Field Rotation Related Soil Productivity Changes, Yield Risk, and Efficiency in Crop Insurance Rating

Mitchell Roznik

Post-Doctoral Fellow
Morrison School of Agribusiness,
W. P. Carey School of Business
Arizona State University
Mesa, AZ
Email: Mitchell.Roznik@asu.edu

Ashok K. Mishra

Kemper and Ethel Marley Chair
Morrison School of Agribusiness,
W. P. Carey School of Business
Arizona State University,
Mesa, AZ
Email: Ashok.K.Mishra@asu.edu

Hongqiang Yan

Post-Doctoral Fellow
Morrison School of Agribusiness,
W. P. Carey School of Business
Arizona State University
Mesa, AZ
Email: hongqiang.yan@asu.edu

Acknowledgement: We thank the seminar and conference participants at the 2022 AAEA meeting in Anaheim, CA, for their suggestions and insights into agricultural policy and productivity.

We remain indebted to USDA-NIFA for funding this project. National Institute of Food and Agriculture, U.S. Department of Agriculture, Grant/Award Number: 2021-67023-33818.

Field Rotation Related Soil Productivity Changes, Yield Risk, and Efficiency in Crop Insurance Rating

Abstract

This study estimates the marginal effect of field rotation–induced changes in soil productivity, measured by the National Commodity Crop Productivity Index (NCCPI), on crop yield distributions and crop insurance premium rates at the farm level. Using a balanced panel of 11,500 corn and 8,700 soybean farm-year observations from geolocated field-level data (2011–2020), we quantify the variation in soil productivity resulting from field rotation selection. Residual-based moment estimation reveals that NCCPI has a significant impact on both the mean and variance of yields. Scenario analysis indicates that a 4.08% increase in corn soil productivity due to rotation reduces premiums by 4.99%, while soybeans exhibit even greater sensitivity to premiums. A reinsurance simulation demonstrates that insurers using field-level NCCPI data achieve lower loss ratios than those relying on uniform ratings. These findings highlight the economic significance of incorporating dynamic soil productivity data into crop insurance underwriting. Accounting for field rotation–related soil variability can improve premium accuracy and enhance actuarial efficiency in risk-based insurance design.

Keywords*:* Weather Data, Google Earth Engine, gSSURGO, Crop Insurance, Risk

JEL Codes : Q22, Q14, Q18, Q54

1. Introduction

Crop insurance has become a central pillar of agricultural policy worldwide, particularly in the face of growing climate uncertainty and increasingly volatile commodity markets. Government-backed programs now cover vast areas of cropland, transferring billions in risk and compensation to producers annually. For instance, the U.S. Federal Crop Insurance Program (FCIP) alone provides over $100 billion in coverage, regularly exceeding traditional farm support programs in fiscal size (Glauber, 2013). Global markets for agricultural insurance are estimated to exceed $30 billion in premiums (Vyas et al., 2021; Wang et al., 2020). This expansion reflects the mounting importance of insurance in mitigating the economic consequences of adverse production shocks. Yet, despite its scale, crop insurance faces persistent efficiency challenges—most notably, adverse selection and moral hazard (Just, Calvin, and Quiggin, 1999; Wu, Goodwin, and Coble, 2020). A major policy and research priority is identifying information that can improve premium

accuracy and insurance performance under conditions of uncertainty, particularly as climate change accelerates production risk (Tack, Coble, and Barnett, 2018).

While prior research has linked numerous environmental features—such as soil texture, pH, bulk density, organic matter, nutrients, and slope—to spatial variation in agricultural productivity (Campbell et al., 1993; Cox et al., 2003; Martín, Bollero, and Bullock, 2005; Kravchenko and Bullock, 2000), less is known about their role in shaping yield risk and insurance outcomes over time at the farm level. Crop productivity indices, such as the National Commodity Crop Productivity Index (NCCPI), provide a standardized method for aggregating these features and are widely used in land valuation (Albers, Sinclair, and Robotham, 2022); however, they are largely overlooked in crop insurance pricing. Only a small number of studies (Woodard, 2014; Woodard and Verteramo-Chiu, 2017; Tsiboe and Tack, 2022) have examined the implications of such indices for temporal yield variability or premium rates at insurable units. Some recent work has also explored how incorporating weather or systemic risk data can improve insurance design (Gong et al., 2023; Liu and Ramsey, 2023). However, few studies have combined dynamic, geolocated soil data with field-level yield observations to examine the year-to-year effects on crop insurance.

The limited empirical evidence presents an important gap. While Tsiboe and Tack (2022) show that soil and topographic data can improve premium accuracy in the absence of yield history, their findings suggest minimal effects once historical yields are established. By contrast, Woodard and Verteramo-Chiu (2017) find more consistent impacts of soil characteristics on rating outcomes. Notably, the methodological choices differ: Tsiboe and Tack (2022) assume constant soil quality over time by applying a buffer-based soil data extraction method, potentially overlooking yield-relevant variation due to crop rotation and field switching. This omission may

mute the observed relationship between soil features and risk. Conversely, Woodard and Verteramo-Chiu (2017) employ methods likely to capture intra-farm soil variability more accurately. These discrepancies raise a key question: Does soil-related crop productivity influence insurance-relevant yield distributions in real-world farm settings where producers rotate fields and vary land use over time? Clarifying this relationship is critical for determining whether soil data can meaningfully enhance premium rating, particularly in high-resolution, field-level contexts.

This study assesses the impact of temporal changes in soil-related crop productivity, resulting from field selection and rotation, on crop yield risk and crop insurance premiums. We utilize a unique, private dataset comprising 11,500 corn and 8,700 soybean field-year observations across 13 U.S. Corn and Soybean Belt states, matched with geolocated field geometries, weather data, irrigation status, and soil characteristics from the gSSURGO dataset. Our empirical strategy departs from previous studies by estimating yield distributions conditional on annually varying NCCPI values, rather than assuming constant soil characteristics. Using Antle's (1983) parametric residual-based moments estimation framework, we estimate how NCCPI affects the mean and variance of yields and derive premium rates based on these distributions. We then simulate a realistic reinsurance scenario in which private insurers use farm-level soil data to set premiums and make retention decisions. At the same time, a pseudo-regulator applies uniform rates without access to field-level soil information. This design enables us to assess the economic and statistical significance of soil productivity in influencing rating outcomes.

The contribution of this study is twofold. First, we introduce a high-resolution, farm-level dataset that enables precise spatial matching of field-level yields with time-varying soil productivity measures. This allows us to capture soil heterogeneity across fields planted by the same farm across years—an aspect missing from earlier studies. Second, we quantify the implications of temporal soil variability for crop insurance performance, showing how real-world field selection decisions (e.g., due to crop rotation) affect the accuracy of premium ratings and the loss-cost ratio. Our scenario analysis complements previous efforts by offering a more dynamic and empirically grounded framework for evaluating the role of soil productivity in risk-based insurance design.

The results show that temporal variation in NCCPI—driven by field switching and crop rotation—substantially affects both the mean and variance of yield distributions. For corn, a 4.08% year-to-year change in NCCPI results in a 4.99% change in premiums under 85% coverage. For soybeans, a 5.51% change in NCCPI results in an 11% change in premiums. These effects suggest that ignoring temporal soil variability can bias premium rates, particularly for farms that frequently change field selection. Overall, our findings indicate that incorporating dynamic, geolocated soil data into rating models can improve both the efficiency and equity of crop insurance programs.

2. Background

2.1. FCIP and the Importance of Accurate Premium Pricing

The United States Department of Agriculture's Risk Management Agency (RMA) administers the FCIP. It provides insurance to farmers to help them manage risks associated with weather, insects, disease, and price fluctuations. The program enables farmers to purchase insurance policies that compensate them when crops are damaged or lost due to natural disasters or market

fluctuations. Premiums for insurance policies are subsidized by the federal government but are based on actuarially fair rating principles. The unsubsidized premium prices should accurately reflect the expected payouts over time. Accurate premium pricing is important for the financial integrity and effectiveness of the crop insurance program.

Crop insurance faces the challenges of adverse selection and moral hazard, and improving policy design may help mitigate these risks. Moral hazard refers to the tendency for insured farmers to take more risks than those without insurance, since they know claims will cover their losses. For example, a farmer with crop insurance may plant on marginal land or use riskier farming practices than they would otherwise. This could result in increased insurance claims and higher costs for the program. Adverse selection refers to the situation in which farmers have a better assessment of their risk than the underwriter and thus advantageously purchase crop insurance (Glauber 2004). Lower-risk farmers may decide they do not need insurance, while higher-risk farmers may opt for the maximum coverage. This can skew the insurance pool toward more claims and losses, making it difficult to set premium pricing. Both moral hazard and adverse selection present challenges for the FCIP; however, these effects are somewhat mitigated due to premium subsidization (Coble et al., 2013).

A large body of research has examined and critiqued the FCIP program, and many of these suggestions have been incorporated into design improvements to enhance program efficiency (Woodard and Verteramo-Chiu, 2017; Rejesus et al., 2015; Finger, 2013; Woodard et al., 2011; Skees and Reed, 1986). Future efforts to improve FCIP efficiency are important as crop insurance rating deficiencies may impact land-use and crop rotation decisions (Woodard, 2012; Chavas and Holt, 1996; Young et al., 2001; Lubowski et al., 2006). A potential area of

improvement for crop insurance premium rating is the inclusion of geolocated soil characteristics data.

2.2. Soil Characteristics and Crop Insurance Premium Pricing

Woodard (2014) investigated the effect of a soil-related crop productivity index on crop insurance premium pricing using a multi-decade field-level dataset of Illinois corn producers. Subsequently, using the same data, Woodard and Verteramo-Chiu (2017) found that including soil productivity information, measured by the NCCPI —a nationally available index, significantly improved crop insurance premium pricing through improved liability setting. Later, the same year, Li et al. (2016) also investigated the potential of enhancing county-level base rate setting using NCCPI and found promising results. The above studies suggest two potential areas for design changes to improve crop insurance premium ratings. First, refine the setting of the expected yield and liability for the insurable unit (farm/field). Second, county-level base rate setting is improved using aggregate soil productivity data. Section 2.3 provides more information on NCCPI.

The USDA-RMA operates and administers the FCIP. There are several types of crop insurance policies, but revenue and yield protection policies are the most popular. The methodologies used to price premiums vary by policy, but certain core components are common. Focusing on the yield protection policies, the premium pricing methodology can be simplified for analysis. Following the summary outlined by Tsiboe and Tack (2022), the simplified RMA rating methodology is as follows.

$$\tilde{y}_{ig} = \ddot{y}_i * C_g \quad (1)$$

$$I_{ig} = max\{0, \tilde{y}_{ig} - y_{it}\} \quad (2)$$

$$P_{ig} = E[I_{ig}] \quad (3)$$

$$P_{ig} = \boldsymbol{R_{ig}} * \tilde{y}_{ig} \quad (4)$$

$$finalPrice = P_{ig} * S_g \quad (5)$$

From Equations (1) to (5), the main components include the approved yield ($\ddot{y}_i$), yield guarantee ($\tilde{y}_{ig}$), coverage level (C_g), indemnity (I_{ig}), premium rate (P_{ig}), and subsidy (S_g). Assuming the output price is equal to one, y_{it} is the yield outcome at the insurable unit in year t. The premium pricing methodology intends to be actuarially fair, whereby the premium rate (P_{ig}) is set equal to the expected indemnities $E[I_{ig}]$. The indemnities, I_{ig} are not known at the time of underwriting, and the price, P_{ig} is set using the continuous rating formula (Equation 6) multiplied by the coverage level, $\tilde{y}_{ig}$.

$$R_{ig} = \alpha_{cg}\left[\overline{\boldsymbol{y}}_{\boldsymbol{\iota}}/\overline{y}_{cr}\right]^{\boldsymbol{\beta_c}} + \delta_{cg} \quad (6)$$

Inspecting the continuous rating formula, Equation 6, $\bar{y}_i$ is the average yield of the insurable unit that is determined by taking an average of the last 4-10 years of yield history and applying adjustments such as yield exclusion, trend, and yield substitution. β_c is the rating exponent and is negative, which implies that farms with higher average yields are less risky, α_{cg} is the county base rate, and δ_{cg} is the county loading factor. Drawing from the simplified premium pricing methodology, Woodard and Verteramo-Chiu (2017) identified mispricing due to the setting of, $\bar{y}_\iota$, the farm (insurable unit) yield, and suggested using soil productivity data could more accurately set the value for the farm yield. Results from Li et al. (2016) showed that the county base rate α_{cg} could be set more accurately by conditioning on soil productivity.

