Systemic Risk, Relative Subsidy Rates and Area Yield Insurance Choice

Xuche Gong, David A. Hennessy, Hongli Feng

Working Paper 22-WP 636
UPDATE
September 2022

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Xuche Gong is PhD Student, Department of Economics, Iowa State University, Ames, Iowa 50014. E-mail: xcgong@iastate.edu.

David A. Hennessy is Professor, Department of Economics, Iowa State University, Ames, Iowa 50014. E-mail: hennessy@iastate.edu.

Hongli Feng is Assistant Professor, Department of Economics, Iowa State University, Ames, Iowa 50014. E-mail: hfeng@iastate.edu.

For questions or comments about the contents of this paper, please contact Xuche Gong, xcgong@iastate.edu.

Received: 12 February 2020 | Accepted: 27 July 2022

DOI: 10.1111/ajae.12342

ARTICLE

Systemic risk, relative subsidy rates, and area yield insurance choice

Xuche Gong | **David A. Hennessy** | **Hongli Feng**

Department of Economics, Iowa State University, Ames, Iowa, USA

Correspondence
Xuche Gong, Department of Economics, Iowa State University, Ames, Iowa, USA.
Email: xcgong@iastate.edu

Abstract

We investigate the nature of crop yield systemic risk and its implications for farmers' area yield insurance choices. A theory-grounded, normalized measure of systemic risk, R^2, is developed and made amenable to empirical analysis through the logit link. The measure is estimated on a large-scale, unit-level corn yield dataset. We find that systemic risk explains less than half of total unit yield variability on average, suggesting that the risk management effectiveness of area yield insurance is low. By relating natural resource endowments with systemic risk, we find that more excessive heat and drought events lead to larger while more excessive precipitation events lead to smaller systemic risk. We then study whether current area yield insurance subsidy rates provide farmers with sufficient subsidy transfers to compensate for the uncovered risk exposure associated with area yield insurance. A new concept, the threshold relative area subsidy rate (TRASR), quantifies a lower bound on the ratio of area yield insurance subsidy rate over individual yield insurance subsidy rate below which risk-averse farmers should always prefer individual yield insurance to area yield insurance. The calibrated TRASR values indicate that current area yield insurance subsidy rates discourage farmers from choosing area yield insurance over individual yield insurance, especially at low area yield insurance coverage levels. We also find that TRASR is positively correlated with systemic risk at the unit-level of analysis, suggesting that farmers who like area yield insurance's risk management features will dislike its transfer implications and vice versa.

KEYWORDS
basis risk, climate change, idiosyncratic risk, individual yield insurance, natural resource endowments, risk management, transfer seeking

JEL CLASSIFICATION
D81, Q12, Q18

1 | INTRODUCTION

Systemic yield risk exists because yields tend to be strongly and positively correlated in a region. This risk can lead to substantial indemnity payments in catastrophic weather years, which is considered a major reason why private crop insurance providers would be unwilling to take on the risk (Miranda & Glauber, 1997). The U.S. Federal Crop Insurance Program (FCIP) has evolved into a public–private partnership in which the federal government shares risks with private providers. Systemic risk has also encouraged the federal government to promote area insurance contracts, which pay indemnity based on average yield or revenue loss in an area rather than on individuals' losses. To the extent that area yield or revenue outcomes correspond with farm-level outcomes, indemnification based on pertinent area statistics can meet farmer financial shortfalls when needed. In addition, compared to individual insurance plans, area insurance plans have the advantages of lower information requirements and less potential for moral hazards (Barnett et al., 2005; Miranda, 1991; Skees et al., 1997). However, as shown in Figure 1, the share of area insurance insured acres in total

