - Rank^{corr}$	$Rank_g^{NN}$	Index	NN policy absolute $\bar{\mathcal{S}}^g$	Rank diff $Rank_g^{NN} - Rank^{corr}$
1	vpdmax7	0.43	1	1	pcpn7	14.50	-1
2	pcpn7	0.39	-1	2	vpdmax7	11.78	1
3	**vpdmin7**	0.38	18	3	**dpt1**	8.37	-15
4	**vpdmax8**	0.38	45	4	**tmin2**	7.15	-38
5	**tmax7**	0.37	55	5	**tmax11**	5.86	-38
6	tmax8	0.33	2	6	**tmax10**	5.70	-54
7	**vpdmin8**	0.29	44	7	**vpdmax1**	5.67	-65
8	**tmin7**	0.24	15	8	tmax8	5.45	2
9	**tmin8**	0.24	63	9	**pcpn6**	5.28	-25
10	**vpdmax6**	0.21	46	10	**pcpn8**	5.13	-3

Table 5: **Out-of-state tests, using adjacent states.** We perform out-of-state tests for the NN-based index insurance. We use Illinois data as the training set to estimate the contract parameters. We use data from three states adjacent to Illinois and that have similar latitudes, namely, Indiana, Kentucky, and Missouri, as the test sets. The NN uses a 3-hidden-layer (64-64-16 neurons) structure. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance policies and the CEW improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurer. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR).

Data	Training set	Test set		
States	Illinois	Indiana	Kentucky	Missouri
Panel A: Utility improvement				
U with insurance	-3.51	-2.92	-9.04	-2.56
U w/o insurance	-4.17	-3.19	-9.71	-2.80
U improvement (%)	15.75%	8.63%	6.82%	8.69%
Panel B: CEW improvement				
CEW with insurance	446.46	469.78	328.29	486.04
CEW w/o insurance	425.04	458.51	319.45	474.67
CEW improvement	21.43	11.28	8.84	11.36
CEW improvement (%)	5.04%	2.46%	2.77%	2.39%
Panel C: Policy characteristics				
Premium	29.21	23.69	14.36	27.55
Coverage	25.03	20.30	12.30	23.61
Insurer's profit	4.18	3.39	2.05	3.94
Panel D: Risk reduction measured by standard deviation				
Std	40.52	51.44	69.37	61.78
Std w/o insurance	74.61	67.99	80.38	79.61
Std reduction (%)	45.69%	24.34%	13.69%	22.40%
Panel E: Risk reduction measured by Value-at-Risk (VaR)				
$\text{VaR}_{5\%}$	400.53	393.78	231.22	400.91
$\text{VaR}_{5\%}$ w/o insurance	326.26	359.30	208.05	361.73
$\text{VaR}_{5\%}$ improvement	74.27	34.48	23.17	39.18
$\text{VaR}_{1\%}$	375.19	346.18	146.12	335.36
$\text{VaR}_{1\%}$ w/o insurance	195.84	264.45	85.21	260.13
$\text{VaR}_{1\%}$ improvement	179.35	81.73	60.91	75.23

Table 6: **Impacts of coverage level.** We consider an NN-based index insurance with various coverage levels. Panel A summarizes utilities with and without (w/o) index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance policies and CEW improvements in dollars and as a percentage. Panel C summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel D summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). "BL" indicates the baseline case studied in Section 4.2.

	Coverage: \$10		Coverage: \$20		Coverage: \$27.16 (BL)		Coverage: \$30		Coverage: \$40	
Data	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement										
U with insurance	-3.64	-3.56	-3.58	-3.50	-3.52	-3.47	-3.57	-3.49	-3.62	-3.53
U w/o insurance	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95
U improvement (%)	13.56%	10.03%	14.83%	11.43%	16.33%	12.33%	15.12%	11.66%	13.99%	10.63%
Panel B: CEW improvement										
CEW with insurance	442.11	444.94	443.96	446.89	446.18	448.18	444.38	447.23	442.73	445.77
CEW w/o insurance	423.89	431.73	423.89	431.73	423.89	431.73	423.89	431.73	423.89	431.73
CEW improvement	18.21	13.21	20.06	15.17	22.28	16.45	20.49	15.50	18.84	14.04
CEW improvement (%)	4.30%	3.06%	4.73%	3.51%	5.26%	3.81%	4.83%	3.59%	4.44%	3.25%
Panel C: Risk reduction measured by standard deviation										
Std	56.18	56.87	48.10	49.31	38.96	42.60	42.48	44.12	41.00	42.04
Std w/o insurance	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78
Std reduction (%)	25.36%	19.65%	36.09%	30.34%	48.24%	39.80%	43.55%	37.66%	45.52%	40.60%
Panel D: Risk reduction measured by Value-at-Risk (VaR)										
$\text{VaR}_{5\%}$	362.22	367.41	377.31	381.66	403.33	391.53	385.05	387.34	382.15	386.06
$\text{VaR}_{5\%}$ w/o insurance	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99
$\text{VaR}_{5\%}$ improvement	38.65	25.42	53.74	39.67	79.75	49.54	61.48	45.35	58.58	44.07
$\text{VaR}_{1\%}$	324.82	321.79	341.75	336.66	373.09	370.54	351.38	343.50	345.11	345.16
$\text{VaR}_{1\%}$ w/o insurance	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56
$\text{VaR}_{1\%}$ improvement	133.88	89.24	150.81	104.10	182.15	137.99	160.44	110.95	154.17	112.61

Table 7: **Impacts of risk aversion.** We consider an NN-based index insurance for farmers with various levels of risk aversion, i.e., relative risk aversion (RRA) of 2, 3.1, 4, and 5. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes CEW with and without (w/o) index insurance policies and certainty equivalent wealth (CEW) improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurers. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). "BL" represents the baseline case studied in Section 4.2. The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

Risk aversion	$\alpha = 0.0051$ (RRA = 2) ($\lambda^* = 1.1203$)		$\alpha = 0.008$ (BL, RRA = 3.1) ($\lambda^* = 1.1668$)		$\alpha = 0.0103$ (RRA = 4) ($\lambda^* = 1.2116$)		$\alpha = 0.0129$ (RRA = 5) ($\lambda^* = 1.2368$)	
Data	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement								
U with insurance	-18.59	-18.37	-3.52	-3.47	-1.02	-1.00	-0.27	-0.26
U w/o insurance	-19.69	-19.15	-4.21	-3.95	-1.45	-1.30	-0.54	-0.45
U improvement (%)	5.60%	4.03%	16.33%	12.33%	29.56%	22.79%	49.35%	41.00%
Panel B: CEW improvement								
CEW with insurance	449.31	451.57	446.18	448.18	442.31	443.83	438.41	440.39
CEW w/o insurance	438.22	443.65	423.89	431.73	408.28	418.72	385.67	399.49
CEW improvement	11.09	7.92	22.28	16.45	34.02	25.11	52.74	40.90
CEW improvement (%)	2.53%	1.79%	5.26%	3.81%	8.33%	6.00%	13.67%	10.24%
Panel C: Policy characteristics								
Premium	28.52	25.76	31.69	28.82	28.26	24.53	27.02	23.34
Coverage	25.46	22.99	27.16	24.70	23.32	20.25	21.84	18.87
Insurer's profit	3.06	2.77	4.53	4.12	4.93	4.28	5.17	4.47
Panel D: Risk reduction measured by standard deviation								
Std	40.12	43.87	38.96	42.60	42.89	46.50	44.35	47.32
Std w/o insurance	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78
Std reduction (%)	46.69%	38.02%	48.24%	39.80%	43.01%	34.30%	41.07%	33.14%
Panel E: Risk reduction measured by Value-at-Risk (VaR)								
$\text{VaR}_{5\%}$	404.57	393.79	403.33	391.53	394.71	388.88	391.59	385.54
$\text{VaR}_{5\%}$ w/o insurance	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99
$\text{VaR}_{5\%}$ improvement	81.00	51.80	79.75	49.54	71.14	46.88	68.02	43.55
$\text{VaR}_{1\%}$	375.49	366.82	373.09	370.54	352.19	347.67	348.37	345.40
$\text{VaR}_{1\%}$ w/o insurance	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56
$\text{VaR}_{1\%}$ improvement	184.55	134.27	182.15	137.99	161.25	115.12	157.43	112.84

Table 8: **Impact of time-varying risk aversion.** We evaluate the index insurance performance with time-varying risk aversion. The farmer's average relative risk aversion is $RRA = 3.1$. If the farmer experiences a loss larger than the 75^{th} percentile in year $t-1$, her risk aversion in year t becomes $3.1 \times (1+x)$; on the contrary, if the farmer experiences a loss lower than the 25^{th} percentile in year $t-1$, her risk aversion in year t is $3.1 \times (1-x)$. Columns 2-7 display results for different risk aversion variations ($x = 0.1, 0.2$, and 0.3). The last two columns correspond to a constant risk aversion of 3.1, which is our baseline model. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes CEW with and without (w/o) index insurance policies and certainty equivalent wealth (CEW) improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurers. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR).