In 2021, Tsiboe and Tack investigated conditioning the weighting exponent, $\boldsymbol{\beta_c}$ (Equation 6) on soil and topographic factors to adjust county-based premiums more accurately to farm insurable units. The authors generated premium rates for Kansas farms using the simplified

RMA premium pricing methodology (Equation 6) and an adjusted methodology that conditioned the weighting parameter on several soil topographic features. Results from the study suggested economically significant improvements for farms with few historical yield records, but no improvement for farms with an established yield history. A limitation of Tsiboe and Tack's (2021) study was that, due to data limitations, the soil factors were not precisely geolocated at the field locations. Rather, the soil factors were extracted from farm buffer areas. Indeed, using soil factors from buffer areas may reduce the effect of soil-related crop productivity on crop insurance premiums, as this measure holds soil-related crop productivity constant over time. In other words, Tsiboe and Tack's (2022) study did not capture the temporal changes in soil-related crop productivity due to field mixing and field selection.

In sum, Woodard and Verteramo-Chiu (2017), Li et al. (2016), and Tsiboe and Tack (2022) investigate the use of NCCPI and other soil characteristics such as soil organic matter, soil texture, and root zone depth, available water storage, slope, exchangeable cations, and soil organic carbon. The current study focuses on the NCCPI index because it is nationally available and links soil characteristics to crop productivity using a biological model informed by agronomists, plant science field specialists, and soil scientists (Albers, Sinclair, and Robotham, 2022).

2.3. National Commodity Crop Productivity Index (NCCPI)

The NCCPI is a national soil productivity interpretation system that uses fuzzy logic modeling to combine the effects of soil, landscape, and climate factors on commodity crop yields. Corn, Soybeans, Small Grains, and Cotton are the four main crop submodels. Each submodel comprises 5-6 subrules that quantify the impacts of properties such as soil chemistry, available water, climate, and topography. The subrules utilize fuzzy logic relationships based on

continuous data or crisp thresholds for categorical data, such as erosion. The fuzzy subrule scores are multiplied, and crisps are added/subtracted. Differences in crop physiology and geography result in variations between the sub-models in terms of the weighting of climate, soil chemistry, and physical property impacts on productivity. The sub-model relationships were derived from yield data analysis, literature review, and boundary line analysis of bivariate plots of soil/climate factors relative to crop yields. The NCCPI methodology uses the following soil and topographic variables: cation exchange capacity, soil pH, soil organic matter, salinity, available water holding capacity, depth to root restricting layers, saturated hydraulic conductivity, linear extensibility percent, coarse fragments/rock volume, soil erosion, surface degradation, parent material, among others. The NCCPI combines a quantitative understanding of how different soils and climates affect crop productivity into a straightforward yet powerful soil productivity index system. The methodology is primarily biological rather than statistical (Albers, Sinclair, and Robotham, 2022). The NCCPI uses a scale ranging from 0.01 to 1.0, with 1.0 representing the highest productivity for non-irrigated commodity crops. The NCCPI model focuses on a relative productivity ranking over periods of years, not for a single year. The NCCPI index evaluates the suitability of soil for various crops and is used by USDA conservation and Farm Bill programs, as well as by other federal agencies and agricultural stakeholders.

2.4. Field Selection and Temporal NCCPI Changes

When an insurance unit comprises multiple fields with differing soils, the average soil-related crop productivity can vary from year to year. Farms may experience varying soil productivity in different years because they select different fields for production. It is worth noting that the soil productivity of various fields differs significantly due to several factors, including land use history, drainage, organic matter, nutrient availability, slope, and terrain, among others. Further,

variations in soil productivity can have significant implications for farming practices, land values, and environmental sustainability. Farmers build organic matter in their fields by adding organic materials, such as compost, manure, and cover crops. Farmers practice crop rotation and mulching, as well as conservation tillage, on their fields to reduce tillage and increase the organic matter content (Machmuller et al. 2015).

The mixing fields could increase yield variability and risk over time. For example, a corn insurance unit may contain three fields – one with highly productive soil, one with moderately productive soil, and one with poorly drained unproductive soil. As these fields are rotated into corn production each year in different proportions, the overall unit's average yield and yield risk may fluctuate more than if the unit contained homogeneous soils. In some years, a higher proportion of poorer ground may be planted, which lowers the overall insured unit yield. The field or plot mixing effect may make it more difficult to estimate average yield and yield risk reliably compared to an insurance unit with more consistent soils. For instance, Figure 1 illustrates an example of an insurable unit (corn) over a four-year period. The soil-related productivity, approximated by NCCPI, was 0.464 in 2017, 0.440 in 2018, 0.702 in 2019, and 0.691 in 2020. A higher NCCPI score indicates higher soil-related productivity. In this example, the variability of the annual NCCPI score for the insurable unit is driven by crop rotation.

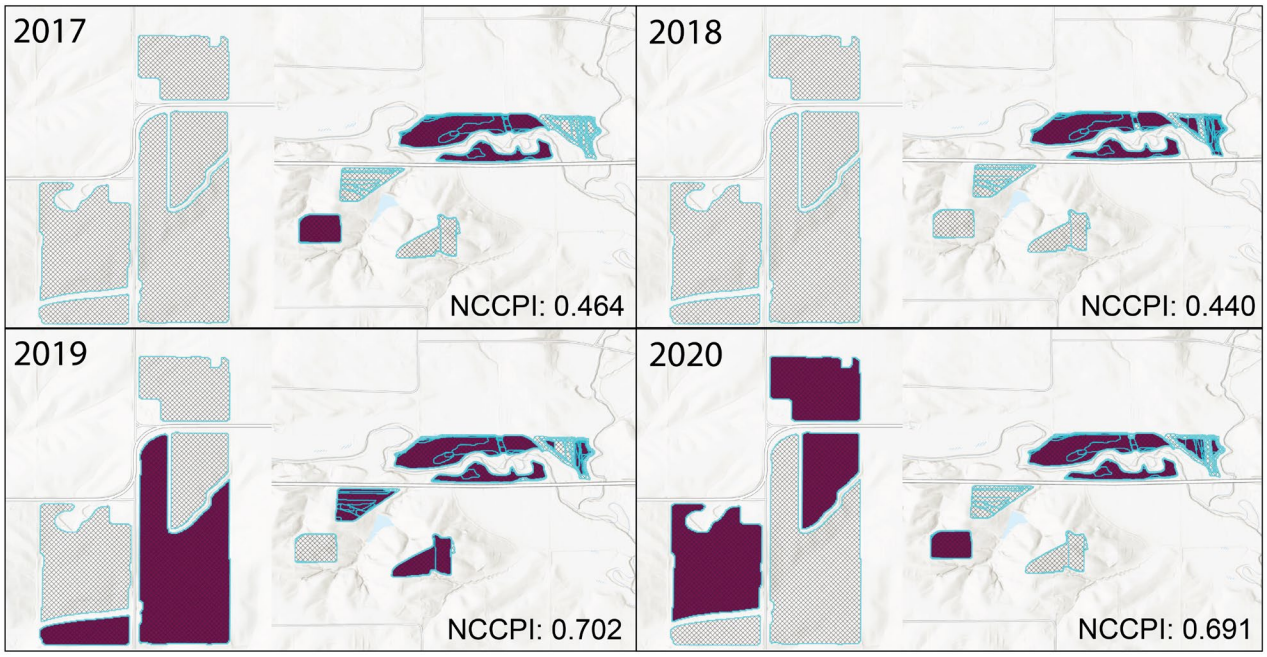


Figure 1: NCCPI temporal change due to field selection, for example, a corn farm. The image has been scaled for viewing purposes. Fields sown with corn are shaded in purple.

3. Data

3.1. Field-level

Several data sources are extracted at the field level and combined to create the data used for analysis. The data includes field-level corn and soybean data, weather data, derived variables, irrigation data, and soil-related productivity data. Short variable descriptions and model variable names are shown in Table 2, and more detailed descriptions are provided below.

3.1.1. Crop Yield Data

Corteva Agriscience provides field-level corn yield data. The unprocessed data comprises an unbalanced panel of 1,247,488 total fields and 253,764 unique farms spanning the period from 2007 to 2020. Each field is accompanied by a geolocated field geometry, giving the precise location. The data was collected through harvest monitors and an agricultural intelligence application. The data are constrained to 13 US states, which represent the vast majority of observations. The states of interest are Iowa, Illinois, Indiana, Maryland, Michigan, Minnesota,

Missouri, Nebraska, New York, Ohio, Pennsylvania, South Dakota, and Wisconsin. Similarly, in the case of soybeans, the unprocessed data is an unbalanced panel of 979,391 total fields and 226,377 unique farms spanning 2007 to 2020. The states of interest are Iowa, Illinois, Indiana, Maryland, Michigan, Minnesota, Missouri, Nebraska, New York, Ohio, Pennsylvania, South Dakota, and Wisconsin.

3.1.2. Weather Data and Derived Variables

Growing season weather conditions heavily influence crop production. Prolonged periods of drought and high temperatures during critical growing periods can negatively affect crop yield (Prasad et al., 2011). The University of Idaho GRIDMET data products are used to create growing degree days. The GRIDMET data product provides daily measurements of various weather variables on a 4-km by 4-km resolution grid (Abatzoglou, 2013). GRIDMET offers more variables than other popular weather data products, including PRISM. GRIDMET delivers a relatively high spatial resolution of 4 km by 4 km and a high frequency of measurement, and it has been available since 1979.

Growing degree days are created using the daily temperature data product from GRIDMET. Growing degree days are created with multiple thresholds, including degree days over 0, 5, 10, 15, 20, 22, 25, 30, 31, and 32°C. This method enables the modeling of more complex growing degree days and follows the implementation steps outlined in Schlenker and Roberts (2006). However, the GRIDMET minimum and maximum daily temperatures are used rather than temperatures from weather stations. The temperature data is extracted at the field location, creating a daily growing degree day value. This series is summed over the growing season from April 1st to October 31st each year to get the field's annual growing degree day value.

3.1.3. Irrigation data

Water availability is a primary factor for commercial crop production, and crop irrigation can significantly affect crop production outcomes. Knowing whether the crop is irrigated at the field level helps model the complex interactions between soil productivity and crop yield. High-resolution irrigation data is collected from NASA's Landsat-derived LGRIP30 data product (Teluguntla et al., 2023). The LGRIP data product is a 30-meter by 30-meter raster layer indicating whether the pixel is irrigated or non-irrigated. The data is extracted at the field level, and the mean of the extracted pixels is taken to derive the mean irrigation score for each field. If the irrigation score is greater than zero, the field is considered irrigated.

3.1.4. gSSURGO Soil Data and NCCPI Productivity Index

The soil-derived NCCPI index is used to model the effect of soil productivity on the statistical moments of crop yield. The data is obtained from the National Resources Conservation Service's gSSURGO data product (Soil Survey Staff, 2020). The soil layer is measured at a 10-meter by 10-meter resolution and is accompanied by a soil score index table that maps the location to 57 soil-quality indexes. The NCCPI is exported as a 10 m resolution raster layer for all states for the NCCPI corn submodel and the soybeans submodel. The NCCPI layers are then extracted from the field locations, and the mean of the extracted pixels is taken as the NCCPI value for the corresponding field. The NCCPI index was selected for the analysis because it has national coverage. It was designed by soil scientists and crop production experts to express the capacity of soil-related crop productivity in a given field.

3.2 Farm-Level Aggregation

Each field in the field-level data has a corresponding mapping to its parent farm. In the data, farms are individual corporate entities or accounts. A large commercial operation can often have

multiple farm entities. These farms are mapped to a higher-level parent, referred to as business entities. The farm-level entity is preferred for this analysis because the same corporate fragmentation is assumed to occur in the crop insurance program. The unbalanced field-level data is aggregated to the farm level by taking an acre-weighted average. Then, the unbalanced farm data is filtered to include only farms with production every year for 10 years, from 2011 to 2020. This filtering operation results in a balanced panel of 11,500 corn farm-year observations. Farms range from 1.2 acres up to 7,600 acres and encompass 1-106 fields of corn. This provides a relatively long panel of continuous corn production. Similarly, the unbalanced farm data for soybeans is filtered to include only farms with production every year for 10 consecutive years, from 2011 to 2020, resulting in a balanced panel of 8,700 farm-year observations. Farms range from 1.4 to 3,571 acres and encompass 1-40 soybean fields. Table 1 shows descriptive statistics of the balanced farm panels used in the analysis. Short variable descriptions and model variable names are shown in Table 2.

Table 1: Descriptive statistics for the farm-level corn data from 2011 to 2020 (10 growing seasons).