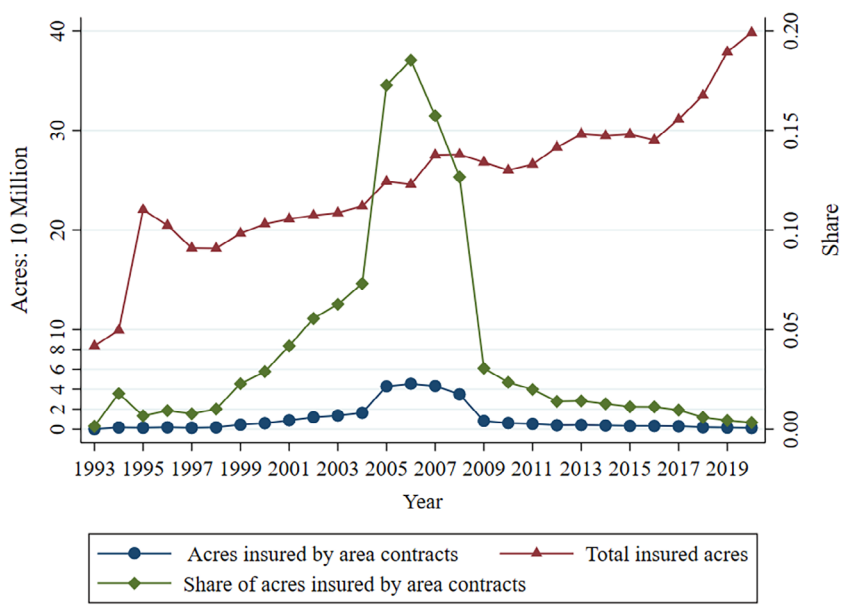

FIGURE 1 Share of acres insured by area insurance contracts for all crops, 1993–2020
Notes: 1. Source: Summary of Business (SOB), 1993–2020, Risk Management Agency (RMA). 2. This figure plots acres insured by area insurance contracts and acres insured by any kind of crop insurance contract for all crops over 1993–2020 (left *y*-axis), and the share of area insurance insured acres in total insured acres for all crops over 1993–2020 (right *y*-axis). 3. Area insurance programs before 2014 are Group Risk Plan, Group Risk Income Protection, and Group Risk Income Protection with Harvest Revenue Option. Area insurance programs starting in 2014 are Area Yield Protection, Area Revenue Protection, and Area Revenue Protection with Harvest Price Exclusion.

insured acres has never exceeded 20% ever since its inception. This observation poses questions regarding the extent to which systemic risk can account for individual farmers' risks and whether farmers view area insurance plans as competitive alternatives to individual plans. In this paper, we seek to answer these questions by first modeling, measuring, and decomposing systemic risk in corn yield across the Greater Midwest. We then apply findings in systemic risk, together with the current FCIP premium subsidy features, to understand low area insurance demand in the United States and likely impacts of climate change for systemic risk.

Following Miranda (1991), we use a linear additive model (LAM) to decompose unit yield into a systemic component that is correlated with area yield and an idiosyncratic component that is uncorrelated with area yield. We then measure systemic yield risk as the proportion of unit yield variation explained by area yield variation. Bounded between 0 and 1, this index provides a unitless measure of the proportion of yield risk that AYP can potentially cover. We also show how LAM enables a further decomposition of systemic risk into three more fundamental components: sensitivity of unit yield to area yield, area yield variance, and a unit's idiosyncratic yield variance. By relating each systemic risk component to climate and land variables, we further explore how systemic yield risk is determined by natural resource endowments.

Using a large-scale, unit-level dataset, we estimate systemic risk for corn yield across 12 major corn production states in the Midwest. We find that the proportion of unit yield risk that is systemic to county yield is 46% on average, suggesting that AYP can provide only modest risk protection. Systemic risk is highest at the southern and western fringes of the Corn Belt but comparatively low in much of Iowa and surrounding areas, indicating a mismatch between major corn production areas and high systemic risk areas. Regions with high systemic yield risk are characterized by large county yield variances and strong unit-to-county yield sensitivities. These regions typically experienced more drought and excessive heat events in the sample years, suggesting that the common downward shift of individual yields in these catastrophic events is the primary cause of high systemic risk.