Risk aversion	$x = 0.1$		$x = 0.2$		$x = 0.3$		Baseline model	
	$RRA = 2.8, 3.1, 3.4$		$RRA = 2.5, 3.1, 3.7$		$RRA = 2.2, 3.1, 4$		RRA = 3.1	
Data	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement								
U with insurance	-3.77	-3.81	-4.65	-4.88	-6.42	-6.91	-3.52	-3.47
U w/o insurance	-4.49	-4.30	-5.47	-5.31	-7.41	-7.28	-4.21	-3.95
U improvement (%)	16.03%	11.34%	14.98%	8.22%	13.40%	5.11%	16.33%	12.33%
Panel B: CEW improvement								
CEW with insurance	445.67	447.62	443.14	444.67	439.89	441.44	446.18	448.18
CEW w/o insurance	424.03	433.82	424.03	433.82	424.03	433.82	423.89	431.73
CEW improvement	21.64	13.80	19.11	10.85	15.86	7.61	22.28	16.45
CEW improvement (%)	5.10%	3.18%	4.51%	2.50%	3.74%	1.75%	5.26%	3.81%
Panel C: Policy characteristics								
Premium	28.08	26.05	31.00	29.98	35.38	33.22	31.69	28.82
Coverage	24.07	22.32	26.57	25.69	30.33	28.47	27.16	24.70
Insurer Profit	4.01	3.72	4.43	4.28	5.06	4.75	4.53	4.12
Panel D: Risk reduction measured by standard deviation								
Std	43.51	46.23	49.36	51.82	56.80	58.32	38.96	42.60
Std w/o insurance	75.27	70.77	75.27	70.77	75.27	70.77	75.26	70.78
Std reduction	42.19%	34.68%	34.42%	26.77%	24.54%	17.59%	48.24%	39.80%
Panel E: Risk reduction measured by Value-at-Risk (VaR)								
$VaR_{5\%}$	388.51	384.33	369.59	375.62	361.21	367.37	403.33	391.53
$VaR_{5\%}$ w/o insurance	323.59	342.59	323.59	342.59	323.59	342.59	323.57	341.99
$VaR_{5\%}$ improvement	64.92	41.73	45.99	33.03	37.61	24.78	79.75	49.54
$VaR_{1\%}$	365.85	354.95	335.21	334.86	312.19	311.37	373.09	370.54
$VaR_{1\%}$ w/o insurance	191.59	232.79	191.59	232.79	191.59	232.79	190.94	232.56
$VaR_{1\%}$ improvement	174.26	122.16	143.61	102.07	120.60	78.58	182.15	137.99

Table 9: **Impacts of insurer's supply curves.** We test the robustness of our results using the upper and lower bounds of [10%, 90%] and [25%, 75%] confidence intervals (CI) of the supply curve estimates. Panel A summarizes utilities with and without (w/o) index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance and the CEW improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurer. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

	[25%, 75%] CI				[10%, 90%] CI			
Supply curve	Lower bound ($\lambda^* = 1.1902$)		Upper bound ($\lambda^* = 1.1502$)		Lower bound ($\lambda^* = 1.2273$)		Upper bound ($\lambda^* = 1.1436$)	
Data	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement								
U with insurance	-3.55	-3.50	-3.51	-3.46	-3.58	-3.51	-3.51	-3.45
U w/o insurance	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95
U improvement (%)	15.60%	11.56%	16.49%	12.51%	15.05%	11.10%	16.60%	12.70%
Panel B: CEW improvement								
CEW with insurance	445.10	447.09	446.41	448.44	444.29	446.44	446.59	448.71
CEW w/o insurance	423.89	431.73	423.89	431.73	423.89	431.73	423.89	431.73
CEW improvement	21.21	15.36	22.52	16.71	20.39	14.71	22.70	16.98
CEW improvement (%)	5.00%	3.56%	5.31%	3.87%	4.81%	3.41%	5.35%	3.93%
Panel C: Policy characteristics								
Premium	28.10	24.49	29.62	26.88	24.34	21.35	30.50	28.03
Coverage	23.61	20.57	25.75	23.37	19.84	17.39	26.67	24.51
Insurer's profit	4.49	3.91	3.87	3.51	4.51	3.95	3.83	3.52
Panel D: Risk reduction measured by standard deviation								
Std	42.13	45.76	40.22	43.52	44.76	47.74	39.78	42.75
Std w/o insurance	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78
Std reduction (%)	44.02%	35.35%	46.56%	38.51%	40.53%	32.55%	47.14%	39.60%
Panel E: Risk reduction measured by Value-at-Risk (VaR)								
$\text{VaR}_{5\%}$	399.03	389.65	400.97	392.40	392.62	387.43	400.74	394.10
$\text{VaR}_{5\%}$ w/o insurance	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99
$\text{VaR}_{5\%}$ improvement	75.45	47.66	77.40	50.41	69.05	45.44	77.17	52.11
$\text{VaR}_{1\%}$	359.78	347.99	372.67	364.31	361.54	350.97	372.56	372.03
$\text{VaR}_{1\%}$ w/o insurance	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56
$\text{VaR}_{1\%}$ improvement	168.84	115.43	181.72	131.76	170.60	118.41	181.62	139.47

Table 10: **Impacts of government subsidies.** We consider an NN-based index insurance with various government subsidy rates, i.e., $\theta = 0, 5\%, 7.5\%, 10\%$, or 14.29%. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance policies and the CEW improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage and profits for the insurer. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). "BL" represents the baseline case studied in Section 4.2. $\theta = 14.29\%$ is the subsidy level such that farmers pay the actuarially fair premium and the government pays related costs and loadings. The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

Subsidy rate	$\theta = 0$ (BL) ($\lambda^* = 1.1668$)		$\theta = 0.05$ ($\lambda^* = 1.2103$)		$\theta = 0.075$ ($\lambda^* = 1.2213$)		$\theta = 0.1$ ($\lambda^* = 1.2678$)		$\theta = 0.1429$ ($\lambda^* = 1.3244$)	
Data	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement										
U with insurance	-3.52	-3.47	-3.52	-3.46	-3.50	-3.45	-3.50	-3.45	-3.50	-3.44
U w/o insurance	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95	-4.21	-3.95
U improvement (%)	16.33%	12.33%	16.43%	12.49%	16.79%	12.61%	16.87%	12.67%	16.84%	12.95%
Panel B: CEW improvement										
CEW with insurance	446.18	448.18	446.33	448.41	446.88	448.57	446.99	448.66	446.94	449.07
CEW w/o insurance	423.89	431.73	423.89	431.73	423.89	431.73	423.89	431.73	423.89	431.73
CEW improvement	22.28	16.45	22.44	16.68	22.98	16.84	23.09	16.93	23.05	17.34
CEW improvement (%)	5.26%	3.81%	5.29%	3.86%	5.42%	3.90%	5.45%	3.92%	5.44%	4.02%
Panel C: Policy characteristics										
Premium	31.69	28.82	30.56	27.87	33.93	31.34	37.17	33.45	37.39	34.57
Coverage	27.16	24.70	25.25	23.02	27.78	25.66	29.32	26.39	28.23	26.10
Insurer's profit	4.53	4.12	5.31	4.84	6.15	5.68	7.85	7.07	9.16	8.47
Panel D: Risk reduction measured by standard deviation										
Std	38.96	42.60	40.68	43.68	39.42	43.34	37.54	42.15	38.57	41.57
Std w/o insurance	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78
Std reduction (%)	48.24%	39.80%	45.95%	38.28%	47.63%	38.77%	50.12%	40.45%	48.75%	41.27%
Panel E: Risk reduction measured by Value-at-Risk (VaR)										
$\text{VaR}_{5\%}$	403.33	391.53	400.95	389.74	401.41	388.91	405.62	394.24	403.85	397.43
$\text{VaR}_{5\%}$ w/o insurance	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99
$\text{VaR}_{5\%}$ improvement	79.75	49.54	77.38	47.75	77.84	46.92	82.05	52.25	80.28	55.44
$\text{VaR}_{1\%}$	373.09	370.54	370.27	359.54	373.35	362.72	379.68	369.60	373.15	370.15
$\text{VaR}_{1\%}$ w/o insurance	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56
$\text{VaR}_{1\%}$ improvement	182.15	137.99	179.33	126.98	182.40	130.17	188.73	137.05	182.21	137.59

Table 11: **Protecting both production and price risks.** We consider an index insurance contracts protecting both production and price risks. The *NN72* contract has the 3-hidden-layer (64-64-16 neurons) structure, as in the baseline model. The *Linear72* is a linear contract with all 72 weather indices. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance policies and the CEW improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurer. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses. The sample period is 1980-2017.

Contract	NN72 ($\lambda^* = 1.1668$)		Linear72 ($\lambda^* = 1.1337$)	
Data	Training	Test	Training	Test
Panel A: Utility improvement				
U with insurance	-1.05	-1.06	-1.69	-1.49
U w/o insurance	-7.19	-5.52	-7.19	-5.52
U improvement (%)	85.42%	80.87%	76.57%	72.99%
Panel B: CEW improvement				
CEW with insurance	597.60	596.66	538.31	553.51
CEW w/o insurance	356.93	389.91	356.93	389.91
CEW improvement	240.67	206.75	181.38	163.60
CEW improvement (%)	67.43%	53.02%	50.82%	41.96%
Panel C: Policy characteristics				
Premium	65.28	53.92	66.96	63.35
Coverage	55.95	46.21	59.06	55.88
Insurer's profit	9.33	7.71	7.90	7.47
Panel D: Risk reduction measured by standard deviation				
Std	105.05	110.16	151.71	151.20
Std w/o insurance	198.50	184.70	198.50	184.70
Std reduction (%)	47.08%	40.36%	23.57%	18.14%
Panel E: Risk reduction measured by Value-at-Risk (VaR)				
$\text{VaR}_{5\%}$	470.94	482.86	358.39	359.16
$\text{VaR}_{5\%}$ w/o insurance	217.08	252.15	217.08	252.15
$\text{VaR}_{5\%}$ improvement	253.86	230.71	141.30	107.01
$\text{VaR}_{1\%}$	336.68	321.77	233.03	266.36
$\text{VaR}_{1\%}$ w/o insurance	-50.07	39.31	-50.07	39.31
$\text{VaR}_{1\%}$ improvement	386.75	282.47	283.10	227.05

Table 12: **Alternative utility functions.** We consider power utility with various levels of risk aversion, i.e., relative risk aversion (RRA) of 2, 3, 4, and 5 and log utility (RRA = 1). The NN uses a 3-hidden-layer (64-64-16 neurons) structure. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes certainty equivalent wealth (CEW) with and without (w/o) index insurance policies and CEW improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurers. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