Variable	**Corn**	**Soybeans**
Number of farms (per year)	1,150	870
Average farm size[1] (number of acres)	238	188
Average fields per farm	3.74	3.31
Average yield in bushels (per farm)	202	60.4
Standard deviation of yield	40.3	12.1
Average NCCPI	0.741	0.684

[1] Average farm size is defined as the average number of acres in production growing the crop of interest over the period.

Source: Corteva Agrisciences Data

Table 2. Description of variables used in the crop yield moment regressions.

Variable	Description
NCCPI	National Commodity Crop Productivity Index (NCCPI3). The variable approximates the soil's capacity for crop production.

dday_0_14C	Accumulated growing degree days greater than 0 °C and less than 15 °C from April 1st to October 31st of each growing season. Values are scaled by 1000.
dday_15_31C	Accumulated growing degree days greater than 15 °C and less than 32 °C from April 1st to October 31st of each growing season. Values are scaled by 1000.
dday_32C+	Accumulated growing degree days greater than or equal to 32 °C from April 1st to October 31st of each growing season. Values are scaled by 1000.
IRR	An indicator variable identifying if the field has irrigation or is non-irrigated. The variable is derived using the LGRIP30 data product.
acres	Number of acres in production growing the crop of interest over the period.

Notes: For more detailed data descriptions, refer to the data section.

Source: Corteva Agrisciences Data

4. Methodology

4.1. Crop Yield Moment Regressions

Following the strategy described in Wang, Rejesus, and Aglasan 2021, the linkages between crop yield risk, growing degree days, and soil productivity are examined. The parametric residual-based framework of Antle (1983) and Antle and Goodger (1984) is employed in this study for estimating stochastic production functions. This method focuses on the mean, variance, and higher moments of the crop yield distribution. An alternative approach is to use quantile regression to assess the effect of soil-related crop productivity on exposure to downside risk (Chavas, 2019). Following the residual-based framework and the conventions of Wang, Rejesus, and Aglasan (2021), the crop production process is represented by the stochastic production function as follows:

$$y = \mu(x) + \varepsilon, \tag{7}$$

where y is the farm corn yield and x is a vector of growing degree day variables, a soil productivity index variable, an irrigation indicator, year indicators, and a geography indicator (county). The error term ε is assumed to have $E(\varepsilon|x) = 0$ and to be independently distributed. The effect of any variable in x on crop yield risk can be assessed by evaluating the moments of the production function. Specifically of interest are the first four moments: mean, variance,

skewness, and kurtosis. The higher moments of the production function represent the production risk and are outlined below.

$$\hat{\epsilon}^i = [y - \mu(x)]^i = M_i(x) + v_i, \ \forall_i = 2,3,4 \quad (8)$$

where $\hat{\epsilon}^i$ is the i^{th}power of the predicted residuals from the regression shown in Equation (1), $M_i(x)$ is the i^{th} moment function and v_i is the error term. When $i = 2$, Equation (2) represents variance, $i = 3$ skewness, and $i = 4$ kurtosis.

For any variable in x the risk-reducing or risk-enhancing effect can be examined. For example, changes in NCCPI crop productivity can be assessed relative to their impact on the variance, which can be categorized as variance-enhancing, variance-neutral, or variance-decreasing. Similarly, the effect of a change in NCCPI can be examined in relation to changes in skewness and kurtosis. Equations (1) and (2) are estimated using the linear moment method proposed by Antle (1983). The method assumes that the moments are parametric linear functions of independent variables as follows:

$$y = x\beta_1 + \varepsilon \quad (9)$$

$$\hat{\epsilon}^i = x\beta_i + v_i, \qquad \forall_i = 2,3,4 \quad (10)$$

Following Wang, Rejesus, and Aglasan (2021), equations (3) and (4) exhibit heteroskedasticity, requiring robust standard errors to be used in estimation. In this study, we apply robust standard errors.

4.2. Empirical Specification of Crop Yield Moment Regressions

The moments of the production function are estimated at the farm level. The mean (first moment) is estimated, then the higher moments – variance, skewness, and kurtosis – are estimated. In this approach, the parameters for the first moment (mean) are estimated separately. Then, the parameters for the remaining higher moments are estimated in a second step. By

separating it into two stages, the mean is estimated before evaluating the higher moments, and the conditional mean is used to center the higher moments.

4.2.1. First-moment specification

$$\begin{aligned} y_{it} = \alpha_{1i} &+ \beta_1 dday0\text{–}14C_{it} + \beta_2 dday15\text{–}31C_{it} + \beta_3 dday32_{it} \\ &+ \beta_4 acres_{it} + \beta_5 nccpi_{it} + \beta_6 irr_{it} \\ &+ \beta_7 (irr_{it}\ x\ nccpi_{it}) + \beta_8 Trend_t + \beta_9 Trend_t\text{^}2 \\ &+ \phi_c + \varepsilon_{it} \end{aligned} \tag{11}$$

where y_{it} is corn yield (bu/acre) for farm i in period t for years 2011 to 2020; $dday0-14C_{it}$ is the cumulative growing degree days from 0C to 14C measured in degrees Celsius; $dday15-31C_{it}$ is the cumulative growing degree days from 15C to 31C; $dday32C_{it}$ is the cumulative growing degree days equal or above 32C; $acres_{it}$ is the number of corn fields grown in year t; $nccpi_{it}$ is the NCCPI index ranging from [0 to 1], with one being highly productive; irr_{it} is an indicator variable for irrigated or non-irrigated; $Trend_t$ and its squared term are included to capture nonlinear time trends(following Tack, Barkley, and Nalley (2015)); ϕ_c are county fixed effects used to account for county differences; ε_{it} is the error term. Robust standard errors are used to remediate heteroskedasticity and facilitate inference.

4.2.2. Higher moment specification

The higher moments are specified similarly to the mean moment equation. The independent variables and functional form are the same as in Equation (11). Equation (12) shows the higher moment specification for the farm-level regression.

$$\hat{\varepsilon}_{it}^{j} = \alpha_{1i} + \beta_1 dday0\text{–}14C_{it} + \beta_2 dday15\text{–}31C_{it} + \beta_3 dday32_{it} \quad (12)$$
$$+ \beta_4 acres_{it} + \beta_5 nccpi_{it} + \beta_6 irr_{it}$$
$$+ \beta_7 (irr_{it}\ x\ nccpi_{it}) + \beta_8 Trend_t + \beta_9 Trend_t\text{^}2$$
$$+ \phi_c + \varepsilon_{it}$$

where $i = 2,3,4$ and represents the i^{th}power, representing the variance, skewness, and kurtosis. The higher moments $\hat{\varepsilon}_{it}^{j}$ are calculated by taking the predicted values from the estimated mean Equation using the actual values for the independent variables. This process is outlined in Equation (8). For example, the predicted higher moments are calculated using the estimated equations (8). The sign of the parameter estimates indicates the directional effect of each variable on yield variability, skewness, and kurtosis:

For variance of yield ($i = 2$), a positive parameter estimate means the variable increases yield variability, while a negative parameter estimate means the variable decreases yield variability. For the skewness of yield (i = 3), a positive parameter estimate indicates that the variable reduces exposure to downside risks, while a negative parameter estimate indicates that the variable increases exposure to downside risks. For kurtosis of yield ($i = 4$), a positive parameter estimate means the variable increases production risk, while a negative parameter estimate means the variable decreases production risk. In summary, the sign of the parameter estimates shows whether each variable increases or decreases yield variability, downside risk exposure, and production risk. A positive estimate increases variance, decreases skewness, and increases kurtosis, while a negative estimate decreases variance, increases skewness, and decreases kurtosis.

4.2.3. Interpreting model results and the effect of soil productivity on crop risk

The parameters of interest are those associated with the *NCCPI* variable. For example, from Equation (12), the parameters β_5 and β_7. After estimating these parameters, marginal effects can be examined relative to the yield distribution's mean, variance, skewness, and kurtosis. Further, a crop insurance premium scenario analysis is conducted to assess the economic impact of NCCPI changes on premium rates.

4.3. Soil-related crop productivity insurance premium: A scenario analysis

We perform a farm-level scenario analysis based on temporal changes in the NCCPI. For each farm, crop yield distributions are estimated under three temporal changes, NCCPI change scenarios: 1) mean change, 2) one standard deviation change, and 3) two standard deviation change. At the farm level, premium rates for each scenario were estimated using the loss-cost ratio method, based on several commonly used yield distributional assumptions from the literature and an 85% coverage level[1]. To model crop yield distribution as a function of soil productivity and other covariates, we employ maximum likelihood estimation (Atkinson, Adams, and Crocker, 1985; Ramirez, Misra, and James, 2003; Liu and Ramsey, 2023), rather than the least squares method. The parametric yield distribution is estimated conditional on values of NCCPI and additional explanatory variables.

Based on the estimated model, we simulate yield distributions under each NCCPI change scenario. In these scenarios, positive change, negative change, and no change are contrasted relative to the farm mean (holding NCCPI constant) to evaluate the economic impacts of soil productivity on crop insurance premiums. A crop yield distribution is fitted for each farm, and

[1] The 85% coverage level was selected because it is the highest coverage election offered in the FCIP. The analysis was rerun using the 80% and 75% coverage levels as robustness checks. Results were similar and the implications remain unchanged.

crop insurance premiums are calculated and then compared between high and low soil productivity changes. The differential in premium rates under the two changes provides an assessment of the economic impact that soil productivity information can have by informing crop yield risk estimates. Comparing high versus low soil premiums quantifies the value of incorporating soil data into crop insurance rating models, enabling the pricing of risk according to expected yield outcomes conditioned on soil-related crop productivity. The scenario analysis follows a five-step process:

Step 1. Estimate the Average Absolute Change and Standard Deviation of Temporal NCCPI Changes within Each Farm. Specifically, we compute the relative change in NCCPI over time within each farm, defined as: $(NCCPI_t - NCCPI_{t-1}) / NCCPI_{t-1}$. These relative changes are then aggregated across all farms and years to construct three key summary statistics, which are used to define temporal NCCPI change scenarios: 1) the mean of the absolute value of relative changes, denoted $\tilde{\mu}$; 2) A one-standard-deviation relative change, denoted $\tilde{\sigma}$; 3) A two-standard-deviation relative change, denoted $2\tilde{\sigma}$.

Step 2. For each farm, obtain the mean NCCPI value over time, denoted as $\overline{NCCPI_i}$. Using the values of $\tilde{\mu}$ and $\tilde{\sigma}$ from Step 1, construct six scenario-specific NCCPI values that simulate positive and negative shifts in soil productivity. The scaled NCCPI values are calculated as follows $(1 \pm \tilde{\mu})\overline{NCCPI_i}$, $(1 \pm \tilde{\sigma})\overline{NCCPI_i}$, and $(1 \pm 2\tilde{\sigma})\overline{NCCPI_i}$. These scenarios are designed to evaluate the sensitivity of crop yield risk to changes in soil productivity.

Step 3. Fit a specific distribution of yield using MLE with the same control variables used in Equation (11).[2] Applying the NCCPI to six scenarios from Step 2, holding all other variables the same as the dataset, to get the conditional distribution of yield.

[2] Due to the computational intensity of the method employed, we include state fixed effects rather than more granular alternatives(county).

Step 4. Integrate the expected loss up to the 85th coverage level and get the Loss Cost Ratio (LCR).

Step 5. Compare the LCR within farms by using the scaled NCCPI versus the mean NCCPI, while holding the average values of other control variables constant within each farm[3]. A ratio greater than one indicates a premium increase (more risk), and less than one shows a premium decrease (less risk). Several parametric distributions have been proposed and applied in modeling crop yields, including the Normal distribution (Just and Weninger, 1999), and the Weibull distribution (Sherrick et al., 2004; Woodard and Sherrick, 2011). In addition, semi-parametric and nonparametric methods have been suggested as alternatives to rigid parametric assumptions (Ker and Coble, 2003; Goodwin and Ker, 1998; Turvey and Zhao, 1999).

For this study, we use the Beta distribution as the primary model of yield outcomes, due to its popularity in the literature (Nelson and Preckel, 1989; Borges and Thurman, 1994; Babcock and Hennessy, 1996; Goodwin, 2009; Claassen and Just, 2011; Hennessy, 2011). The Beta distribution is well-suited to modeling yields due to its finite support and flexibility in accommodating both positive and negative skewness. To assess the robustness of our results, we also estimate yield distributions using two alternative distributions: the Gamma distribution (Gallagher, 1987) and the Lognormal distribution (Day, 1965; Claassen and Just, 2011). These three distributions also differ in their ability to capture skewness in yield outcomes. The Beta distribution can accommodate both left and right skewness within a bounded interval, making it flexible for modeling yield variability. In contrast, both the Gamma and Lognormal distributions are inherently right-skewed, with the degree of skewness decreasing as their respective shape parameters increase. This property is particularly useful in capturing the asymmetrical risk often

[3] The fitted distribution is conditional on the final year in the sample.

observed in agricultural yield data. This set of distributions allows us to evaluate whether our inferences are sensitive to distributional assumptions.