Given AYP's limited risk management effectiveness, we then turn to investigating whether the current AYP subsidy level delivers sufficient transfers to compete with YP. Although FCIP is mainly intended to provide farmers with risk management tools, it is widely accepted that subsidy transfers at least partly motivate farmers to insure (Du et al., 2017; Goodwin & Smith, 2013). We introduce a novel concept, the Threshold Relative Area Subsidy Rate or TRASR, which is defined as the ratio of the AYP subsidy rate over the YP subsidy rate such that the two insurance plans provide equal expected net returns. As YP provides better risk protection, TRASR constitutes a lower bound on the relative subsidy level needed to entice risk-averse farmers to choose AYP over YP. By calibrating TRASR values for corn farmers in the 12 chosen Midwest states and comparing these values to current relative subsidy rates for AYP over YP, we find that the current premium subsidy schedule discourages corn farmers from choosing AYP over YP, especially at low AYP contract coverage levels. We also find that TRASR is positively correlated with systemic risk at the unit level of analysis, suggesting a decision dilemma: Farmers who enjoy good risk protection from AYP are more likely to find current AYP subsidy rates financially unrewarding, whereas those who find current AYP subsidy rates financially rewarding are less likely to be well protected by AYP.

This paper makes several significant and practical contributions to the literature. First, it adds to the measurement and conceptual understanding of systemic risk. Previous studies generally measure systemic risk in terms of correlations within insurers' portfolios (Goodwin & Hungerford, 2015; Hayes et al., 2003; Miranda & Glauber, 1997) or yield correlations within a given region (Du et al., 2018; Tack & Holt, 2016; Wang & Zhang, 2003). Insurer portfolio analyses are functions of pertaining crop insurance plans and choices, and so are limited and indirect in shedding light on the underlying nature of systemic yield risk. Yield correlation analyses hew closer to the problem primitives. However, using a long panel dataset for Illinois and Kansas farms, Zulauf et al. (2013) finds that farm-to-county yield correlation has limited ability to explain the share of a farm's loss that is systemic with county yield. Our systemic risk measure is

more structural in that it considers both how closely unit yield varies with county yield and what that means for total yield variation. The study closest to ours in modeling systemic risk is Claassen and Just (2011), which uses a sample of unit yields together with ANOVA methods to break down unit yield variance into systemic, random, and interaction components. Our paper departs from Claassen and Just (2011) by decomposing systemic risk into more fundamental components, a distinction that enables us to study how changes in unit yield and county yield variability affect systemic risk.

Second, this paper provides new perspectives on farmers' apparently low demand for area insurance plans in the United States. The existing literature generally compares AYP with YP in effects on yield variance (Barnett et al., 2005; Miranda, 1991; Stigler & Lobell, 2021) or certainty equivalent revenues (Awondo & Datta, 2018; Deng et al., 2007). These methods depend on assumptions about the yield distribution or the utility function and underestimate the effect of basis risk on farmers' AYP choices. Basis risk exists for AYP buyers because they will not receive any indemnity when individual yield loss occurs but area yield performs well. This type of risk is generally considered a major cause of low participation in weather index insurance worldwide (Clarke, 2016; Hill et al., 2016; Keller & Saitone, 2022). This paper circumvents the limitations of the existing literature by directly studying AYP's risk management effectiveness and premium subsidy transfers. Our systemic yield risk estimates suggest limited risk management effectiveness for AYP contracts and thus a high level of basis risk, whereas the calibrated TRASR values demonstrate that the current premium subsidy schedule favors YP over AYP in general. This paper further advances the literature by revealing a positive correlation between systemic yield risk and TRASR, suggesting that, in the presence of subsidized YP, AYP is typically unable to simultaneously meet farmers' risk management and transfer-seeking needs.