Utility function	Log Utility		Power Utility							
Risk aversion	RRA = 1 ($\lambda^* = 1.0633$)		RRA = 2 ($\lambda^* = 1.1060$)		RRA = 3 ($\lambda^* = 1.1383$)		RRA = 4 ($\lambda^* = 1.1947$)		RRA = 5 ($\lambda^* = 1.2128$)	
Data	Training	Test	Training	Test	Training	Test	Training	Test	Training	Test
Panel A: Utility improvement										
U with insurance	6.12×10^{-06}	6.12×10^{-06}	-2.22×10^{-03}	-2.21×10^{-03}	-2.50×10^{-06}	-2.48×10^{-06}	-3.80×10^{-09}	-3.74×10^{-09}	-7.43×10^{-12}	-7.23×10^{-12}
U w/o insurance	6.10×10^{-06}	6.10×10^{-06}	-2.29×10^{-03}	-2.29×10^{-03}	-2.85×10^{-06}	-2.85×10^{-06}	-5.75×10^{-09}	-5.75×10^{-09}	-1.86×10^{-11}	-1.86×10^{-11}
U improvement (%)	0.20%	0.30%	2.83%	3.38%	12.17%	13.02%	33.90%	34.95%	60.11%	61.22%
Panel B: CEW improvement										
CEW with insurance	452.99	455.67	450.24	452.79	447.03	449.21	444.25	446.61	428.23	431.28
CEW w/o insurance	447.36	447.36	437.50	437.50	418.94	418.94	386.98	386.98	340.33	340.33
CEW improvement	5.63	8.31	12.74	15.29	28.09	30.27	57.27	59.63	87.89	90.95
CEW improvement (%)	1.26%	1.86%	2.91%	3.50%	6.71%	7.23%	14.80%	15.41%	25.83%	26.72%
Panel C: Policy characteristics										
Premium	22.07	19.62	27.01	23.79	26.30	23.15	29.50	25.98	36.20	32.70
Coverage	20.76	18.45	24.42	21.51	23.10	20.33	24.69	21.74	29.85	26.97
Insurer's profit	1.31	1.17	2.59	2.28	3.20	2.81	4.81	4.23	6.35	5.74
Panel D: Risk reduction measured by standard deviation										
Std	43.98	46.88	40.87	43.92	43.16	46.14	41.66	44.27	55.02	56.57
Std w/o insurance	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78	75.26	70.78
Std reduction (%)	41.57%	33.76%	45.70%	37.95%	42.66%	34.81%	44.65%	37.45%	26.90%	20.07%
Panel E: Risk reduction measured by Value-at-Risk (VaR)										
$VaR_{5\%}$	396.61	390.79	403.15	393.09	395.27	390.76	394.05	390.00	356.58	351.16
$VaR_{5\%}$ w/o insurance	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99	323.57	341.99
$VaR_{5\%}$ improvement	73.04	48.80	79.58	51.10	71.70	48.77	70.48	48.00	33.01	9.17
$VaR_{1\%}$	368.04	357.16	372.99	363.47	356.07	349.75	368.48	368.09	297.32	295.67
$VaR_{1\%}$ w/o insurance	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56	190.94	232.56
$VaR_{1\%}$ improvement	177.09	124.61	182.05	130.91	165.13	117.19	177.54	135.53	106.38	63.12

Online Appendix

A An example of an overfitted solution (3)

Let's consider a special case of problem (3) as an illustrating example. Let $\{(\boldsymbol{x}_j, y_j)\}_{j=1,...,n}$ be a realized sample of $(\boldsymbol{X}, Y)$. Consider the minimization problem:

$$\begin{cases} \min_{I \in \mathcal{I}} & -\frac{1}{n}\sum_{j=1}^{n} U[w - y_j + I(\boldsymbol{x}_j) - (1-\theta)\pi_e(I)] \\ \text{s.t.} & P_L \leq \pi_e(I) = \frac{\lambda}{n}\sum_{j=1}^{n} I(\boldsymbol{x}_j) \leq P_U, \end{cases} \tag{A.1}$$

where $\mathcal{I} := \{I : \mathbb{R}^p \mapsto \mathbb{R}^+ | I \text{ is measurable}\}$. For simplicity we also replace the budget constraint by:

$$P_L = P_U = P = \lambda \frac{1}{n}\sum_{j=1}^{n} I(\boldsymbol{x}_j)$$

Then we have the following proposition:

Proposition 1 (Jensen's inequality)**.** *For any concave utility function U and any deterministic function I such that $P = \lambda \frac{1}{n}\sum_{j=1}^{n} I(\boldsymbol{x}_j)$, we have:*

$$\begin{aligned} \frac{1}{n}\sum_{j=1}^{n} U[w - y_j + I(\boldsymbol{x}_j) - (1-\theta)P] &\leq U\left[\frac{1}{n}\sum_{j=1}^{n}\{w - y_j + I(\boldsymbol{x}_j) - (1-\theta)P\}\right] \\ &= U\left[w - \frac{1}{n}\sum_{j=1}^{n} y_j + \frac{P}{\lambda} - (1-\theta)P\right], \end{aligned}$$

with equality if and only if $I(\boldsymbol{x}_j) - y_j$ is a constant. Therefore the optimal solution I^* is given by,

$$I^*(\boldsymbol{x}) = \begin{cases} y_j + P/\lambda - \frac{1}{n}\sum_{j=1}^{n} y_j, & \textit{if } \boldsymbol{x} = \boldsymbol{x}_j, j = 1, 2, ..., n, \\ \textit{any arbitrary number}, & \textit{otherwise.} \end{cases} \tag{A.2}$$

This solution is "overfitted": although it mathematically optimizes problem (A.1), it says nothing about what the amount of indemnity should be for a new data sample. In fact, the problem comes from the fact that the admissible functional space $\mathcal{I}$ is too large and contains functions that are not constrained, or smooth enough. On the other hand, for instance, we constrain the space $\mathcal{I}$ to be the space of linear functions, then equations (A.2) cannot be satisfied for all indices $j = 1, ..., n$. Such a solution is too smooth, and usually result in a poor fit. Thus the challenge is to find a trade-off between these two extreme cases, that is, to propose suitable functional constraints on the $\mathcal{I}$.

B Feasible sets

We want to choose a feasible set, $\mathcal{I}_0$, so as to strike a balance between flexibility and stability. $\mathcal{I}_0$ should be large enough to include candidate payoff functions that capture intricate (nonlinear, nonmonotonic) relationships between the high-dimensional indices and losses, yet $\mathcal{I}_0$ should also exclude those very "ill-behaved" ones in $\mathcal{I}$, which are sensitive to the sample data and cannot be an appropriate insurance contract. The trade-off between contract flexibility and robustness is illustrated in Figure B.1. The blue star in $\mathcal{I}$ is the global optimal contract, which may not be obtained.[35] The red triangle illustrates a highly unstable, overfitted contract, which we want to avoid. The dotted-grey circle area, $\tilde{\mathcal{I}}_0 \subset \mathcal{I}$, is a set of piecewise linear contracts. Although quite stable, $\tilde{\mathcal{I}}_0$ is far away from the blue star due to its restrictive functional form. Our goal is to expand the boundary of the feasible set towards $\mathcal{I}_0$, and obtain the optimal contract that falls within the intersection area, which is represented by the blue diamond. This optimal contract sacrifices a little stability but achieves much more flexibility and hence a large amount of basis risk reduction.

[35]This is due to the fact that we replace the expectation in problem (1) with its empirical counterpart.

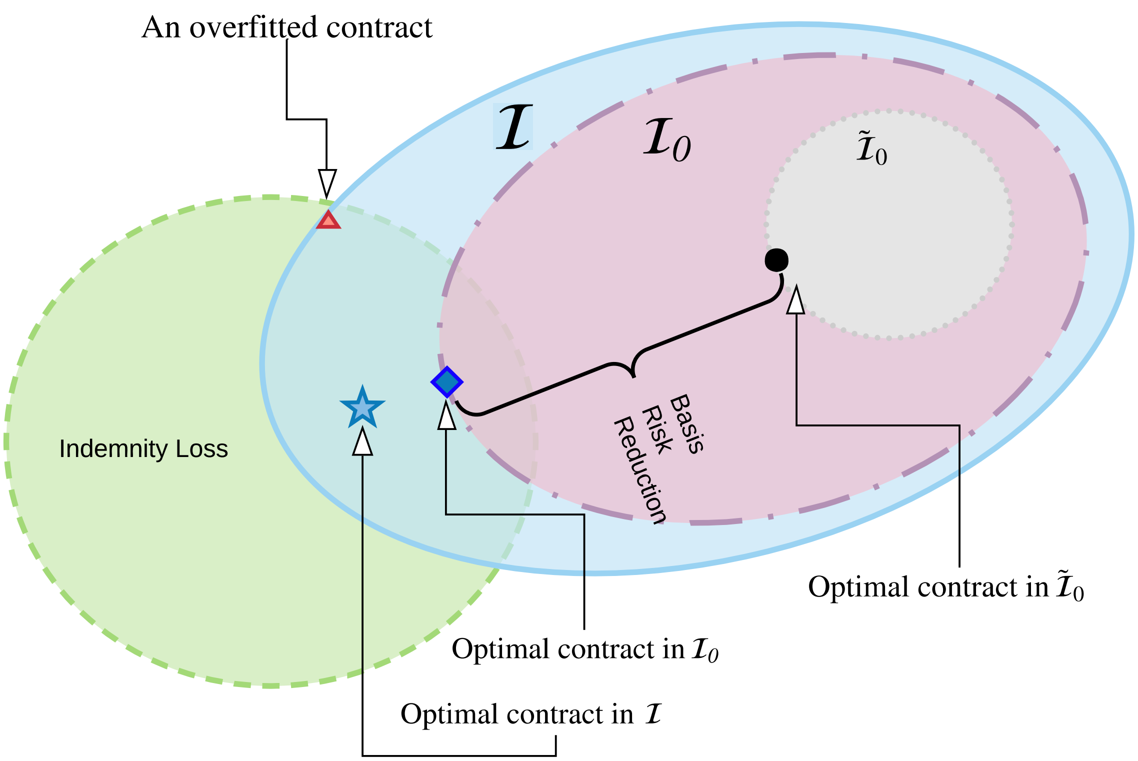
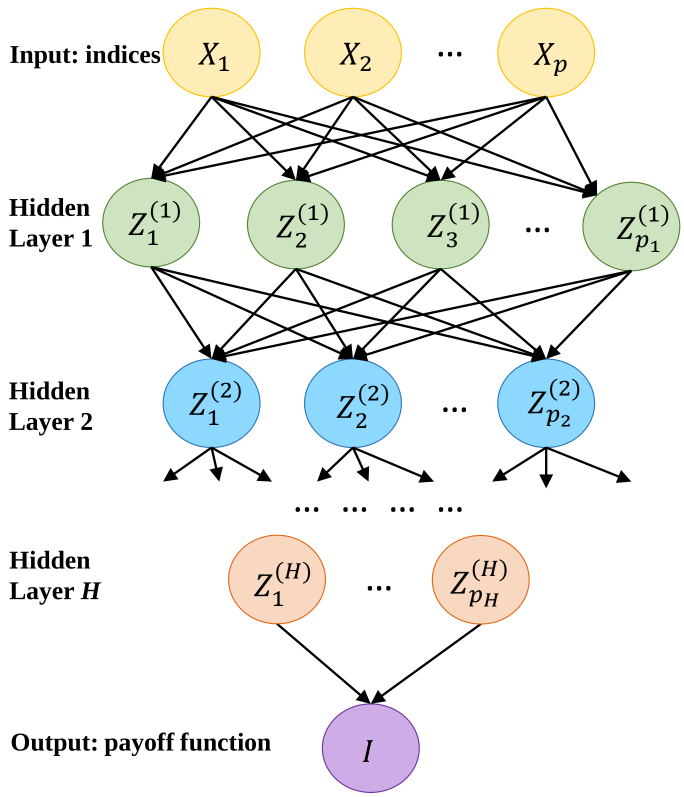