We briefly outline the estimation strategy below[4]. Let the response vector be given by: $y = \left(y_{\{1,1\}}, \ldots, y_{\{1,T\}}, y_{\{2,1\}}, \ldots, y_{\{n,T\}}\right)^T$, the fully parametric model is specified as:

$$g_k\left(\theta_k^{(it)}\right) = \eta_k^{(it)} = \alpha_{1i} + \beta_k^{(1)} \text{dday_0-14C}_{it} + \tag{13}$$

$\beta_k^{(2)}$ dday_15-31C$_{it}$ + $\beta_k^{(3)}$ dday_32C$_{it}$ + $\beta_k^{(4)} acres_{it}$ +

$\beta_k^{(5)}$ nccpi$_{it}$ + $\beta_k^{(6)}$ irri$_{it}$ + $\beta_k^{(7)}$ (irri$_{it}$ × nccpi$_{it}$) + $\beta_k^{(8)} Trend_t$ +

$\beta_k^{(9)} Trend_t\text{^}2 + \phi_s + \varepsilon_{it}$.

Then, the log-likelihood function is given by:$l(\theta) = \sum_{i=1}^{n} \sum_{t=1}^{T} \log f\left(y_{it} \mid \theta^{(it)}\right)$, where each parameter $\theta_k^{(it)} = g_k^{-1}\left(\eta_k^{(it)}\right)$ is obtained from its corresponding linear predictor via the inverse link function: $\theta_k^{(it)} = g_k^{-1}\left(\eta_k^{(it)}\right)$. We estimate the parameters μ and σ for each distribution. Since variants of parameterizations for the probability density functions (PDFs) exist across the literature, we follow the formulations provided in Rigby et al. (2019) and list the specific probability density functions (PDFs) used in our analysis.

The Beta distribution, selected as our primary model due to its finite support and flexibility in modeling skewness, is specified as:

$f_Y(y|\mu,\sigma) = \frac{1}{B\left(\frac{\mu(1-\sigma^2)}{\sigma^2}, \frac{(1-\mu)(1-\sigma^2)}{\sigma^2}\right)} y^{\frac{\mu(1-\sigma^2)}{\sigma^2}-1}(1-y)^{\frac{(1-\mu)(1-\sigma^2)}{\sigma^2}-1}$, with $E(Y) = \mu$, and $Var(Y) = \sigma^2\mu(1-\mu)$.

[4] We thank an anonymous reviewer for pointing out that the strategies we implement differ from those used by the RMA in practice. However, given the limited time span of the data, it is challenging to estimate individual yield distributions using a linear spline model or other alternatives with only ten years of observations. Instead, we adopt a panel model with fixed effects. By including state fixed effects, we account for broad spatiotemporal heterogeneity. These controls absorb a substantial portion of the farm- and field-level variation, while the remaining cross-sectional heterogeneity is captured through farm-specific variables such as weather conditions and irrigation status.

To use the Beta distribution (which requires 0 < y < 1), we rescale the yield data: $y_{it}^{\text{scaled}} = \frac{y_{it}}{y_{max}+0.001}$, and the estimated quantities and predicted values were subsequently transformed back to the original scale. The Gamma distribution is defined as:

$f_Y(y|\mu,\sigma) = \frac{y^{1/\sigma^2-1}e^{-y/(\sigma^2\mu)}}{(\sigma^2\mu)^{1/\sigma^2}\Gamma(1/\sigma^2)}$, with $E(\mathrm{Y}) = \mu,$ and $\mathrm{Var}(Y) = \sigma^2\mu^2$. Finally, the Lognormal distribution, which accommodates right-skewed outcomes, is expressed as:

$f_Y(y|\mu,\sigma) = \frac{1}{y\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\log y-\mu)^2}{2\sigma^2}\right)$, with mean and variance given $\mathrm{E(Y)} = \exp\left(\mu + \frac{\sigma^2}{2}\right)$, and $\mathrm{Var(Y)} = \exp(2\mu+\sigma^2)\,(\exp(\sigma^2)-1)$. The LCR method is used in Step 4 to derive premium rates and assumes an underlying Gamma distribution of yields. The actuarially fair premium rate is set to the expected LCR:

$$E(LCR) = AF\ Premium\ Rate = \int_0^{\bar{y}_i \cdot C_g} Max\left(0, \bar{y}_i \cdot C_g - y\right) \cdot f(y)dy/\left(\bar{y}_i \cdot C_g\right) \quad (14)$$

where C_g is the coverage election level, $f(y)$ is the corn yield distribution, y is the yield outcome, and $\bar{y}_i$ is the conditional expected yield. The $\bar{y}_i$ is approximated by taking the conditional mean for the farm using the estimated PDFs above. $\bar{y}_i \cdot C_g$ represents the yield guarantee. The $E(LCR)$ is sensitive to the parameterization of the yield distribution $f(y)$.

To provide standard errors and confidence intervals for the estimated ratios under all scenarios, we implement bootstrap inference. Bootstrapping is well-suited for the panel data structure in our setting, where we have many farms (cross-sectional units) but a fixed time dimension (T = 10). Following Wooldridge (2010), we assume random sampling in the cross-sectional dimension. That is, in each bootstrap replication, we treat the entire time series for each farm as a single observation. We resample N = 11,500 corn farms and N = 8,700 soybean farms with replacement, where each selected farm contributes its full set of observations over time.

Importantly, we do not resample individual periods within a farm; only farms are resampled. This approach aligns with standard bootstrap procedures for panel data with large N and small T, providing a realistic approximation to the finite-sample distribution under these conditions. We use B=1000 bootstrap replications to compute the standard deviation (as the bootstrap estimate of standard error) and construct the 95% confidence intervals for the estimates of interest. Note that we assume the temporal changes are not affected by the bootstrap sampling procedure; we retain the mean and standard deviation in the original sample of the NCCPI temporal change throughout the bootstrap process.

4.4 Out-of-Sample Predictive Accuracy and Retain–Cede Rating Game

To demonstrate the predictive advantage of incorporating soil data, we first compare the semi-deviation of predictive distributions from models with and without NCCPI. We assess the out-of-sample predictive accuracy of the estimated mean yields using a 10-fold cross-validation procedure, where the dataset is randomly partitioned into ten approximately equal-sized subsamples (folds). In each iteration, nine folds are used for model estimation, and the remaining fold is used for out-of-sample prediction. To specifically evaluate the risk of overestimation, we consider two directional risk measures: 1) **Semi-Deviation**, defined as Semi-Deviation $= \sqrt{\frac{1}{N}\sum_{it:\,\hat{y}_{it}-y_{it}>0}(\hat{y}_{it}-y_{it})^2}$, where $\hat{y}_{it}$ *is the predicted mean yield for farm* i *in year* t*, and* y_{it} is the actual yield at harvest. This measure captures the variability of prediction errors when the predicted yield exceeds the realized yield. 2). Absolute Semi-Deviation, defined as Absolute Semi-Deviation $= \frac{1}{N}\sum_{it:\,\hat{y}_{it}-y_{it}>0}|\hat{y}_{it}-y_{it}|$, which measures the average magnitude of overestimation errors. These directional risk measures focus exclusively on cases where the predicted mean yield exceeds the realized yield (i.e., $\hat{y}_{it} > y_{it}$). Such overestimation is particularly costly in crop insurance, as it leads to systematic underpricing of premiums and

exposes insurers to unexpected indemnity payments when actual yields fall below guaranteed thresholds. This approach aligns with the structure of the LCR calculation (see Equation (14)). It is well established that soil characteristics influence the moments of the yield distribution. Motivated by this, we incorporate a soil productivity index (NCCPI) into our models to potentially improve predictive accuracy. Next, we shall evaluate the contribution of soil information through an out-of-sample retained-cede rating game, assessing whether incorporating soil variables improves pricing efficiency and adverse selection outcomes.

A substantial amount of literature exists on rating efficiency in crop insurance. A popular framework in this literature is the out-of-sample retain-cede rating game, first introduced by Ker and McGowan (2000). In this game, two players, typically representing different rating methodologies, compete through adverse selection, each estimating premium rates using their model to identify which contracts to retain or cede. This framework has since been widely applied in subsequent studies, including Ker and Coble (2003), Racine and Ker (2006), Harri et al. (2011), Annan et al. (2014), Tolhurst and Ker (2015), Ker, Tolhurst, and Liu (2016), Park, Brorsen, and Harri (2018), Ramsey (2020), Liu and Ker (2020) and Liu and Ford Ramsey (2023).

This game is motivated by the institutional structure of the US crop insurance program. At the same time, private insurers do not set premium rates. The private insurers participate in underwriting outcomes by estimating internal premium rates and deciding which contracts to cede back to the government and which to retain. However, given the relatively short period of our ten-year panel, estimating individual-level yield distributions directly is challenging[5].

[5] For example, Liu and Ker (2020) implement the rating game using relatively long historical data, estimating premiums for 1998 based on data from 1955 to 1997, and then rolling forward year by year over a 20-year period. In contrast, our shorter ten-year panel limits the use of such rolling estimation, requiring an alternative approach.

Following the framework of Ker and McGowan (2000), we adopt a panel model with fixed effects. The estimation approach for farm-specific yield distributions is consistent with that outlined in Section 4.3. To evaluate the value of soil information in premium rating, we compare the loss ratios with and without the inclusion of NCCPI data. In our simulation, we play the role of a private insurer using the additional NCCPI data to estimate conditional yields. We contrast this with a pseudo-agency, which does not incorporate NCCPI information and relies solely on weather and other standard controls[6]. Analogous to the RMA, the pseudo-agency sets premium rates, but private insurers decide which contracts to cede back to the pseudo-agency and which to retain. This simulation allows us to assess the extent to which soil productivity data enables private companies to generate economically and statistically significant excess rents through soil productivity-based adverse selection.

Following the framework of Ker and McGowan (2000), we treat the pseudo-agency as playing a role analogous to the RMA. Private insurers do not set premium rates but do share the accompanying underwriting gains or losses. As a result, they estimate their premium rates to decide which contracts to cede back to the pseudo-agency and which to retain. In our simulation, we assume the role of a private insurer, using the NCCPI data to estimate yields and attempt to capture economic rent through adverse selection against the pseudo-agency. For simplicity, we assume that insurance companies either retain or cede the full underwriting gain or loss associated with each policy. A risk-neutral insurance company will maintain policies for which its internal premium estimates are lower than the pseudo-agency rates and cede those for which its estimates are higher.

[6] The pseudo-agency uses controls including dday_0-14C$_{it}$, dday_15-31C$_{it}$, , dday_32C$_{it}$, $acres_{it}$, and irri$_{it}$, along with state fixed effects and trend terms.

We then compare the selection patterns and resulting loss ratios to assess whether incorporating soil data improves pricing efficiency and enhances the insurer's ability to identify mispriced contracts. A risk-neutral insurance company retains policies it deems desirable (i.e., when $\hat{\pi}^{NCCPI} < \hat{\pi}^{No-NCCPI}$) and cedes those it deems undesirable (i.e., when $\hat{\pi}^{NCCPI} > \hat{\pi}^{No-NCCPI}$). Using the actuarially fair premium rate in Ker and McGowan (2000), for a yield insurance contract or policy that guarantees a percentage (denoted C_g) of the expected yield (denoted $\hat{y}_{it}$) is defined as

$$\pi_{\text{it}} = \Pr\left(Y < \hat{y}_{it} \cdot C_g\right) \cdot \left(\hat{y}_{it} \cdot C_g - E\left(Y \mid Y < \hat{y}_{it} \cdot C_g\right)\right) \tag{15}$$

where $C_g \in [0,1]$ is the coverage level, the expectation operator (E) and (P) are taken with respect to the conditional yield we estimated. [7] For the rating game analysis, we restrict our attention to yield distributions that follow a Beta distribution. To calculate the loss ratio, we weight farms by their actual total crop acreage, assuming either full insurance coverage of all planted acres or that all farmers participate under the same insured proportion.