Third, this paper provides opportunities for projecting future systemic risk patterns and area insurance demand. Recent studies have found that climate change would likely increase yield variability, crop insurance demand, and FCIP costs (Annan & Schlenker, 2015; Ray et al., 2015; Tack et al., 2018), but little is known about how systemic risk will evolve and what role area insurance could play in the presence of climate change. Previous studies have found that yield correlations will be higher in extreme weather years (Du et al., 2018; Tack & Holt, 2016). Our results confirm earlier findings that more excessive heat and drought events will lead to higher systemic risk. However, we also find that more excessive precipitation events will lead to lower systemic risk, perhaps because floods are localized to low-lying land. As of 2020, the Federal National Climate Assessments for both the Midwest and the Great Plains project a hotter climate with more drought and more flooding events.[1] Thus, our findings suggest that climate change implications for future systemic risk patterns depend on the specific extreme weather events that an area is likely to incur. Given this and the finding of a positive correlation between systemic yield risk and TRASR, how climate change affects future AYP demand patterns will also be area specific.

The remainder of this paper proceeds by first presenting our conceptual framework for modeling systemic risk and TRASR. Some propositions and conjectures are then developed, and these inferences are examined in the empirical sections. We conclude with a summary of major findings as well as some discussions on policy implications and items for future research.

2 | CONCEPTUAL FRAMEWORK

Our focus is on yield risk so we assume throughout that price is non-random. To avoid unnecessary notation, we set output price equal to 1 and ignore it henceforth. All yield, premium, subsidy, and indemnity variables used in this paper are on a per-acre basis.

[1]See https://nca2018.globalchange.gov/chapter/21/ and https://nca2018.globalchange.gov/chapter/22/.

2.1 | Modeling systemic risk

Following Miranda (1991), we apply LAM to characterize the relationship between unit yield and county yield,[2]

$$\widetilde{y}_i = \mu_i + \beta_i\left(\widetilde{y}_c - \mu_c\right) + \varepsilon_i. \tag{1}$$

Here $\widetilde{y}_i \in \left[0, y_i^u\right]$ and $\widetilde{y}_c \in \left[0, y_c^u\right]$ are, respectively, unit yield and county yield variables, where subscript i denotes insurance unit and c denotes county. The two yield variables are continuously distributed over the associated yield range with respective expected values $\mu_i = \mathrm{E}\left(\widetilde{y}_i\right)$ and $\mu_c = \mathrm{E}\left(\widetilde{y}_c\right)$, where $\mathrm{E}(\cdot)$ is the expectation operator. The error term has mean zero, $\mathrm{E}(\varepsilon_i) = 0$, and is uncorrelated with county yield, $\mathrm{Cov}\left(\widetilde{y}_c, \varepsilon_i\right) = 0$, where $\mathrm{Cov}(\cdot)$ is the covariance operator. LAM has been widely used in crop insurance analysis and farm-level policy studies to simulate farm yield from county yield when farm-level data are unavailable (Carriquiry et al., 2008; Coble & Dismukes, 2008). Ramaswami and Roe (2004) show that whenever systemic and idiosyncratic risks are additive in individual yields then LAM can be obtained by applying the law of large numbers.

By way of Equation (1), we decompose unit yield deviation from expectation into a systemic component, $\beta_i\left(\widetilde{y}_c - \mu_c\right)$, which is correlated with county yield, and an idiosyncratic part, ε_i, which is uncorrelated with county yield. Coefficient β_i measures the sensitivity of unit yield to county yield. Because $\beta_i \le 0$ is rare for crop production, we assume that $\beta_i > 0$ throughout the paper. Letting $\sigma_i^2 = \mathrm{Var}\left(\widetilde{y}_i\right)$, $\sigma_c^2 = \mathrm{Var}\left(\widetilde{y}_c\right)$ and $\sigma_{\varepsilon_i}^2 = \mathrm{Var}(\varepsilon_i)$, where $\mathrm{Var}(\cdot)$ is the variance operator, we have $\sigma_i^2 = \mathrm{Var}\left(\widetilde{y}_i - \mu_i\right) = \beta_i^2 \mathrm{Var}\left(\widetilde{y}_c - \mu_c\right) + \mathrm{Var}(\varepsilon_i) = \beta_i^2 \sigma_c^2 + \sigma_{\varepsilon_i}^2$.