Figure B.1: **Feasible sets and optimal contracts.** This figure compares three different feasible sets and their corresponding optimal contracts. The dashed-green circle area represents the indemnity loss, which is the actual loss experienced by the policyholder. The general feasible set, $\mathcal{I}$, is represented by the solid-blue circle area and the blue star denotes the global optimal contract. The dotted-grey circle area, $\tilde{\mathcal{I}}_0$, is a feasible set of all piecewise linear contracts. The black dot at the edge of $\tilde{\mathcal{I}}_0$ is the optimal piecewise linear contract, i.e., the contract with the smallest basis risk within $\tilde{\mathcal{I}}_0$. The dotted-blue area, $\mathcal{I}_0$, represents the feasible set we explore. Its optimal contract is denoted by the blue diamond. The red triangle illustrates an overfitted contract.

C Neural network structure

Figure C.1 illustrates a neural network with H-hidden layers.



$$\boldsymbol{X}_{Input} = \boldsymbol{X}^T = (X_1, X_2, \ldots, X_p)^T$$

$$\boldsymbol{Z}^{(1)} = f_0(\boldsymbol{\alpha}^{(0)} + \boldsymbol{\omega}^{(0)} \boldsymbol{X}_{Input})$$

$$\boldsymbol{Z}^{(2)} = f_1(\boldsymbol{\alpha}^{(1)} + \boldsymbol{\omega}^{(1)} \boldsymbol{Z}^{(1)})$$

$$\boldsymbol{Z}^{(H)} = f_{H-1}(\boldsymbol{\alpha}^{(H-1)} + \boldsymbol{\omega}^{(H-1)} \boldsymbol{Z}^{(H-1)})$$

$$I = f_H(\boldsymbol{\alpha}^{(H)} + \boldsymbol{\omega}^{(H)} \boldsymbol{Z}^{(H)})$$

Figure C.1: **An illustration of neural networks with H-hidden layers.** This is an example of the fully-connected architecture in which neurons between two adjacent layers are fully pairwise connected, but neurons within a layer have no connections. f_h is an activation function; $\boldsymbol{\alpha}^{(h)}$ and $\boldsymbol{\omega}^{(h)}$ are parameters of the linear combination, $h = 1, 2, \ldots, H$.

D Data Summary

Tables D.1 and D.2 show summary statistics of the 72 weather indices used for empirical analysis. Statistics including mean, standard deviation, minimum, 25^{th} and 75^{th} percentiles, and maximum, are presented. The sample period is 1925-2018.

Table D.1: Summary statistics of weather indices. See descriptions for index variables abbreviations in Table 1.

	pcpn1	pcpn2	pcpn3	pcpn4	pcpn5	pcpn6	pcpn7	pcpn8	pcpn9	pcpn10	pcpn11	pcpn12
Mean	53.36	49.69	77.37	98.32	106.95	108.91	94.97	87.04	86.25	74.84	74.97	62.08
StD	41.73	31.35	42.99	47.51	54.25	52.62	47.61	45.58	55.35	45.60	44.94	39.53
Min	0.03	0.06	6.86	16.60	7.71	4.28	3.97	2.30	0.21	0.02	4.28	3.90
Q_{25}	26.11	27.27	45.38	63.49	67.21	71.22	61.45	55.14	47.90	42.48	42.02	35.36
Q_{75}	68.76	65.37	99.57	125.26	133.62	138.28	120.13	109.56	111.79	98.32	100.17	78.26
Max	411.75	260.57	372.58	413.99	365.64	419.15	347.60	371.80	339.07	310.42	322.15	338.16
	dpt1	dpt2	dpt3	dpt4	dpt5	dpt6	dpt7	dpt8	dpt9	dpt10	dpt11	dpt12
Mean	-7.66	-5.91	-1.57	3.64	9.58	14.97	17.61	17.02	12.60	6.00	0.06	-5.13
StD	3.50	3.36	2.72	2.31	2.43	1.81	1.47	1.51	1.87	2.20	2.28	3.18
Min	-20.75	-17.26	-10.10	-3.96	2.64	9.19	12.93	12.49	7.32	-3.59	-8.08	-16.12
Q_{25}	-9.89	-8.19	-3.52	2.03	7.89	13.76	16.55	16.01	11.33	4.70	-1.42	-7.16
Q_{75}	-5.26	-3.46	0.20	5.27	11.26	16.25	18.62	18.08	13.90	7.39	1.69	-2.89
Max	2.11	2.14	7.22	10.09	16.84	20.35	22.35	21.52	18.51	12.57	6.71	4.80
	tmax1	tmax2	tmax3	tmax4	tmax5	tmax6	tmax7	tmax8	tmax9	tmax10	tmax11	tmax12
Mean	1.73	4.24	10.37	17.75	23.59	28.58	30.71	29.74	26.27	19.68	11.11	3.88
StD	3.77	3.87	3.60	2.60	2.36	2.02	2.00	1.99	2.17	2.54	2.86	3.39
Min	-9.45	-8.17	-0.42	9.11	16.72	23.08	24.23	24.04	18.94	9.35	2.53	-7.20
Q_{25}	-0.82	1.58	7.82	15.93	21.90	27.17	29.40	28.39	24.68	18.07	9.16	1.67
Q_{75}	4.42	7.02	12.84	19.62	25.25	29.85	31.89	30.99	27.79	21.19	13.13	6.27
Max	12.51	15.14	21.65	25.03	30.58	35.78	38.13	37.82	33.48	28.53	18.97	13.41

Table D.2: Summary statistics of weather indices (cont'd). See descriptions for index variables abbreviations in Table 1.

	tmin1	tmin2	tmin3	tmin4	tmin5	tmin6	tmin7	tmin8	tmin9	tmin10	tmin11	tmin12
Mean	-7.84	-5.88	-0.77	5.22	10.89	16.11	18.22	17.10	12.86	6.56	0.45	-5.23
StD	3.81	3.71	2.92	2.22	2.17	1.74	1.56	1.75	1.89	2.02	2.30	3.40
Min	-21.01	-19.34	-12.21	-2.61	4.93	10.41	13.32	11.29	7.58	-0.48	-7.93	-17.39
Q_{25}	-10.30	-8.28	-2.73	3.68	9.33	14.94	17.15	15.95	11.53	5.27	-1.01	-7.32
Q_{75}	-5.13	-3.16	1.14	6.80	12.38	17.32	19.30	18.27	14.15	7.92	2.03	-2.89
Max	1.70	3.34	9.21	11.92	18.07	22.11	23.33	24.05	19.71	13.04	8.15	4.00
	vpdmax1	vpdmax2	vpdmax3	vpdmax4	vpdmax5	vpdmax6	vpdmax7	vpdmax8	vpdmax9	vpdmax10	vpdmax11	vpdmax12
Mean	2.86	3.84	6.78	11.81	16.26	21.03	22.87	21.18	18.68	12.94	6.45	3.27
StD	1.09	1.42	2.11	2.43	3.15	4.11	4.90	4.46	3.95	3.05	1.82	1.08
Min	0.81	1.21	2.16	6.19	8.64	13.19	12.41	12.41	9.04	4.49	2.16	0.89
Q_{25}	2.00	2.76	5.28	9.96	14.07	18.07	19.73	18.04	15.92	11.04	5.13	2.44
Q_{75}	3.52	4.71	8.07	13.48	18.17	23.17	25.11	23.57	20.97	14.43	7.58	3.96
Max	7.18	9.80	15.98	19.94	28.31	37.81	44.68	43.10	36.38	28.28	13.71	7.05
	vpdmin1	vpdmin2	vpdmin3	vpdmin4	vpdmin5	vpdmin6	vpdmin7	vpdmin8	vpdmin9	vpdmin10	vpdmin11	vpdmin12
Mean	0.45	0.52	0.78	1.33	1.59	1.87	1.61	1.16	1.10	0.96	0.72	0.48
StD	0.19	0.21	0.29	0.37	0.49	0.72	0.75	0.65	0.51	0.35	0.23	0.19
Min	0.01	0.01	0.05	0.28	0.06	0.08	0.00	0.00	0.00	0.00	0.07	0.01
Q_{25}	0.30	0.35	0.57	1.09	1.28	1.36	1.10	0.70	0.76	0.74	0.55	0.33
Q_{75}	0.60	0.68	0.97	1.56	1.84	2.25	2.01	1.50	1.35	1.16	0.88	0.63
Max	1.05	1.18	2.04	2.98	4.29	6.39	6.13	5.40	3.82	3.18	1.78	1.32

E Data splitting by preserving the temporal order

In this section, we test the robustness by splitting data with preserved temporal order. In particular, we split the sample into three disjoint time periods, with the earliest 70% data as training set, then the next 15% as validation set, and the last 15% as test set. We train NN models for each given set of hyperparameters and select the optimal model based on validation set. Then we compare the out-of-sample performances of the optimal NN-based contract and other simpler index insurance contracts in the test set. The results are summarized in Table E.1, which are similar to those reported in Table 3. For example, we find that NN-based index insurance significantly improves farmers' utility and CEW, and reduces risks.