Let H denote the set of all policies (11,500 for corn and 8,700 for soybeans), S the set of policies the insurance company retains, and S^c the set of policies it cedes. To estimate the yield distributions, we employ a 10-fold cross-validation procedure, where the dataset is randomly partitioned into ten equal-sized subsamples. In each iteration, nine folds are used for model estimation and the remaining fold is used for out-of-sample prediction. For a given set of policies $S \subset H$, the loss ratio is defined as:

$$\text{Loss } Ratio_S = \frac{\sum_{it \in S} \max\left(0, \hat{y}_{it} \cdot C_g - y_{it}\right) \cdot acres_{it}}{\sum_{it \in S} \hat{\pi}_{it}^{no-nccpi} \cdot acres_{it}}. \tag{16}$$

[7] Equations (14) and (15) provide integral and expectation-based formulations of the actuarially fair premium rate, respectively. While expressed differently, they are mathematically equivalent.

Specifically, the program-level loss ratio is calculated by aggregating across all policies in H, while the pseudo-agency loss ratio is calculated by aggregating only over ceded policies in over S^c, following the loss ratio formula defined in Equation (16). We use approximate randomization to assess the statistical significance of the insurer's selection activities. This testing approach simulates the distribution of a test statistic under the null hypothesis when the exact null distribution is unknown. The null hypothesis states that incorporating the NCCPI into reinsurance decisions does not result in lower loss ratios for the insurance company – in other words, NCCPI is uninformative, and the conditional yield distributions are equivalent across retained and ceded policies. Under this hypothesis, the insurer has zero expected gains from using NCCPI and is therefore indifferent between retaining and ceding any given policy.

To simulate the null, we randomly assign retention decisions across policies, retaining each policy with probability equal to the original retention rate (as shown in Tables 7 and 8, Column 3). This ensures that the expected retention rate under the null matches the observed rate. For each situation, we compute the insurer's simulated loss ratio, repeating this process 1,000 times to generate a distribution of loss ratios under the null. We then compare the observed loss ratio to this null distribution. The p-value is calculated as the proportion of simulated loss ratios that are less than the observed statistic. A statistically significant result (i.e., rejection of the null) indicates that conditioning retention decisions on NCCPI leads to significantly lower loss ratios, suggesting that insurers can capture both economically and statistically significant excess rents through soil-productivity-based adverse selection.

5. Results and Discussion

The analysis was separated into two parts: 1) statistical moment regression analysis and 2) premium rate scenario analysis. First, the effect of soil productivity, measured by the NCCPI, on

the distribution of farm-level crop yields was examined. Then, a premium rate scenario analysis was implemented to determine the economic significance of the statistical findings.

5.1. Effect of soil-related crop productivity on the moments of crop yield: Results from moment regression

Table 3 presents the moment regression results for farm-level corn. The variables of interest for the analysis are NCCPI and IRR*NCCPI, representing the NCCPI crop productivity index and the interaction with irrigation ($irrigated = 1$). The mean NCCPI value for the sample of 11,500 farm-year observations is 0.741. According to the regression results, a 5% increase in soil-related crop productivity for a non-irrigated farm is associated with a yield increase of 3.09 bushels per acre (1.53%) and a variance decrease of 2.06%. Examining the mean regression (M1), the weather-related variables, number of acres (acres), soil-related variables, and irrigation variables have statistically significant effects with the expected signs. Furthermore, the variance regression (M2) results indicate that the number of acres comprising a farm significantly reduces farm yield variance. This finding is consistent with expectations, as larger farms are likely to possess greater risk management capabilities to stabilize yields compared to smaller farms. The impacts of soil productivity on M3 and M4 are statistically insignificant.[8]

[8] The results presented include all observations from the balanced panel. Individual state regressions were estimated to ensure results held across states. The results were similar, and the implications remain unchanged.

Table 3: Corn farm-level regression estimates for the effect of soil-related crop productivity on the first four moments of crop yield.

	M1	M2	M3	M4
Weather				
dday_0_14C	48.406***	-1352.209***	19105.548	-11087112.059**
	(5.169)	(288.374)	(35378.160)	(4801482.716)
dday_15_31C	51.554***	-1586.963***	26562.716	-18881749.694***
	(4.377)	(248.411)	(32866.462)	(4899657.879)
dday_32C	-759.653***	16404.478***	-272826.625*	93635472.943***
	(24.687)	(1335.720)	(161414.823)	(21195327.969)
Crop Productivity				
NCCPI	81.124***	-987.791***	47348.081	-2760138.950
	(5.510)	(302.774)	(35568.517)	(4736589.766)
Irrigation				
IRR	19.018***	61.247	24641.148	-1362088.156
	(6.483)	(350.701)	(40024.809)	(5196198.456)
NCCPI × IRR	-24.522***	-101.562	-18023.538	851862.326
	(8.615)	(458.605)	(52473.531)	(7018837.003)
Other Factors				
acres	0.002***	-0.298***	3.638	-3918.831***
	(0.001)	(0.052)	(7.286)	(1431.707)
Time Trend				
	3.246***	181.209***	-2023.829	913942.945**
	(0.603)	(31.340)	(3458.252)	(420879.797)
Trend-squared	-0.168***	-12.420***	31.969	-88719.107**
	(0.051)	(2.770)	(319.712)	(41524.841)
Num.Obs.	11,500	11,500	11,500	11,500
R-squared	0.442	0.082	0.025	0.040

Note: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. County fixed effects are included but not reported.

Source: Authors' calculations

Table 4 presents the moment regression results for farm-level soybeans. The results in Table 4 are similar, albeit with a different magnitude compared to the corn field-level result reports above. The results show a similar effect of NCCPI and IRR*NCCPI variables. A 5% increase in soil-related crop productivity for a non-irrigated farm is associated with a yield increase of 0.96 bu/acre (1.59%) and a variance decrease of 5.53%. Examining the mean regression (M1), the weather-related variables, soil-related variables, and irrigation variables have statistically significant effects with the expected signs. Further, the variance regression (M2) results show

that the number of acres that make up the farm significantly reduces the farm-yield variance (although the magnitude is smaller compared to the corn results. Finally, the impact of soil-related crop productivity on M4 is statistically significant.

Table 4: Soybeans farm-level regression estimates for the effect of soil-related crop productivity on the first four moments of crop yield.

Variable	M1	M2	M3	M4
Weather				
dday_0_14C	15.892***	-145.362***	-5305.186*	-414500.840**
	(2.071)	(50.115)	(2995.286)	(209849.790)
dday_15_31C	13.604***	19.681	1199.957	5797.913
	(1.507)	(28.839)	(1395.667)	(83472.559)
dday_32C	-144.572***	621.974***	-17707.708***	484113.699
	(8.218)	(152.264)	(6233.799)	(314118.899)
Crop Productivity				
NCCPI	27.683***	-238.053***	-1200.413	-369331.355**
	(2.254)	(45.907)	(2348.461)	(158224.527)
Irrigation				
IRR	9.255***	-36.218	2061.045	-26253.128
	(2.286)	(47.909)	(2212.709)	(118471.076)
NCCPI × IRR	-13.027***	59.505	-3273.572	55617.772
	(3.304)	(70.654)	(3279.322)	(175433.938)
Other Factors				
acres	0.003***	-0.055***	0.204	-73.046***
	(0.001)	(0.009)	(0.334)	(16.542)
Time Trend				
	0.172	3.894	26.069	23519.178*
	(0.240)	(4.663)	(224.788)	(13527.799)
Trend-squared	0.050**	-0.312	-16.953	-2516.276*
	(0.020)	(0.416)	(21.602)	(1385.544)
Number of Obs.	8,700	8,700	8,700	8,700
R-squared	0.355	0.065	0.027	0.032

Note: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. County fixed effects are included but not reported.
Source: Authors' calculations

5.2. Spatial and Temporal Variability of Soil-Related Crop Productivity (NCCPI)

The spatial and temporal variability in NCCPI was calculated using the raw data. For instance, Figure 2 displays the NCCPI values for each farm year from 2011 to 2020, encompassing 11,500 corn farm-year observations. Results show that NCCPI values, representing soil-related crop

productivity, differ substantially among farms, ranging from 0.14 for low productivity to 0.97 for high productivity. Figure 3 shows the within-farm temporal variability of NCCPI over the study period. The histogram depicts material changes in NCCPI by farm year. Analysis revealed the mean absolute annual change in NCCPI was 4.08% over time, with a standard deviation of 7.23% for the temporal NCCPI changes. Overall, the examination of the data revealed considerable spatial differences (Figure 2) in soil productivity between farms, as measured by NCCPI, as well as temporal variability within farms from year to year (Figure 3).

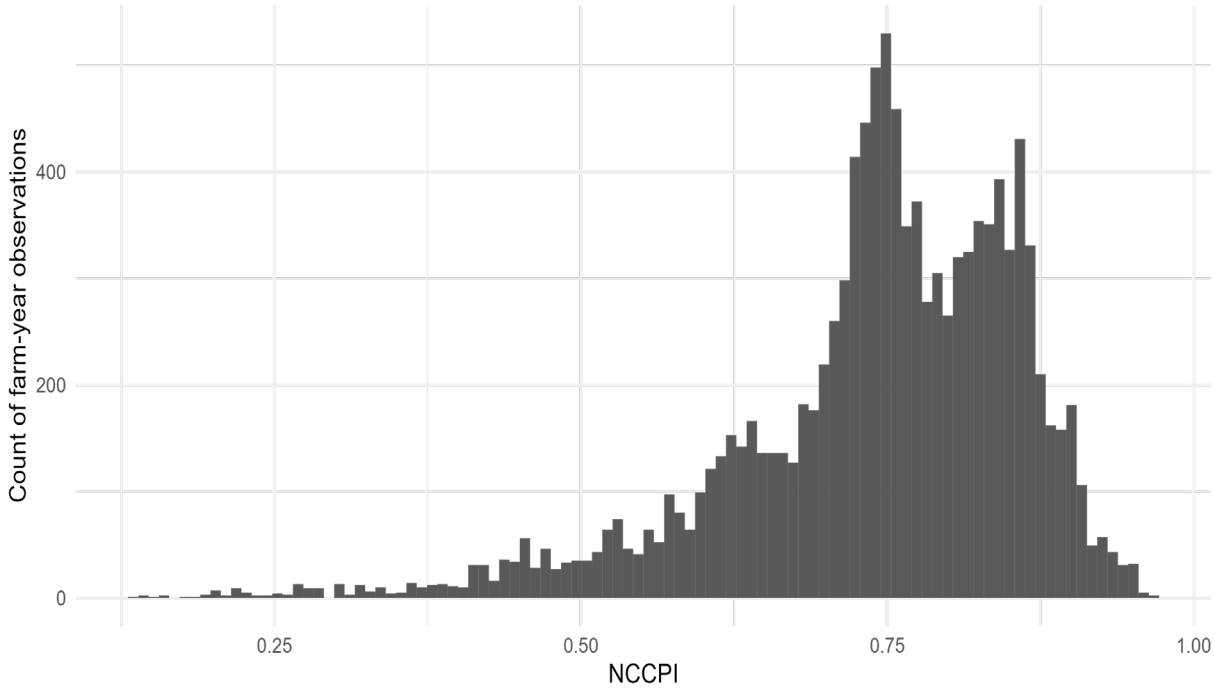


Figure 2: Histogram of 11,500 corn farm-year NCCPI values from 2011-2020.
Source: Corteva Agrisciences Data

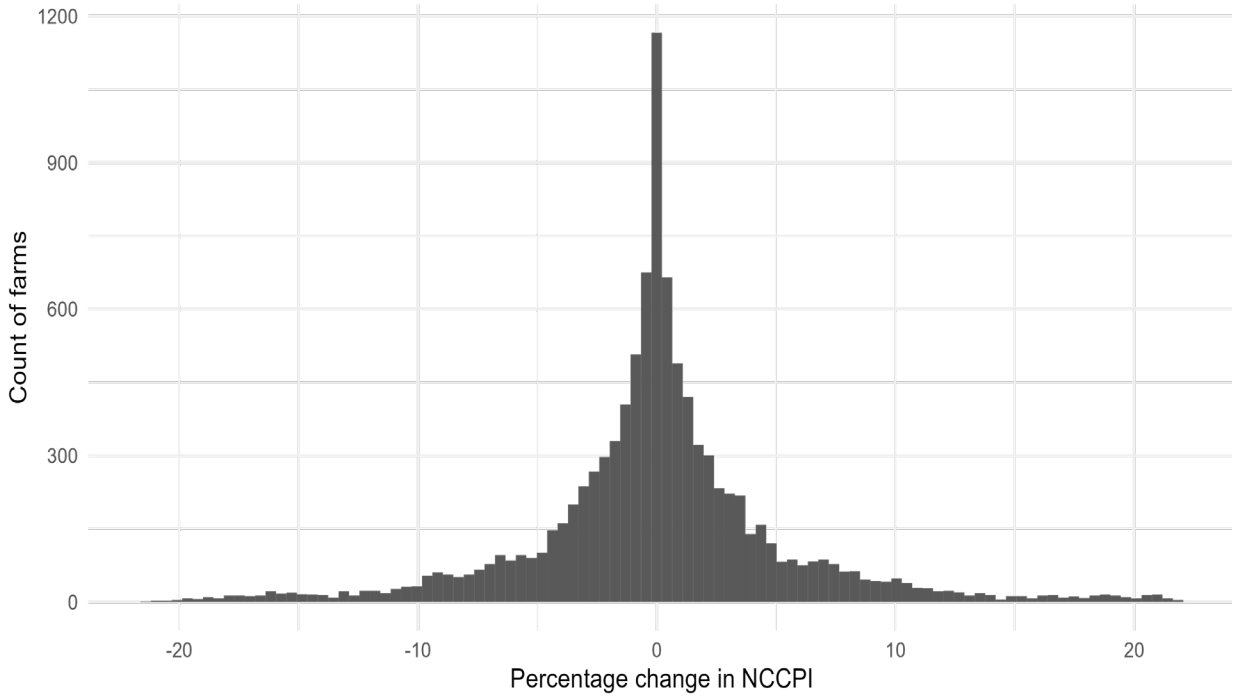


Figure 3: Annual temporal NCCPI percentage changes for corn farms. The x-axis is limited to three standard deviations for illustration.