Systemic risk, labeled as R_i^2, is then modeled as the fraction of unit yield variation that can be explained by county yield variation,

$$R_i^2 = \frac{\mathrm{Var}\left[\beta_i\left(\widetilde{y}_c - \mu_c\right)\right]}{\mathrm{Var}\left(\widetilde{y}_i - \mu_i\right)} = \frac{\beta_i^2 \sigma_c^2}{\beta_i^2 \sigma_c^2 + \sigma_{\varepsilon_i}^2} = \frac{1}{1 + \sigma_{\varepsilon_i}^2 / \left(\beta_i^2 \sigma_c^2\right)}. \tag{2}$$

As $R_i^2 \in [0, 1]$, it provides a straightforward absolute measurement of systemic risk. Values $R_i^2 > 0.5$ indicate that systemic risk is the dominant yield risk source faced by the farmer so that AYP has the potential to remove the majority of total risk.

Equation (2) also shows that our systemic risk measure can be decomposed into three more fundamental components: (i) the square of unit yield's sensitivity to county yield, β_i^2; (ii) county yield variance, σ_c^2; and (iii) idiosyncratic yield variance, $\sigma_{\varepsilon_i}^2$. Proposition 1 provides simple inferences that can be extracted from the equation.

> **Proposition 1.** Ceteris paribus, R_i^2 increases with (i) an increase in unit yield's sensitivity to county yield, β_i; (ii) an increase in county yield variance, σ_c^2; and (iii) a decrease in idiosyncratic yield variance, $\sigma_{\varepsilon_i}^2$.

Proposition 1 provides indications on which counties and insurance units are likely to have high systemic risk. The most discernible is that, ceteris paribus, systemic risk will be higher in counties with larger county yield variances. We can readily identify these counties because county yield data can be accessed from the National Agricultural Statistical Service (NASS). In addition, although idiosyncratic yield variance is not observable, we might in general expect that units in regions where

[2]Here, we use "unit" as a general concept. There are several insurance unit types, such as optional unit and enterprise unit. See Bulut (2020) for more information.

heterogeneous approaches to production are taken are more likely to display large idiosyncratic yield variance and so low systemic risk. Were AYP the only insurance choice, farmers would tend to adopt practices common in an area to align their risks with area risk and obtain better protection from AYP (Chambers & Quiggin, 2002).

A common notion about area insurance is that it should work best in homogenous areas where yield correlations are strong (Barnett et al., 2005). We comment next on the relationship between systemic risk and the correlation coefficient between unit yield and county yield. Denoting $\theta_i = \mathrm{Cov}(\widetilde{y}_i, \widetilde{y}_c)/(\sigma_i \sigma_c)$ as the correlation coefficient between $\widetilde{y}_i$ and $\widetilde{y}_c$, we can rewrite the beta coefficient as $\beta_i = \mathrm{Cov}(\widetilde{y}_i, \widetilde{y}_c)/\mathrm{Var}(\widetilde{y}_c) = \theta_i \sigma_i \sigma_c / \sigma_c^2 = \theta_i \sigma_i / \sigma_c$.[3] Defining $\chi_i = \sigma_{\varepsilon_i}^2 / \sigma_i^2$ as the share of idiosyncratic yield variation in total unit yield variation, we obtain.

$$R_i^2 = \frac{1}{1 + (1/\theta_i^2)\chi_i}. \tag{3}$$

Proposition 2. Ceteris paribus, R_i^2 increases with (i) an increase in the correlation between unit yield and county yield, θ_i; and (ii) a decrease in the share of total unit yield variation that is accounted for by idiosyncratic yield variation, χ_i.

Proposition 2 implies that a strong unit-to-county yield correlation may not lead to high systemic risk when the share of idiosyncratic yield variation in total unit yield variation is large. Were county yield variation small but unit yield variation large then AYP only removes a small proportion of yield risk even when unit yield is highly correlated with county yield.