Table E.1: **Comparing various insurance contracts with preserved temporal order.** This table summarizes performances of seven insurance contracts in the test set when splitting data by preserving the temporal order. The seven contracts include (1) a linear insurance contract with one weather index (*Linear1*); (2) a linear insurance contract with five weather indices (*Linear5*); (3) a quadratic insurance contract with five weather indices (*Quadratic5*); (4) a cubic insurance contract with five weather indices (*Cubic5*); (5) an NN-based contract with five weather indices (*NN5*); (6) a linear insurance contract with 72 weather indices (*Linear72*); and (7) optimal NN-based contract with 72 weather indices (*NN72*). Panel A summarizes expected utilities with and without (w/o) index insurance and the percentage of utility improvement. Panel B reports certainty equivalent wealth (CEW) with and without (w/o) index insurance and the CEW improvement in dollars and as a percentage. Panel C summarizes policy characteristics including policy premium, coverage, and profits of the insurer. Panel D summarizes the risk reduction effect of an index insurance policy, measured by the standard deviation of wealth. Panel E summarizes the tail risk reduction, measured by the 5%- and 1%-level value-at-risk (VaR). The risk loading parameter at equilibrium (λ^*) for each contract is reported in parentheses.

Contract	Linear1	Linear5	Quadratic5	Cubic5	NN5	Linear72	NN72
	($\lambda^* = 1.0628$)	($\lambda^* = 1.0829$)	($\lambda^* = 1.0855$)	($\lambda^* = 1.0747$)	($\lambda^* = 1.0640$)	($\lambda^* = 1.1337$)	(BL, $\lambda^* = 1.1668$)
Panel A: Utility improvement							
U with insurance	-5.53	-4.98	-4.91	-4.85	-4.81	-4.56	-4.36
U w/o insurance	-6.00	-6.00	-6.00	-6.00	-6.00	-6.00	-6.00
U improvement (%)	7.77%	17.05%	18.17%	19.10%	19.75%	23.97%	27.34%
Panel B: CEW improvement							
CEW with insurance	389.70	402.96	404.66	406.08	407.10	413.85	419.52
CEW w/o insurance	379.59	379.59	379.59	379.59	379.59	379.59	379.59
CEW improvement	10.10	23.37	25.07	26.49	27.51	34.26	39.93
CEW improvement (%)	2.66%	6.16%	6.60%	6.98%	7.25%	9.02%	10.52%
Panel C: Policy characteristics							
Premium	14.98	12.03	13.87	17.60	15.98	29.81	26.08
Coverage	14.10	11.10	12.78	16.38	14.44	26.30	22.35
Insurer Profit	0.89	0.92	1.09	1.22	1.55	3.52	3.73
Panel D: Risk reduction measured by standard deviation							
Std	91.23	84.14	82.79	81.63	81.38	73.14	68.18
Std w/o insurance	97.65	97.65	97.65	97.65	97.65	97.65	97.65
Std reduction	6.58%	13.84%	15.22%	16.41%	16.67%	25.10%	30.18%
Panel E: Risk reduction measured by Value-at-Risk (VaR)							
$\text{VaR}_{5\%}$	274.47	289.10	292.54	298.14	303.69	294.58	310.75
$\text{VaR}_{5\%}$ w/o insurance	258.65	258.65	258.65	258.65	258.65	258.65	258.65
$\text{VaR}_{5\%}$ improvement	15.82	30.45	33.89	39.49	45.04	35.93	52.10
$\text{VaR}_{1\%}$	78.96	131.71	132.92	130.68	134.93	196.49	225.10
$\text{VaR}_{1\%}$ w/o insurance	58.19	58.19	58.19	58.19	58.19	58.19	58.19
$\text{VaR}_{1\%}$ improvement	20.77	73.52	74.73	72.49	76.74	138.31	166.92

F Correlations between production losses and weather indices

Figure F.1 shows the correlation coefficients between production losses and 72 weather indices.

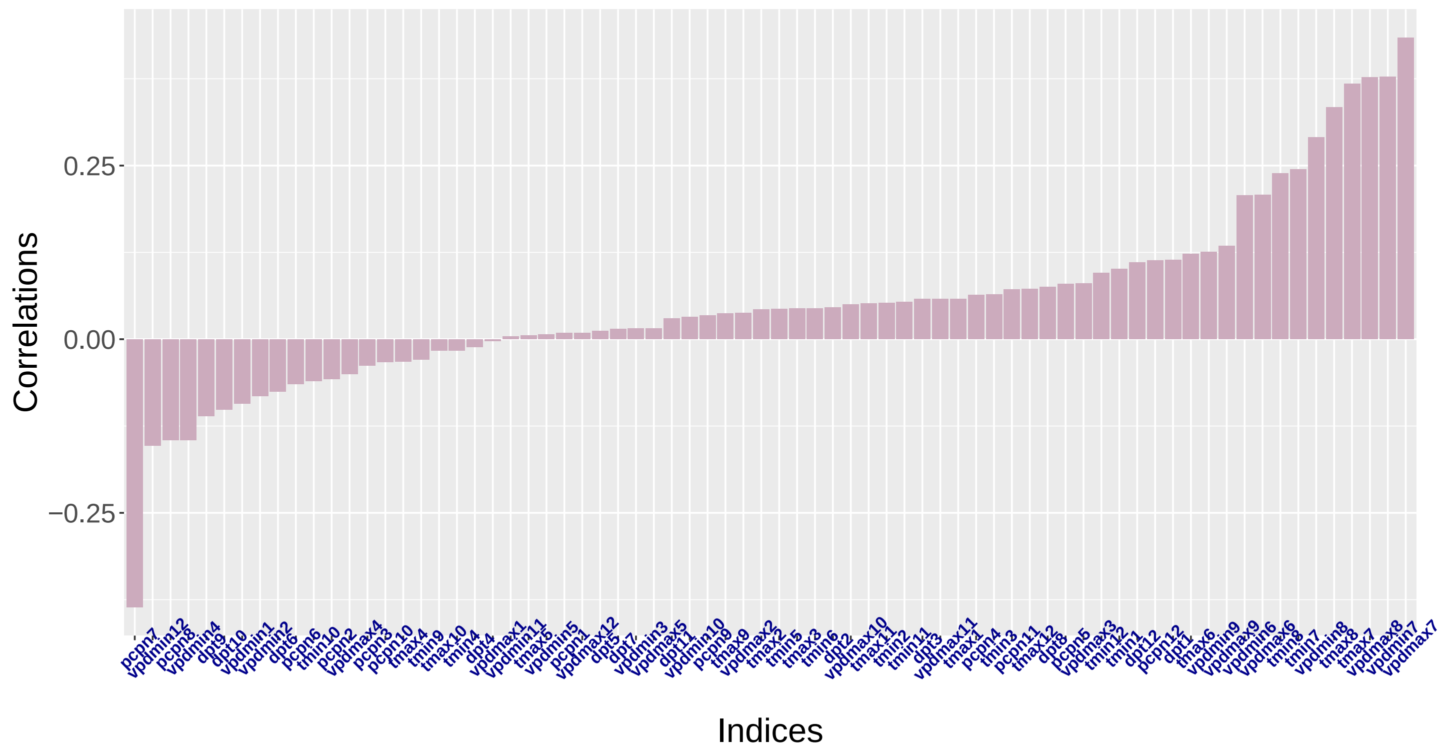
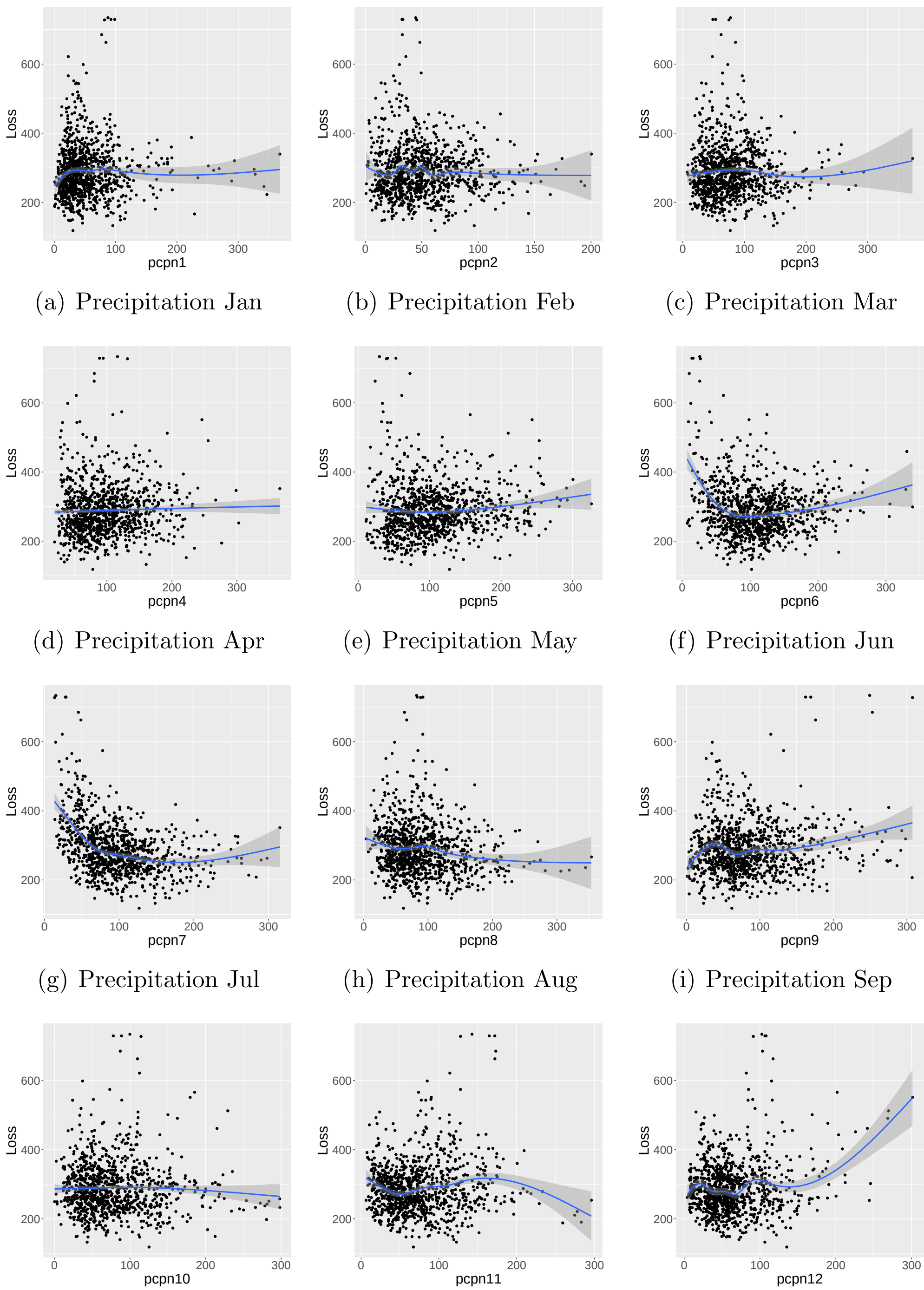