Source: Corteva Agrisciences Data

Appendix Figures 1 and 2 show the spatial and temporal variability in soybean NCCPI. Appendix Figure 1 displays the NCCPI values for each farm year from 2011 to 2020, encompassing 8,700 farm-year soybean observations. In contrast, Appendix Figure 2 shows the within-farm temporal variability of NCCPI over the study period. The results are slightly different than corn farm-field observations.

5.3 Soil-Related Productivity and Insurance Premium Scenario Analysis

Six NCCPI change scenarios were examined to assess the economic effects of temporal variability in NCCPI on crop insurance premiums. 1) positive mean absolute change, 2) negative mean absolute change, 3) positive one standard deviation change, 4) negative one standard deviation change, 5) positive two standard deviation change, and 6) negative two standard deviation change. Premium rates for each scenario were generated using the loss-cost ratio

method at the farm level, assuming a Beta distribution[9] and an 85% coverage level.[10] These were compared to premiums with NCCPI held at the farm mean, with ratios below one indicating lower premiums and above one indicating higher premiums.

Table 5 outlines the results of the corn crop scenario. At an 85% coverage level, estimates (assuming the yield follows a beta distribution) show that a mean 4.08% increase in NCCPI results in a 4.99% reduction in insurance premiums. In contrast, a 4.08% decrease in NCCPI results in a 5.19% increase in insurance premiums. Furthermore, a one-standard-deviation change in NCCPI, corresponding to a 7.23% increase, resulted in an 8.70% reduction in insurance premiums. In comparison, a 7.23% decrease in NCCPI led to a 9.34% increase in insurance premiums. A two-standard-deviation increase (14.45%) in NCCPI results in a 16.78% reduction in insurance premiums. Finally, a 14.45% decrease in NCCPI results in a 19.36% increase in insurance premiums. Indeed, the results are robust with different assumptions about the yield distribution, and the results remain quite similar. For the elasticity measure, a 1% change in the NCCPI leads to approximately a 1.3% change in the premium rate. Moreover, decreases in NCCPI tend to result in slightly larger changes in the insurance premium compared to equivalent increases. Generally, the scenario analysis above demonstrates that temporal NCCPI variability has a substantial economic effect on crop insurance premiums for corn farmers.

Table 5: Corn scenario analysis results and premium rate impact at the 85% coverage level

NCCPI Scenario	Beta Distribution	Gamma Distribution	Lognormal Distribution
Mean Change			

[9] A normal distribution was also run as a robustness check. Results were similar and implications remain unchanged.
[10] The 85% coverage level was selected because it is the highest coverage election offered in the FCIP. The analysis was rerun using the 80% and 75% coverage levels as robustness checks. Results were similar and the implications remain unchanged.

4.08% NCCPI increase	-0.0499***	-0.0511***	-0.0527***
	(0.0109)	(0.0110)	(0.0121)
	[-0.0715, -0.0284]	[-0.0736, -0.0305]	[-0.0781, -0.0299]
4.08% NCCPI decrease	0.0519***	0.0533***	0.0551***
	(0.0118)	(0.0119)	(0.0132)
	[0.0292, 0.0756]	[0.0315, 0.0783]	[0.0312, 0.0833]
One-SD Change			
7.23% NCCPI increase	-0.0870***	-0.0889***	-0.0917***
	(0.0187)	(0.0188)	(0.0207)
	[-0.1241, -0.0498]	[-0.1272, -0.0532]	[-0.1351, -0.0523]
7.23% NCCPI decrease	0.0934***	0.0960***	0.0992***
	(0.0214)	(0.0217)	(0.0241)
	[0.0524, 0.1370]	[0.0564, 0.1421]	[0.0559, 0.1513]
Two-SD Change			
14.45% NCCPI increase	-0.1678***	-0.1711***	-0.1762***
	(0.0350)	(0.0349)	(0.0384)
	[-0.2358, -0.0972]	[-0.2403, -0.1028]	[-0.2546, -0.1011]
14.45% NCCPI decrease	0.1936***	0.1994***	0.2063***
	(0.0458)	(0.0466)	(0.0519)
	[0.1088, 0.2890]	[0.1163, 0.2993]	[0.1153, 0.3199]

Note: The coverage level was set to the highest insurable amount, 85%. A ratio of unchanged NCCPI of less than one indicates lower premium rates, while a ratio greater than one indicates higher premium rates. Standard errors in parentheses; 95% confidence intervals in brackets; * p<0.10, ** p<0.05, *** p<0.01.
Source: Authors' calculations.

The results are similar for the soybean analysis (Table 6). Indeed, the effects are almost doubled in the case of soybeans compared to corn. For example, at an 85% coverage level, a 5.51% increase in NCCPI results in a 10.92% reduction in insurance premiums. In contrast, a 5.51% decrease in NCCPI results in an approximately 11.89% increase in insurance premiums. A one standard deviation increase (a 9.82% increase) in NCCPI results in an 18.81% reduction in insurance premiums. In contrast, a 9.82% decrease in NCCPI resulted in a 21.89% increase in insurance premiums. Finally, a two-standard-deviation (a 19.63%) increase in NCCPI reduces insurance premiums by 34.75%, while a 19.63% decrease in NCCPI increases insurance premiums by 47.06%. With different assumptions about the yield distribution, the results remain quite similar. For the elasticity measure, a 1% change in the NCCPI leads to approximately a 2% change in the premium rate. However, we should note that the magnitude of the impact is greater

than the estimated effect on corn, as discussed above. Thus, overall, the scenarios in the case of soybeans also demonstrate temporal NCCPI variability and its substantial economic impact on crop insurance premiums.

Table 6: Soybeans scenario analysis results and premium rate impact at the 85% coverage level

NCCPI Scenario	Beta Distribution	Gamma Distribution	Lognormal Distribution
Mean Change			
5.51% NCCPI increase	-0.1092***	-0.0977***	-0.1024***
	(0.0205)	(0.0172)	(0.0191)
	[-0.1430, -0.0649]	[-0.1320, -0.0641]	[-0.1414, -0.0668]
5.51% NCCPI decrease	0.1189***	0.1055***	0.1112***
	(0.0241)	(0.0200)	(0.0223)
	[0.0688, 0.1602]	[0.0681, 0.1472]	[0.0710, 0.1584]
One-SD Change			
9.82% NCCPI increase	-0.1881***	-0.1688***	-0.1767***
	(0.0342)	(0.0289)	(0.0319)
	[-0.2431, -0.1134]	[-0.2255, -0.1119]	[-0.2407, -0.1163]
9.82% NCCPI decrease	0.2189***	0.1938***	0.2046***
	(0.0455)	(0.0377)	(0.0423)
	[0.1251, 0.2978]	[0.1245, 0.2732]	[0.1295, 0.2949]
Two-SD Change			
19.63% NCCPI increase	-0.3475***	-0.3145***	-0.3277***
	(0.0586)	(0.0501)	(0.0549)
	[-0.4366, -0.2167]	[-0.4094, -0.2130]	[-0.4337, -0.2182]
19.63% NCCPI decrease	0.4706***	0.4143***	0.4394***
	(0.1033)	(0.0854)	(0.0964)
	[0.2620, 0.6536]	[0.2615, 0.5979]	[0.2727, 0.6529]

Note: The coverage level was set to the highest insurable amount, 85%. A ratio of unchanged NCCPI of less than one indicates lower premium rates, while a ratio greater than one indicates higher premium rates. Standard errors in parentheses; 95% confidence intervals in brackets; * p<0.10, ** p<0.05, *** p<0.01.

Source: Authors' calculations.

The results from the scenario analysis suggest that the temporal dynamics of soil-related crop productivity (i.e., year-to-year variation in soil productivity of the field selected for crop production) may have a large impact on crop insurance premiums and crop yield risk. These results align closely with Woodard & Verteramo-Chiu's 2017 findings. The authors found evidence of substantial crop insurance mispricing when soil characteristics are not included. The

analysis results suggest that the temporal dynamics of soil-related crop productivity can have a significant impact on crop yield risk and crop insurance premiums. Therefore, these temporal effects may explain some of Woodard's (2014) observed mispricing and the divergence of results compared to Tsiboe and Tack (2022). The large temporal variability observed may help unravel the modest effects of Tsiboe and Tack (2022), which, due to data limitations, missed the temporal aspect of soil-related crop productivity.

Employing the data from Tsiboe and Tack (2022), the dampened effect of soil-related crop productivity on crop insurance premiums is expected, as the temporal dynamics are not present because soil productivity is held fixed through time. One may expect that when the temporal effects are captured, they may average over time (i.e., measured at geolocated fields). However, it may take years of crop yield history to capture a stable average soil productivity effect, as crop rotation changes over time and fields are brought in and out of production. For example, a producer may alter their farm footprint and change their soil-related productivity by purchasing a new parcel of land or forfeiting a rented parcel (field). The results of this study indicate that temporal changes in field rotation-related soil productivity influence premium pricing, even with a relatively established history of data (10 years).

5.4 Simulation Results: Modified Insurance Rating Game

Results for the out-of-sample (cross-validation) prediction are reported in Appendix Tables 1 and 2, indicating that incorporating NCCPI in the estimation of yield distributions reduces prediction errors. These results are robust across the three alternative distributions considered, with particularly strong and comparable performance under the beta and gamma distributions. Given the estimated yield distributions from the previous section, we simulate out-of-sample retain-cede decisions for 11,500 farm-year observations of corn and 8,700 farm-year observations of

soybeans. Three coverage levels – 85%, 75%, and 65% – are simulated to assess whether insurance companies can generate economically and statistically significant excess rents from the use of NCCPI data.Tables 7 and 8 report the results of the rating game comparing an insurance company that incorporates NCCPI data with a pseudo-agency that excludes it, for corn and soybeans, respectively. From left to right, the columns present: the coverage level, the number of contracts, the percentage of contracts retained and ceded, the loss ratio for all contracts, the loss ratio for retained contracts (private loss ratio), the loss ratio for ceded contracts (pseudo-agency loss ratio), and the p-value from the approximate randomization test. Note that p-values close to zero support the efficacy of the insurance company using NCCPI data, while p-values close to one indicate stronger performance by the pseudo-agency, which does not incorporate NCCPI data in contract rating. Across all three coverage levels in the soybean simulations (see Table 8), the insurance company retains more than half of the policies. At each coverage level, the loss ratios consistently follow the same ranking: the pseudo-agency loss ratio is the highest, followed by the total program loss ratio, with the private insurer's loss ratio being the lowest. More importantly, statistically significant evidence of a lower loss ratio is found when using NCCPI data to choose which contracts to retain versus cede, at the 1% significance level for all three coverage levels.

The simulation results for corn (see Table 7) are broadly consistent with those for soybeans. However, statistical significance varies by coverage level: the test result is significant at the 1% level for 85% coverage, at the 10% level for 75% coverage, and at the 20% level for 65% coverage. Across all three coverage levels in the soybean simulations (see Table 8), the insurance company retains more than half of the policies. At each coverage level, the loss ratios consistently follow the same ranking: the pseudo-agency loss ratio is the highest, followed by the

total program loss ratio, with the private insurer's loss ratio being the lowest. More importantly, statistically significant evidence of a lower loss ratio is found when using NCCPI data to choose which contracts to retain versus cede, at the 1% level, across all three coverage levels.