2.2 | Systemic risk and natural resource endowments

In this subsection we investigate how systemic risk varies with natural resource endowments. This information will allow us to develop insights on the roles of land quality and weather conditions in determining systemic risk and on how the risk protection function of area insurance will evolve in the presence of climate change.

Rather than directly model systemic risk as a function of natural resource endowment variables, labeled as Z_c, we allow each of the three systemic risk components to be a function of Z_c. The aggregate effect of natural resource endowments on systemic risk can be obtained as follows. Letting $\tau_i^2 = \sigma_{\varepsilon_i}^2 / (\beta_i^2 \sigma_c^2)$, then $R_i^2 = 1/(1+\tau_i^2)$ and $R_i^2/(1-R_i^2) = 1/\tau_i^2$. Taking the natural log of both sides of the last equation generates.

$$\ln\left(\frac{R_i^2}{1-R_i^2}\right) = -2\ln(\tau_i) = 2\ln[\beta_i(Z_c)] + 2\ln[\sigma_c(Z_c)] - 2\ln[\sigma_{\varepsilon_i}(Z_c)]. \tag{4}$$

Because the logistic transformation is monotonic, the effect of Z_c on R_i^2 is qualitatively the same as the effect of Z_c on $\ln\left[R_i^2/(1-R_i^2)\right]$. Equation (4) shows that the effects of natural resource endowments on systemic risk pass through $\beta_i(\cdot)$, $\sigma_c(\cdot)$, and $\sigma_{\varepsilon_i}(\cdot)$. Thus, if a given natural resource endowment variable has the same effect on $\beta_i(\cdot)$ and $\sigma_c(\cdot)$, then these two effects are concordant with each other. However, were a natural resource endowment variable to have the same effects on $\sigma_{\varepsilon_i}(\cdot)$ and either one of $\beta_i(\cdot)$ and $\sigma_c(\cdot)$, then these two effects would offset.

[3]See Part I of online Supplementary Appendix (Appendix S1) on how to derive $\beta_i = \theta_i \sigma_i / \sigma_c$. This expression is also derived in Miranda (1991). See its Equation (13), p. 235.

2.3 | Modeling YP and AYP contracts

Before introducing our definition of TRASR, we first outline how YP and AYP work. We choose Yield Protection and Area Yield Protection as, respectively, our generic YP and AYP plans. These are the two major yield insurance plans currently available in the United States.

Unit i in county c with random yield $\widetilde{y}_i$ and choosing to purchase YP will receive indemnity payments in the form,

$$\widetilde{n}_i = \max\left(\phi_i \overline{y}_i - \widetilde{y}_i, 0\right), \tag{5}$$

where $\widetilde{n}_i$ is realized YP indemnity payment, $\overline{y}_i$ is the insurer's reference "expected" unit yield, as established by Risk Management Agency (RMA),[4] and ϕ_i is the YP coverage level chosen by the farmer with $\phi_i \in \{0.5, \ldots, 0.85\}$ where evaluations are in 5% increments. Thus, YP pays indemnities whenever $\widetilde{y}_i$ falls below the policy protection amount, $\phi_i \overline{y}_i$.

Similarly, AYP pays indemnities whenever the county average yield is lower than the policy-protected county yield level. Compared with the YP indemnity function, the AYP indemnity function takes a more complicated form,

$$\widetilde{n}_c = \rho \max\left[\min\left(\frac{\overline{y}_c \phi_c - \widetilde{y}_c}{\phi_c - l}, \overline{y}_c\right), 0\right], \tag{6}$$

where $\widetilde{n}_c$ is the realized AYP indemnity payment, $\overline{y}_c$ is the reference expected county average yield, and ϕ_c is the AYP coverage level with $\phi_c \in \{0.7, \ldots, 0.9\}$ also available in 5% increments. Protection factor ρ, with $\rho \in [0.8, 1.2]$, allows the insured to adjust the AYP liability to better match expected individual losses. The loss limit factor, l, with fixed value 0.18, allows the insured's entire loss to be paid when the county loss equals 82% of the expected county yield, but no additional indemnity is payable when the county loss exceeds 82% of the expected county yield (see Bulut and Collins (2014), p. 415, note 1).