Figure F.1: **Correlations between production losses and weather indices.**

G Nonlinear relationships between production losses and weather indices

This appendix collects scatterplots of all 72 weather indices against the crop losses, using 1,000 random draws from the sample. The blue curve is fitted by a generalized additive model. The shadow area indicates 95% confidence interval. We can see that most weather indices have intricate nonlinear relationships with crop losses, and this complexity could not be adequately captured by linear models that are used by most existing index insurance design framework. The nonlinearity suggests inadequacy of those index insurance with simple structures, and calls for more sophisticated models.



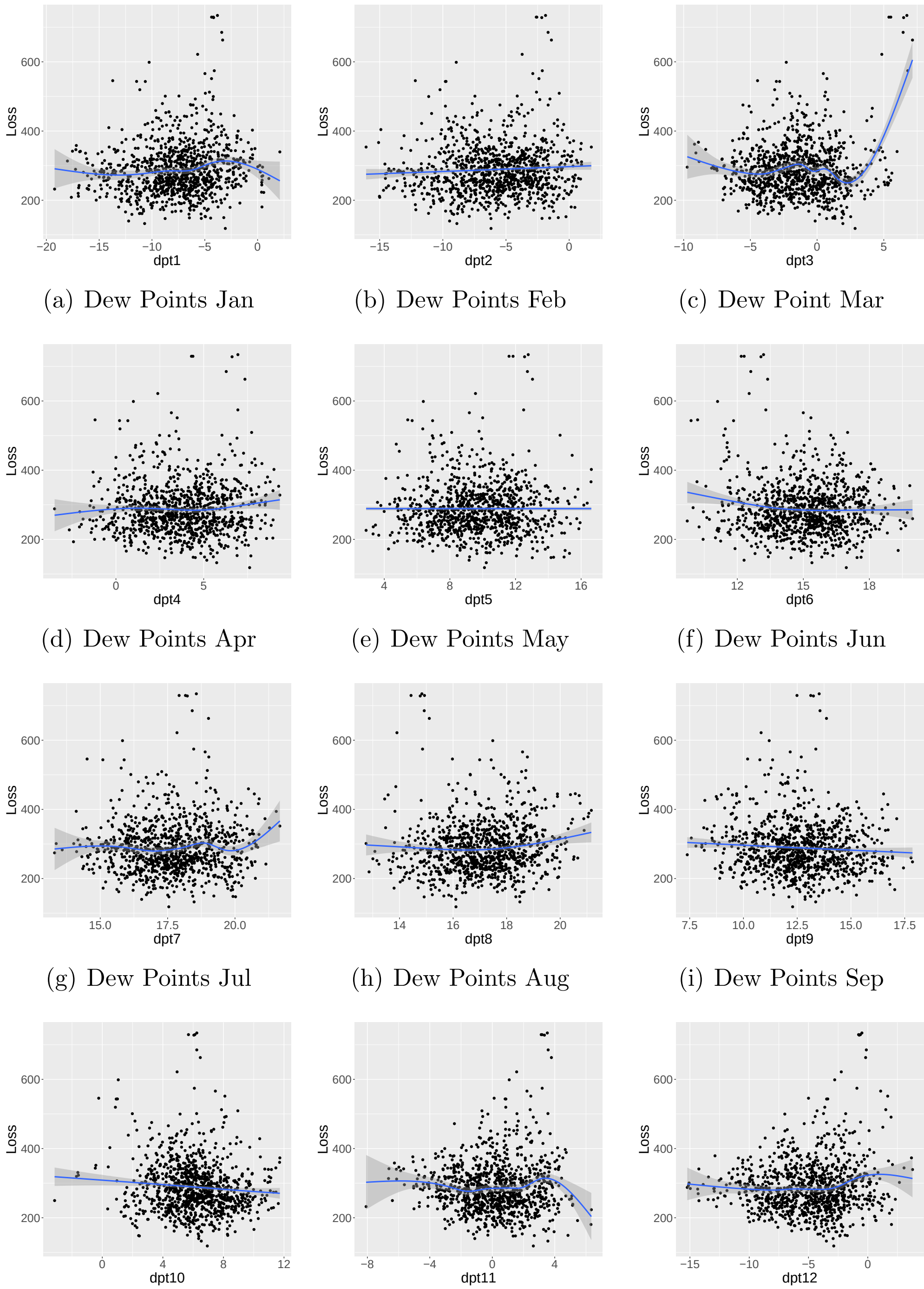
(a) Precipitation Jan (b) Precipitation Feb (c) Precipitation Mar

(d) Precipitation Apr (e) Precipitation May (f) Precipitation Jun

(g) Precipitation Jul (h) Precipitation Aug (i) Precipitation Sep

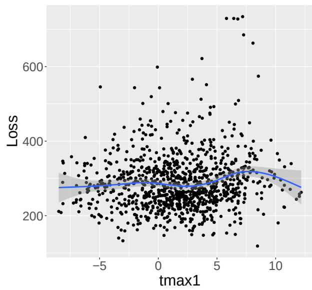
(j) Precipitation Oct (k) Precipitation Nov (l) Precipitation Dec

Figure G.1: Scatterplots of precipitation (Jan-Dec) with crop losses.

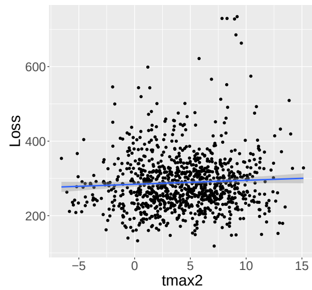


(a) Dew Points Jan
(b) Dew Points Feb
(c) Dew Point Mar
(d) Dew Points Apr
(e) Dew Points May
(f) Dew Points Jun
(g) Dew Points Jul
(h) Dew Points Aug
(i) Dew Points Sep
(j) Dew Points Oct
(k) Dew Points Nov
(l) Dew Points Dec

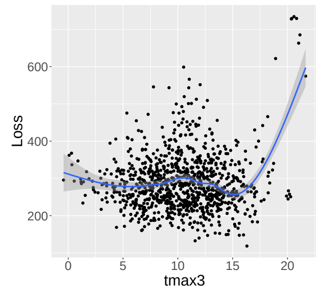
Figure G.2: Scatterplots of dew point temperature (Jan-Dec) with crop losses.