Table 7: Simulation Results: Retention Rates and Loss Ratios for Corn Contracts

Coverage Level	Total Contracts	Retained (%)	Ceded (%)	Loss Ratio (All)	Loss Ratio (With NCCPI)	Loss Ratio (Without NCCPI)	*p*-Value
65%	11,500	67.4%	32.6%	1.483	1.402	1.597	0.172
75%	11,500	67.9%	32.1%	1.043	0.983	1.129	0.084
85%	11,500	68.9%	31.1%	0.850	0.752	0.993	0.000

Source: Authors' calculations.

Table 8: Simulation Results: Retention Rates and Loss Ratios for Soybean Contracts

Coverage Level	Total Contracts	Retained (%)	Ceded (%)	Loss Ratio (All)	Loss Ratio (With NCCPI)	Loss Ratio (Without NCCPI)	*p*-Value
65%	8,700	62.9%	37.1%	1.281	0.999	1.853	0.009
75%	8,700	63.0%	37.0%	1.032	0.814	1.441	0.000
85%	8,700	63.5%	36.5%	0.900	0.713	1.236	0.000

Source: Authors' calculations.

5.5 Summary of Results for Corn and Soybean

Statistical analysis revealed that the soil productivity index has a significant impact on the mean and variance of corn and soybean yields. The impact of the NCCPI index on skewness and kurtosis was statistically insignificant for corn. The statistical analysis revealed that the soil productivity index has a significant impact on the distribution of soybean yields, as indicated by the mean, variance, and kurtosis. The effect on skewness was statistically insignificant. A rating scenario analysis was conducted to assess the economic implications of the temporal variability of NCCPI changes on crop insurance premiums. The insurance premiums were simulated under different scenarios of changing NCCPI over time for both crops. Findings suggest that temporal

NCCPI changes (field rotation) from year-to-year field selection meaningfully affect crop insurance premiums. The results indicate that NCCPI has a statistically and economically significant effect on the moments of crop yield distributions and resultant crop insurance premiums. An out-of-sample retain-cede rating game was conducted to evaluate the superior performance of the model incorporating soil productivity data. The results are consistent with the scenario analysis, indicating that using soil productivity information in crop insurance rating leads to greater efficiency compared to models that exclude it.

6. Summary and Discussion

This study utilized detailed, geolocated farm-level data to examine the relationships between soil productivity, crop yield risk, and crop insurance premiums. This study aimed to evaluate the statistical and economic effects of field rotation (temporal changes) on soil-related crop productivity, introduced through field selection, on crop yield risk and subsequent crop insurance premiums. The results extend the literature in two ways. First, by assessing the effect of soil-related crop productivity on the distribution of farm-level crop yields. Second, by evaluating the potential economic significance of temporal changes in soil-related crop productivity for crop insurance pricing. A dataset comprising 11,500 corn farm-years and 8,500 soybean farm-years, with soil and weather data extracted from precise field locations, was analyzed using residual-based regression techniques.

The analysis demonstrated that soil-related crop productivity, as measured by the National Commodity Crop Productivity Index (NCCPI), has a significant impact on the mean and variance of crop yields. Increased NCCPI was associated with higher average yields and lower yield variability. However, NCCPI did not exhibit a statistically significant relationship with the skewness or kurtosis of corn yields. Results were similar for soybeans, with the addition

of a statistically significant effect on kurtosis. Examination of the data revealed substantial spatial differences in NCCPI between farms and temporal variability in NCCPI within farms over time. On average, NCCPI changed by 4.08% year-to-year, due to crop rotation and land expansion/contraction decisions.

The scenario analysis calculated crop insurance premiums under different NCCPI change assumptions to quantify economic impacts. Results showed that field rotation or temporal changes in NCCPI can have a meaningful effect on premiums. At the 85% coverage level, a 1% change in NCCPI leads to approximately a 1.3% change in the premium rate when evaluated at the farm mean. Similar patterns are observed in the soybean scenario analysis, where the estimated elasticity is higher, around 2. These results indicate that the effect of soil productivity on premium rates is more pronounced for soybeans than for corn. Furthermore, the simulated insurance rating game demonstrates that incorporating NCCPI improves rating efficiency, particularly at higher coverage levels. The research indicates that soil-related productivity has statistically and economically significant relationships with crop yield distributions and insurance premiums. The findings suggest that temporal dynamics (yearly changes due to field selection) in field soil crop productivity can substantially alter yield risk over time. From a crop insurance perspective, farm-level soil data could help improve rating accuracy and reduce mispricing.

The results have several important implications. First, establishing historical yield trends alone may be insufficient for crop insurance rate-making if the temporal dynamics (such as field rotation) of soil productivity are ignored. Second, soil data can be readily incorporated into existing rating methods, such as liability setting and continuous rating formulas, to improve pricing precision. Third, crop rotation decisions may have unintended impacts on yield risk and

insurance costs that are not currently accounted for. However, some limitations of this study should be noted. The data consisted of commercial farms; thus, results may differ for smaller operations. Additionally, it is unclear whether the results can be generalized to other major crops, such as wheat and cotton, which may have different relationships between soil productivity and yields. Future research could investigate mispricing compared to the actual RMA premium rates. A similar analysis using field-level data collected and maintained by the RMA could be useful in examining the extent and current state of mispricing. Farm-level insurable unit RMA premium rates could be compared with premium rates conditioned on soil data that contains the observed temporal effects (i.e., extracted from geolocated fields).

References

Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modeling. I*nternational Journal of Climatology*, 33(1), 121-131.

Adhikari, S., Knight, T. O., & Belasco, E. J. (2012). Evaluation of crop insurance yield guarantees and producer welfare with upward-trending yields. *Agricultural and Resource Economics Review*, 41(3), 367-376.

Albers, M. A., Sinclair, R. R., & Robotham, M. P. (2022). National Commodity Crop Productivity Index (NCCPI) user guide v3.0. USDA NRCS Soil and Plant Science Division.

Annan, F., Tack, J., Harri, A., & Coble, K. (2014). Spatial pattern of yield distributions: implications for crop insurance. *American Journal of Agricultural Economics*, 96(1), 253-268.

Anthony, P., Malzer, G., Sparrow, S., & Zhang, M. (2012). Soybean yield and quality in relation to soil properties. *Agronomy Journal*, 104(5), 1443-1458.

Antle, J. M. (1983). Testing the stochastic structure of production: A flexible moment-based approach. *Journal of Business & Economic Statistics* 1(3): 192–201.

Antle, J. M. and Goodger, W. (1984). Measuring stochastic technology: The case of Tulare milk production. *American Journal of Agricultural Economics* 66(3): 342–350.

Atkinson, S. E., Adams, R. M., & Crocker, T. D. (1985). Optimal measurement of factors affecting crop production: maximum likelihood methods. *American Journal of Agricultural Economics*, 67(2), 414-418.

Babcock, B. A., & Hennessy, D. A. (1996). Input demand under yield and revenue insurance. *American Journal of Agricultural Economics*, 78(2), 416-427.

Borges, R. B., & Thurman, W. N. (1994). Marketing quotas and random yields: marginal effects of inframarginal subsidies on peanut supply. *American Journal of Agricultural Economics*, 76(4), 809-817.

Campbell, C. A., Selles, F., Zentner, R. P., & McConkey, B. G. (1993). Available water and nitrogen effects on yield components and grain nitrogen of zero‐till spring wheat. Agronomy Journal, 85(1), 114-120.

Chavas, J. P. (2019). Adverse shocks in agriculture: the assessment and management of downside risk. *Journal of Agricultural Economics,* 70(3), 731-748.

Chavas, J. P., & Holt, M. T. (1996). Economic behavior under uncertainty: A joint analysis of risk preferences and technology. The Review of Economics and Statistics, 329-335.

Claassen, R., & Just, R. E. (2011). Heterogeneity and distributional form of farm‐level yields. *American Journal of Agricultural Economics*, 93(1), 144-160.

Coble, K. H., Knight, T. O., Pope, R. D., & Williams, J. R. (1997). An expected‐indemnity approach to the measurement of moral hazard in crop insurance. *American Journal of Agricultural Economics*, 79(1), 216-226.

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, 40(1), 79-96.

Coble, K.H. and Barnett, B.J., 2013. Why do we subsidize crop insurance? *American Journal of Agricultural Economics*, 95(2), pp.498-504.

Corwin, D. L., Lesch, S. M., Shouse, P. J., Soppe, R. I. C. H. A. R. D., & Ayars, J. E. (2003). Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agronomy Journal*, 95(2), 352-364.

Cox, M. S., Gerard, P. D., Wardlaw, M. C., & Abshire, M. J. (2003). Variability of selected soil properties and their relationships with soybean yield. *Soil Science Society of America Journal*, 67(4), 1296-1302.

Day, R. H. (1965). Probability distributions of field crop yields. *American Journal of Agricultural Economics*, 47(3), 713-741.

Di Virgilio, N., Monti, A., & Venturi, G. (2007). Spatial variability of switchgrass (Panicum virgatum L.) yield as related to soil parameters in a small field. *Field Crops Research*, 101(2), 232-239.

Finger, R. (2013). Investigating the performance of different estimation techniques for crop yield data analysis in crop insurance applications. *Agricultural Economics*, 44(2), 217-230.

Gallagher, P. (1987). US soybean yields: estimation and forecasting with nonsymmetric disturbances. *American Journal of Agricultural Economics*, 69(4), 796-803.

Glauber, J. W. (2004). Crop insurance reconsidered. *American Journal of Agricultural Economics*, 86(5), 1179-1195.

Glauber, J. W. (2013). The growth of the federal crop insurance program, 1990—2011. *American Journal of Agricultural Economics*, 95(2), 482-488.

Goodwin, B. K. (2009). Payment limitations and acreage decisions under risk aversion: a simulation approach. *American Journal of Agricultural Economics*, 91(1), 19-41.

Goodwin, B. K., & Ker, A. P. (1998). Nonparametric estimation of crop yield distributions: implications for rating group‐risk crop insurance contracts. *American Journal Of Agricultural Economics*, 80(1), 139-153.

Gong, X., Hennessy, D. A., & Feng, H. (2023). Systemic risk, relative subsidy rates, and area yield insurance choice. *American Journal of Agricultural Economics*, 105(3), 888-913.

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(3), 707-717.

Hennessy, D. A. (2011). Modeling stochastic crop yield expectations with a limiting beta distribution. *Journal of Agricultural and Resource Economics*, 177-191.

Just, R. E., & Weninger, Q. (1999). Are crop yields normally distributed? *American Journal of Agricultural Economics*, 81(2), 287-304.

Just, R. E., Calvin, L., & Quiggin, J. (1999). Adverse selection in crop insurance: Actuarial and asymmetric information incentives. *American Journal of Agricultural Economics*, 81(4), 834-849.

Kane, D. A., Bradford, M. A., Fuller, E., Oldfield, E. E., & Wood, S. A. (2021). Soil organic matter protects US maize yields and lowers crop insurance payouts under drought. *Environmental Research Letters,* 16(4), 044018.

Ker, A. P., & Coble, K. (2003). Modeling conditional yield densities. *American Journal of Agricultural Economics*, 85(2), 291-304.

Ker, A. P., & McGowan, P. (2000). Weather-based adverse selection and the US crop insurance program: The private insurance company perspective. *Journal of Agricultural and Resource Economics*, 386-410.

Ker, A. P., Tolhurst, T. N., & Liu, Y. (2016). Bayesian estimation of possibly similar yield densities: implications for rating crop insurance contracts. *American Journal of Agricultural Economics*, 98(2), 360-382.

Kravchenko, A. N., & Bullock, D. G. (2000). Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy Journal*, 92(1), 75-83.

Liu, Y., & Ker, A. P. (2020). Rating crop insurance contracts with nonparametric Bayesian model averaging. *Journal of Agricultural and Resource Economics*, 45(2), 244-264.

Liu, Y., & Ramsey, A. F. (2023). Incorporating historical weather information in crop insurance rating. *American Journal of Agricultural Economics*, 105(2), 546-575.

Li, X., Tack, J. B., Coble, K. H., & Barnett, B. J. (2016). Can crop productivity indices improve crop insurance rates? Selected Paper prepared for presentation at the 2016 Agricultural & Applied Economics Association Annual Meeting, Boston, Massachusetts.

Lubowski, R. N., Bucholtz, S., Claassen, R., Roberts, M. J., Cooper, J. C., Gueorguieva, A., & Johansson, R. C. (2006). Environmental effects of agricultural land-use change: the role of economics and policy (No. 1477-2016-121102).