2.4 | Deriving TRASR

This subsection illustrates the importance of relative premium subsidy rates in encouraging farmers to choose AYP over YP. Because YP provides better risk protection than AYP, under the actuarially fair premium assumption, it follows that risk-averse farmers will always choose YP over AYP whenever no premium subsidy is provided.

To see this, let π_i and π_c denote, respectively, YP and AYP premiums, whereas s_i and s_c denote respective subsidy rates. Under the actuarial fairness assumption, that is, $\pi_i = \mathrm{E}(\widetilde{n}_i)$ and $\pi_c = \mathrm{E}(\widetilde{n}_c)$, then the expected net returns from purchasing YP and AYP are

$$\mathrm{E}\left(\widetilde{y}_i^i\right) = \mathrm{E}\left[\widetilde{y}_i + \widetilde{n}_i - (1 - s_i)\pi_i\right] = \mathrm{E}\left(\widetilde{y}_i\right) + s_i \mathrm{E}(\widetilde{n}_i) \text{ and}$$

$$\mathrm{E}\left(\widetilde{y}_i^c\right) = \mathrm{E}\left[\widetilde{y}_i + \widetilde{n}_c - (1 - s_c)\pi_c\right] = \mathrm{E}\left(\widetilde{y}_i\right) + s_c \mathrm{E}(\widetilde{n}_c), \tag{7}$$

respectively. Thus, when no premium subsidy is offered, that is, when $s_i = s_c = 0$, then the expected net returns from purchasing YP and AYP are the same and equal $\mathrm{E}\left(\widetilde{y}_i\right)$. Risk-averse farmers then will never choose AYP over YP as YP provides better risk protection.

[4]RMA establishes each unit's expected yield based on its actual production history (APH) yield, which is a simple average of 4–10 years of verified yield history on the insured unit. See Plastina and Edwards (2017) for more details on calculating APH yields. Also see Skees et al. (1997) for how RMA sets up county yield expectation.

Only when premium subsidies are introduced and the condition $s_c \mathrm{E}(\widetilde{n}_c) > s_i \mathrm{E}(\widetilde{n}_i)$ is satisfied will risk-averse farmers have incentives to choose AYP over YP. Then, the ratio

$$s^*_{c|i} = s_c / s_i = \mathrm{E}(\widetilde{n}_i)/\mathrm{E}(\widetilde{n}_c) \tag{8}$$

provides an intuitive lower bound on the relative subsidy rate at which a risk-averse or risk-neutral farmer will be indifferent between the two subsidized contracts. We call $s^*_{c|i}$ the threshold relative area subsidy rate (TRASR), the relative subsidy rate of AYP over YP at which the expected net returns from purchasing YP and AYP are the same.[5] Figure A1 in Part II of Appendix S1 illustrates how TRASR is related to a farmer's insurance choice. When the relative subsidy rate is below TRASR, then YP dominates AYP as YP provides better risk protection and higher expected net return; risk-averse or risk-neutral farmers should always prefer YP over AYP. When the relative subsidy rate is above TRASR, then AYP provides higher expected net returns and may dominate. In this case, farmers' insurance choices depend on their risk aversion levels.