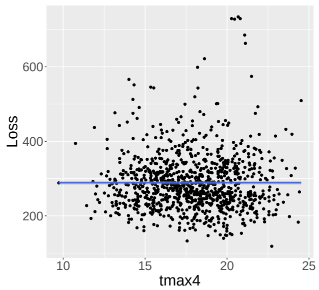
(a) Max Temperature Jan



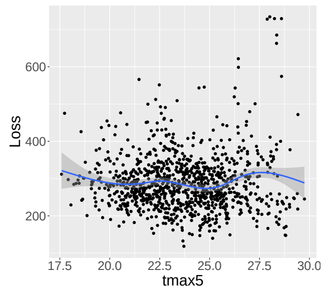
(b) Max Temperature Feb



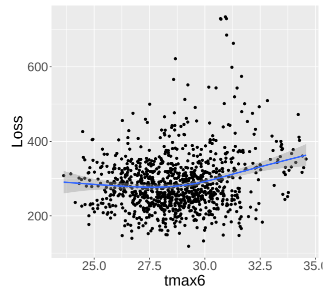
(c) Max Temperature Mar



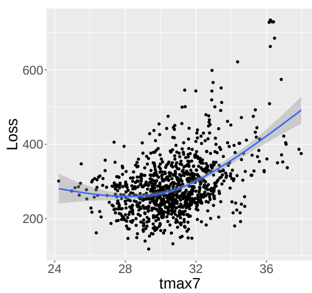
(d) Max Temperature Apr



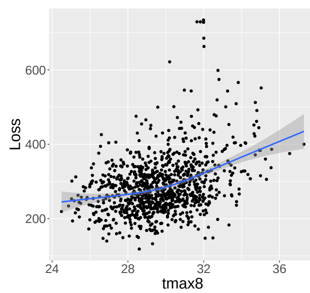
(e) Max Temperature May



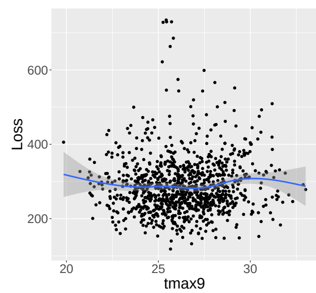
(f) Max Temperature Jun



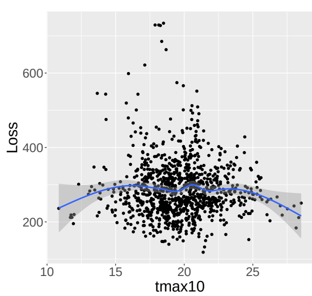
(g) Max Temperature Jul



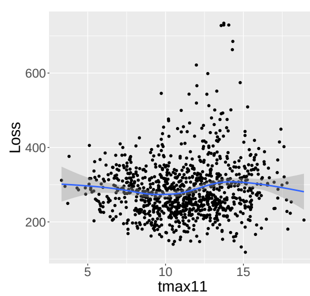
(h) Max Temperature Aug



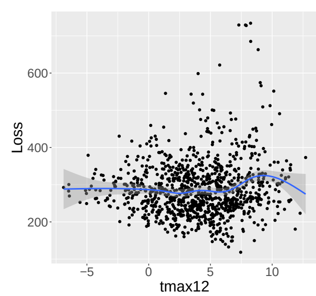
(i) Max Temperature Sep



(j) Max Temperature Oct

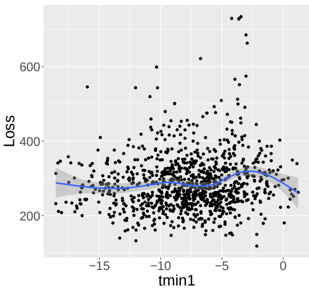


(k) Max Temperature Nov

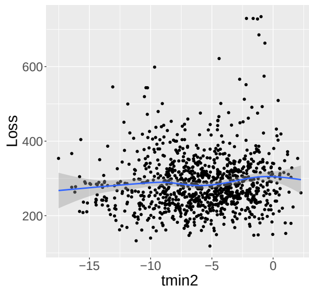


(l) Max Temperature Dec

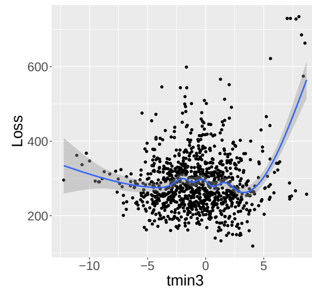
Figure G.3: Scatterplots of maximum temperature (Jan-Dec) with crop losses.



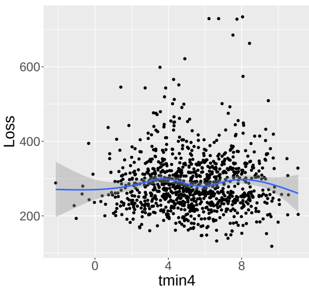
(a) Min Temperature Jan



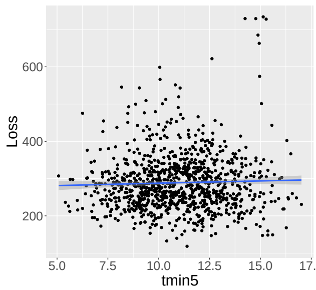
(b) Min Temperature Feb



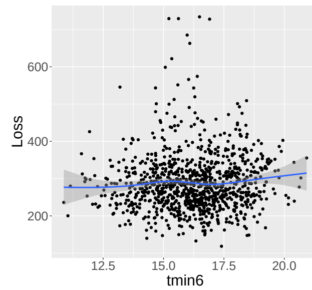
(c) Min Temperature Mar



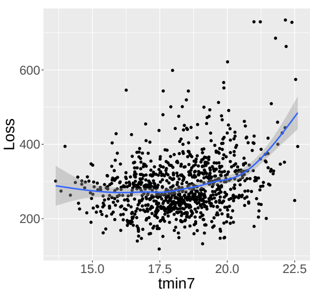
(d) Min Temperature Apr



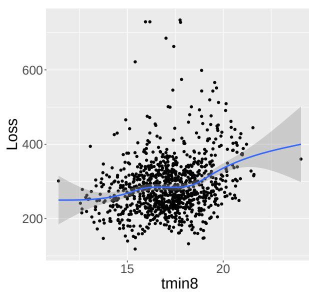
(e) Min Temperature May



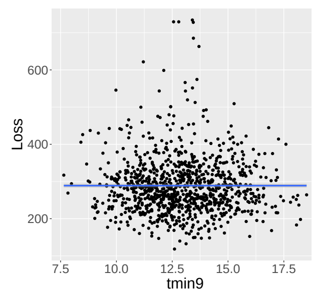
(f) Min Temperature Jun



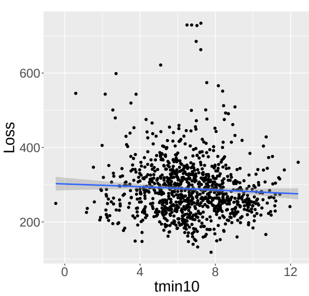
(g) Min Temperature Jul



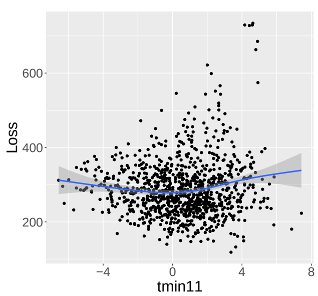
(h) Min Temperature Aug



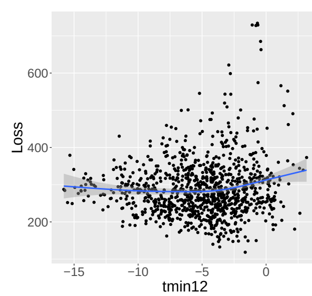
(i) Min Temperature Sep



(j) Min Temperature Oct

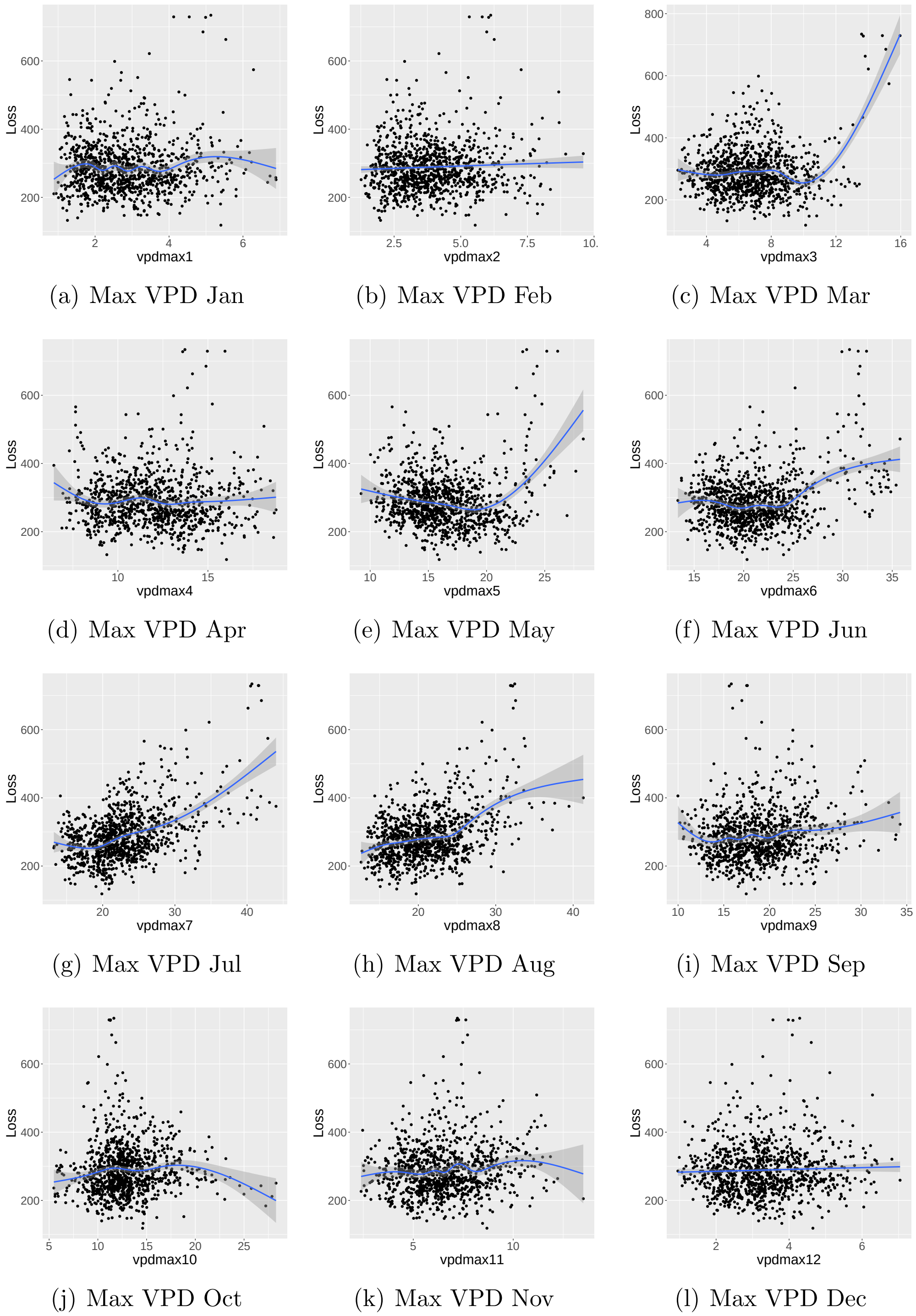


(k) Min Temperature Nov



(l) Min Temperature Dec

Figure G.4: Scatterplots of minimum temperature (Jan-Dec) with crop losses.



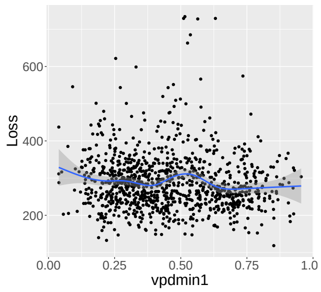
(a) Max VPD Jan (b) Max VPD Feb (c) Max VPD Mar

(d) Max VPD Apr (e) Max VPD May (f) Max VPD Jun

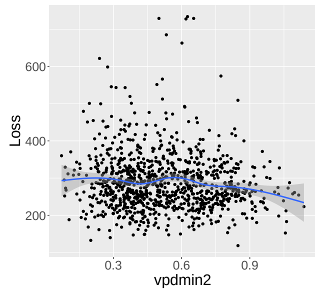
(g) Max VPD Jul (h) Max VPD Aug (i) Max VPD Sep

(j) Max VPD Oct (k) Max VPD Nov (l) Max VPD Dec

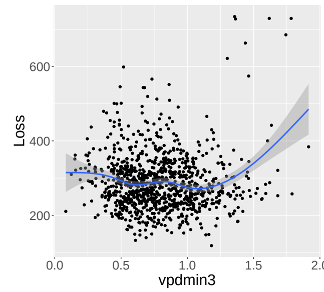
Figure G.5: Scatterplots of maximum vapor pressure deficit (Jan-Dec) with crop losses.