Machmuller, M., Kramer, M., Cyle, T., Hill, N., Hancock, D., & Thompson, A. (2015). Emerging land use practices rapidly increase soil organic matter. Nat Commun 6, 6995 https://doi.org/10.1038/ncomms7995

Mahul, O., & Stutley, C. J. (2008). Government support to agricultural insurance. Challenges and options for developing countries. Washington, DC: The World Bank, 12.

Martín, N. F., Bollero, G., & Bullock, D. G. (2005). Associations between field characteristics and soybean plant performance using canonical correlation analysis. *Plant and Soil*, 273, 39-55.

Nelson, C. H., & Preckel, P. V. (1989). The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*, 71(2), 370-378.

Nelson, C. H., & Preckel, P. V. (1989). The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*, 71(2), 370-378.

Nyiraneza, J., Cambouris, A. N., Ziadi, N., Tremblay, N., & Nolin, M. C. (2012). Spring wheat yield and quality related to soil texture and nitrogen fertilization. *Agronomy Journal*, 104(3), 589-599.

Oldfield, E. E., Bradford, M. A., & Wood, S. A. (2019). Global meta-analysis of the relationship between soil organic matter and crop yields. *Soil*, 5(1), 15-32.

Pardhasaradhi Teluguntla, Prasad Thenkabail, Adam Oliphant, Murali Gumma, Itiya Aneece, Daniel Foley, and Richard McCormick (2023). Landsat-Derived Global Rainfed and Irrigated-Cropland Product 30 m V001. 2023, distributed by NASA EOSDIS Land Processes DAAC. Accessed 2023-04-01.

Park, E., Brorsen, B. W., & Harri, A. (2019). Using Bayesian Kriging for spatial smoothing in crop insurance rating. *American Journal of Agricultural Economics*, 101(1), 330-351.

Prasad, P. V. V., Pisipati, S. R., Momčilović, I., & Ristic, Z. (2011). Independent and combined effects of high temperature and drought stress during grain filling on plant yield and chloroplast EF‐Tu expression in spring wheat. *Journal of Agronomy and Crop Science*, 197(6), 430-441.

Racine, J., & Ker, A. (2006). Rating crop insurance policies with efficient nonparametric estimators that admit mixed data types. *Journal of Agricultural and Resource Economics*, 27-39.

Ramirez, O. A., Misra, S., & Field, J. (2003). Crop‐yield distributions revisited. American *Journal of Agricultural Economics*, 85(1), 108-120.

Ramsey, A. F. (2020). Probability distributions of crop yields: a Bayesian spatial quantile regression approach. *American Journal of Agricultural Economics*, 102(1), 220-239.

Rejesus, R. M., Coble, K. H., Miller, M. F., Boyles, R., Goodwin, B. K., & Knight, T. O. (2015). Accounting for weather probabilities in crop insurance rating. J*ournal of Agricultural and Resource Economics*, 306-324.

Rigby, R. A., Stasinopoulos, M. D., Heller, G. Z., & De Bastiani, F. (2019). Distributions for modeling location, scale, and shape: Using GAMLSS in R. Chapman and Hall/CRC.

Schlenker, W., & Roberts, M. J. (2006). Nonlinear effects of weather on corn yields. *Review of Agricultural Economics*, 28(3), 391-398.

Sene, M., Vepraskas, M. J., Naderman, G. C., & Denton, H. P. (1985). Relationships of soil texture and structure to corn yield response to subsoiling. *Soil Science Society of America Journal*, 49(2), 422-427.

Skees, J. R., & Reed, M. R. (1986). Rate making for farm‐level crop insurance: Implications for adverse selection. *American Journal of Agricultural Economics*, 68(3), 653-659.

Smith, V. H., & Goodwin, B. K. (1996). Crop insurance, moral hazard, and agricultural chemical use. *American Journal of Agricultural Economics*, 78(2), 428-438.

Soil Survey Staff. The Gridded Soil Survey Geographic (gSSURGO) Database for Multiple States. United States Department of Agriculture, Natural Resources Conservation Service (2020). Available online at https://gdg.sc.egov.usda.gov/. Accessed 2023-04-01.

Tack, J., Barkley, A., & Nalley, L. L. (2015). Effect of warming temperatures on US wheat yields. *Proceedings of the National Academy of Sciences*, 112(22), 6931-6936.

Tack, J., Coble, K., & Barnett, B. (2018). Warming temperatures will likely induce higher premium rates and government outlays for the US crop insurance program. *Agricultural Economics*, 49(5), 635-647.

Tolhurst, T. N., & Ker, A. P. (2015). On technological change in crop yields. *American Journal of Agricultural Economics*, 97(1), 137-158.

Tsiboe, F., & Tack, J. (2022). Utilizing Topographic and Soil Features to Improve Rating for Farm‐Level Insurance Products. *American Journal of Agricultural Economics*, 104(1), 52-69.

Turvey, C. G., & Zhao, J. (1999). Parametric and nonparametric crop yield distributions and their effects on all-risk crop insurance premiums (No. 1620-2016-134858).

Vyas, S., Dalhaus, T., Kropff, M., Aggarwal, P., & Meuwissen, M. P. (2021). Mapping global research on agricultural insurance. *Environmental Research Letters*, 16(10), 103003.

Wang, H. H., Tack, J. B., & Coble, K. H. (2020). Frontier studies in agricultural insurance. The Geneva Papers on Risk and Insurance-Issues and Practice, 45, 1-4.

Wang, R., Rejesus, R. M., & Aglasan, S. (2021). Warming temperatures, yield risk and crop insurance participation. *European Review of Agricultural Economics*, 48(5), 1109-1131.

Woodard, J. D., & Sherrick, B. J. (2011). Estimation of mixture models using cross‐validation optimization: implications for crop yield distribution modeling. *American Journal of Agricultural Economics*, 93(4), 968-982.

Woodard, J. D., & Verteram-Chiu, L. J. (2017). Efficiency impacts of utilizing soil data in the pricing of the federal crop insurance program. *American Journal of Agricultural Economics*, 99(3), 757-772.

Woodard, J. D., Sherrick, B. J., & Schnitkey, G. D. (2011). Actuarial impacts of loss cost ratio ratemaking in US crop insurance programs. *Journal of Agricultural and Resource Economics*, 211-228.

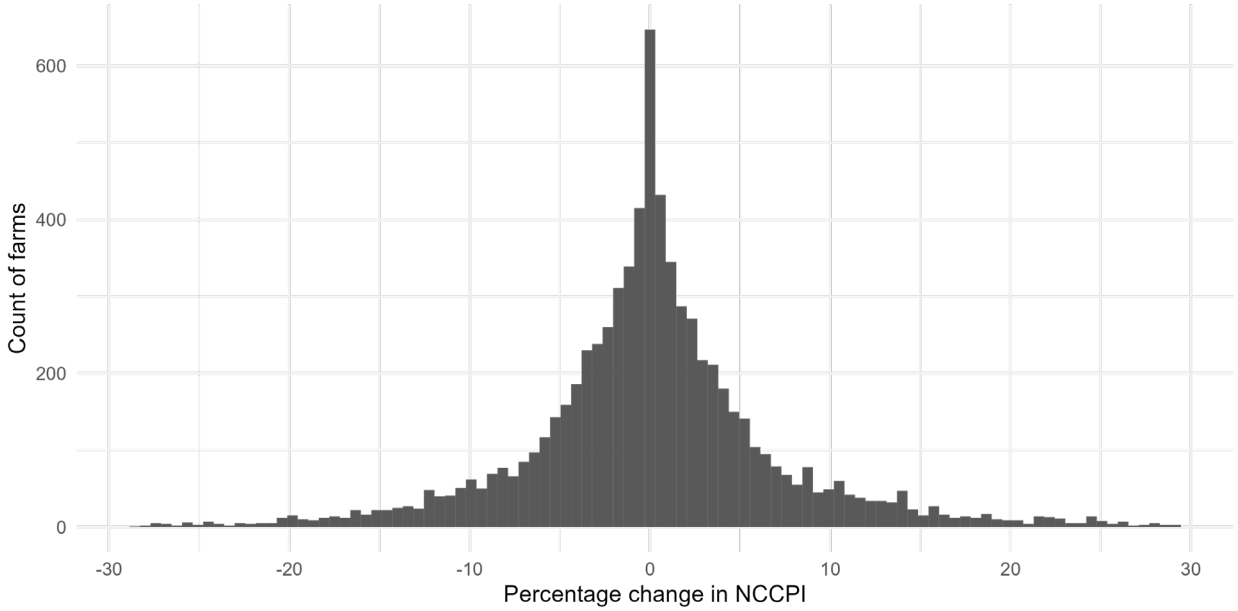
Woodard, J. D. (2014). Impacts of weather and time horizon selection on crop insurance ratemaking: A conditional distribution approach. *North American Actuarial Journal*, 18(2), 279-293.

Wooldridge, J. M. (2010). Econometric analysis of cross-section and panel data. MIT press.

Wu, S., Goodwin, B. K., & Coble, K. (2020). Moral hazard and subsidized crop insurance. *Agricultural Economics*, 51(1), 131-142.

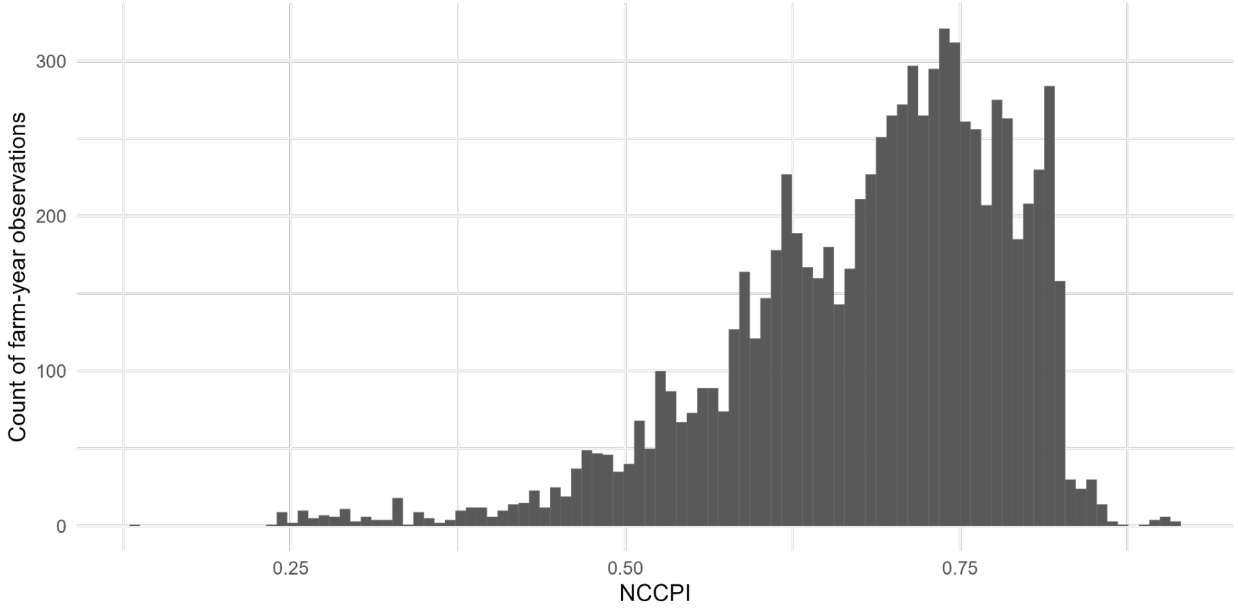
Young, C. E., Vandeveer, M. L., & Schnepf, R. D. (2001). Production and price impacts of US crop insurance programs. *American Journal of Agricultural Economics*, 83(5), 1196-1203.

Appendix



Appendix Figure 1: Histogram of 8,700 soybean farm-year NCCPI values from 2011-2020.

Source: Authors' calculations.



Appendix Figure 2: Annual temporal NCCPI percentage changes for soybean farms. The x-axis is limited to three standard deviations for illustration.

Source: Authors' calculations.

Appendix Table 1: Predictive Accuracy for Corn Yields

Beta	With NCCPI	Without NCCPI
Semi-RMSE	25.454	26.041
Semi-MAE	12.727	13.050
Gamma		
Semi-RMSE	25.917	26.505
Semi-MAE	12.963	13.297
Lognormal		
Semi-RMSE	26.810	27.467
Semi-MAE	13.423	13.799

Source: Authors' calculations.

Appendix Table 2: Predictive Accuracy for Soybean Yields

Beta	With NCCPI	Without NCCPI
Semi-RMSE	7.939	8.125
Semi-MAE	4.204	4.272
Gamma		
Semi-RMSE	7.933	8.115
Semi-MAE	4.202	4.266
Lognormal		
Semi-RMSE	8.065	8.219
Semi-MAE	4.284	4.339

Source: Authors' calculations.