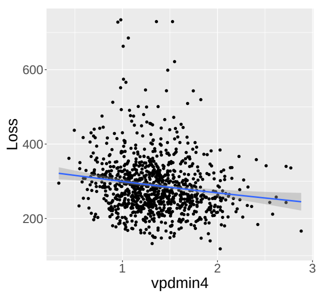
(a) Min VPD Jan



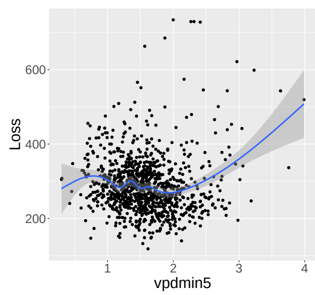
(b) Min VPD Feb



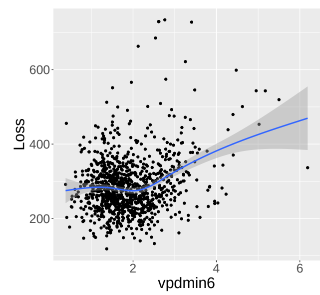
(c) Min VPD Mar



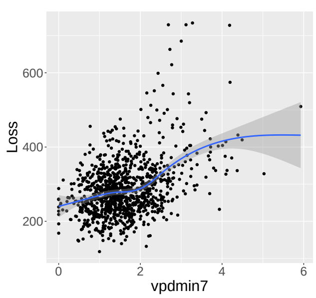
(d) Min VPD Apr



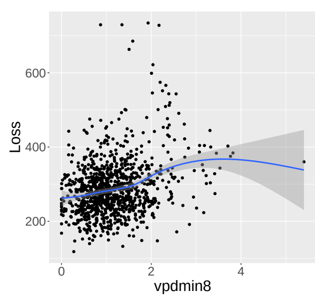
(e) Min VPD May



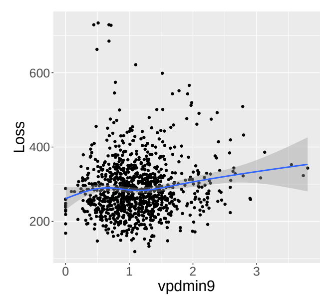
(f) Min VPD Jun



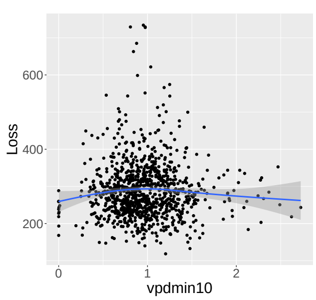
(g) Min VPD Jul



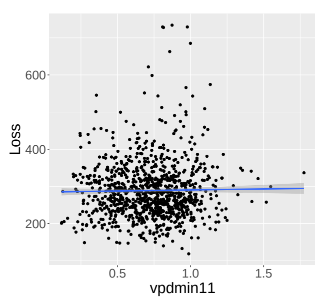
(h) Min VPD Aug



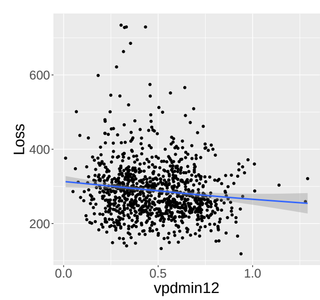
(i) Min VPD Sep



(j) Min VPD Oct



(k) Min VPD Nov



(l) Min VPD Dec

Figure G.6: Scatterplots of minimum vapor pressure deficit (Jan-Dec) with crop losses.

H Additional gradient-based sensitivity analysis

H.1 Ranking weather indices according to gradient-based sensitivities

This appendix displays ranking of all weather indices based on their gradient-based sensitivities to insurance payoffs from the NN-based index insurance contract and their absolute correlations with production losses, in Figure H.1. The number on top of each bar is the rank difference between using the sensitivity analysis and the absolute correlation (see Table F.1). We can see from Figure H.1 that some weather indices are impactful in terms of both absolute correlation and sensitivities (those ranked high with small rank differences, e.g., vpdmax7, pcpn7), whereas some weather indices are impactful based on sensitivities but not correlations (those ranked high with larger rank differences, e.g., tmin2, tmax10). From the perspective of designing effective index insurance contracts, those weather indices with large absolute value of correlations are not necessarily the most important ones.

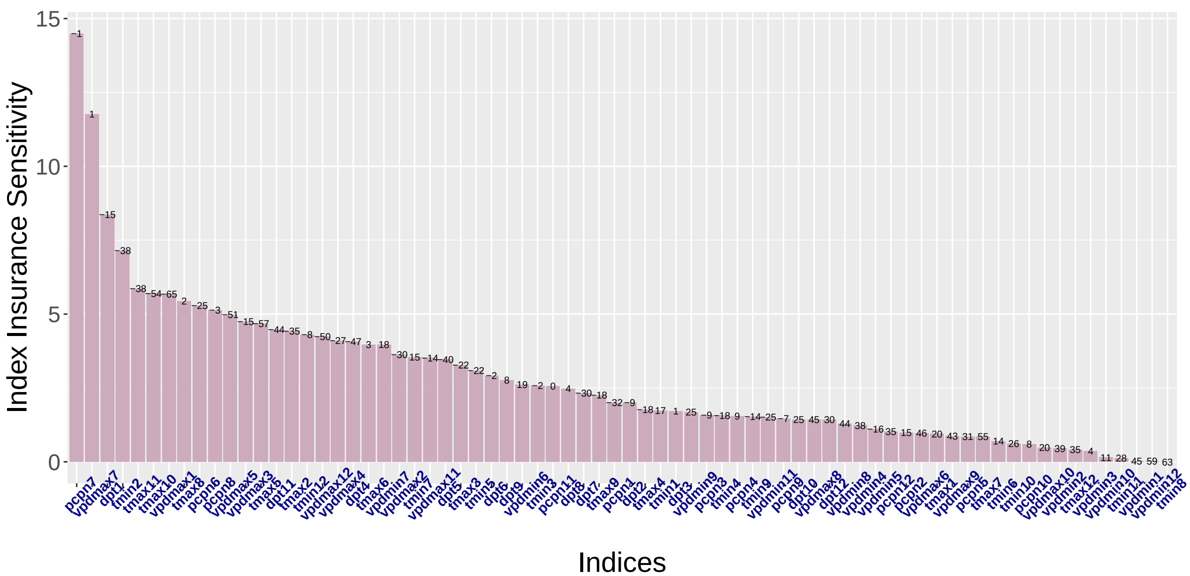


Figure H.1: **Rankings of indices according to index insurance sensitivities.** The number on top of each bar is the rank difference between using the sensitivity analysis criterion and the absolute correlation criterion.

H.2 The impacts of dimensionality

In this section, we investigate the impact of dimensionality on the NN-based index insurance performance. We consider models with the most important 1, 18, 36, 54, and 72 weather indices. The index importance is ranked based on the gradient-based sensitivity analysis discussed in Section 4.3. We see that using only one index improves the utility by 3.11% in the test set. Adding more weather indices significantly improves the model performances. For example, the model with 36 weather indices improves the utility by 11.66% in the test set. This analysis demonstrates the importance of including higher dimensional inputs in the NN-based index insurance contract.

Table H.1: **Comparing models with various number of weather indices.** We evaluate the performance of the NN-based index insurance with different number of weather indices in the test set, using 1, 18, 36, 54, and 72 weather indices. Panel A summarizes utilities with and without (w/o) different index insurance policies and the percentage of utility improvement. Panel B summarizes CEW with and without (w/o) index insurance policies and certainty equivalent wealth (CEW) improvements in dollars and as a percentage. Panel C summarizes policy characteristics, including premiums, coverage, and profits for the insurers. Panel D summarizes the risk reduction effect of different index insurance policies, measured by the standard deviation of wealth. Panel E summarizes the risk reduction at the tail, measured by the 5%- and 1%-level value-at-risk (VaR). "BL" represents the baseline case studied in Section 4.2.

Contract	72 indices (BL)	54 indices	36 indices	18 indices	One index
Panel A: Utility improvement					
U with insurance	-3.47	-3.47	-3.49	-3.52	-3.83
U w/o insurance	-3.95	-3.95	-3.95	-3.95	-3.95
U improvement (%)	12.33%	12.17%	11.66%	11.06%	3.11%
Panel B: CEW improvement					
CEW with insurance	448.18	447.95	447.22	446.38	435.68
CEW w/o insurance	431.73	431.73	431.73	431.73	431.73
CEW improvement	16.45	16.22	15.49	14.66	3.96
CEW improvement (%)	3.81%	3.76%	3.59%	3.39%	0.92%
Panel C: Policy characteristics					
Premium	28.82	23.67	20.36	23.10	22.10
Coverage	24.70	20.28	17.45	19.80	20.80
Insurer Profit	4.12	3.38	2.91	3.30	1.31
Panel D: Risk reduction measured by standard deviation					
Std	42.60	45.28	47.97	48.48	64.96
Std w/o insurance	70.78	70.78	70.78	70.78	70.78
Std reduction	39.80%	36.03%	32.22%	31.50%	8.21%
Panel E: Risk reduction measured by Value-at-Risk (VaR)					
$\text{VaR}_{5\%}$	391.53	387.98	383.83	381.73	349.60
$\text{VaR}_{5\%}$ w/o insurance	341.99	341.99	341.99	341.99	341.99
$\text{VaR}_{5\%}$ improvement	49.54	45.99	41.84	39.74	7.61
$\text{VaR}_{1\%}$	370.54	351.80	349.06	333.10	268.78
$\text{VaR}_{1\%}$ w/o insurance	232.56	232.56	232.56	232.56	232.56
$\text{VaR}_{1\%}$ improvement	137.99	119.25	116.50	100.55	36.23