By substituting in YP and AYP indemnity functions, we can develop an explicit form for $s^*_{c|i}$. First note that by assuming that yield expectations established by RMA perfectly match actual expectations, that is, $\overline{y}_i = \mu_i$ and $\overline{y}_c = \mu_c$, Equations (1) and (5) then jointly imply $\widetilde{n}_i = \max\left[\beta_i\left(\mu_c - \widetilde{y}_c\right) - \mu_i(1-\phi_i) - \varepsilon_i, 0\right]$. This new equation presents YP indemnities as a function of county yield and idiosyncratic yield. The YP indemnity function can be rewritten as

$$\widetilde{n}_i = \begin{cases} \beta_i\left(\mu_c - \widetilde{y}_c\right) - \mu_i(1-\phi_i) - \varepsilon_i, & \text{whenever } \widetilde{y}_c < \mathrm{M}_i(\varepsilon_i); \\ 0, & \text{whenever } \widetilde{y}_c \geq \mathrm{M}_i(\varepsilon_i), \end{cases} \tag{9}$$

where we use $\mathrm{M}_i(\varepsilon_i) = \mu_c - [\mu_i(1-\phi_i) + \varepsilon_i]/\beta_i$ to denote the trigger value of $\widetilde{y}_c$, given ε_i, below which YP pays strictly positive indemnities.

Similarly, the AYP indemnity function can be rewritten as

$$\widetilde{n}_c = \begin{cases} \mu_c \rho, & \text{whenever } \widetilde{y}_c < \mathrm{M}^{\mathrm{l}}_{\mathrm{c}}; \\ \alpha_c\left(\mu_c \phi_c - \widetilde{y}_c\right), & \text{whenever } \mathrm{M}^{\mathrm{l}}_{\mathrm{c}} \leq \widetilde{y}_c < \mathrm{M}_c; \\ 0, & \text{whenever } \widetilde{y}_c \geq \mathrm{M}_c, \end{cases} \tag{10}$$

where we use $\mathrm{M}_c = \mu_c \phi_c$ to denote the trigger value of $\widetilde{y}_c$ below which AYP pays strictly positive indemnities, and we use $\mathrm{M}^l_c = \mu_c l$ to denote the lower bound on $\widetilde{y}_c$ below which AYP always pays its maximum indemnity level, $\mu_c \rho$. We further use $\alpha_c = \rho/(\phi_c - l)$ to denote the inverse of the slope of the AYP indemnity function's middle component given in Equation (10). Because $0.8 \leq \rho \leq 1.2$, $0.7 \leq \phi_c \leq 0.9$ and $l = 0.18$, it follows that $\alpha_c \geq 0.8/0.72$, so $\alpha_c > 1$ always holds and α_c scales up AYP payments.

As demonstrated in Part III of Appendix S1, some transformations provide an explicit form for $s^*_{c|i}$:

$$s^*_{c|i} = \frac{\beta_i \int_{\underline{\varepsilon}_i}^{\overline{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F\left(\widetilde{y}_c\right) d\widetilde{y}_c dG(\varepsilon_i)}{\alpha_c \int_{\mathrm{M}^l_c}^{\mathrm{M}_c} F\left(\widetilde{y}_c\right) d\widetilde{y}_c}, \tag{11}$$

[5]TRASR also constitutes a lower bound on the relative subsidy rate needed to induce loss-averse farmers to choose AYP over YP. Loss aversion is found to be an important factor affecting agricultural producers' choices over risky outcomes (Feng et al., 2020; Liu, 2013; Lampe & Würtenberger, 2020). Losses of equal size outweigh gains for loss-averse farmers. Due to basis risk, loss-averse farmers would avoid AYP when YP provides the same expected net return as AYP.

where $F(\cdot)$ is the cumulative density function (CDF) for $\widetilde{y}_c$, $G(\cdot)$ is the CDF for ε_i, $\overline{\varepsilon}_i = y_i^u - \mu_i + \beta_i \mu_c$ is the upper bound on ε_i and $\underline{\varepsilon}_i = -\mu_i - \beta_i (y_c^u - \mu_c)$ is the lower bound on ε_i. Thus, $s_{c|i}^*$ is a function of two random variables, $\widetilde{y}_c$ and ε_i, as well as a parameter set. In the following subsection we seek to understand how $s_{c|i}^*$ is affected by systemic risk variables.

2.5 | TRASR and systemic risk

Although $s_{c|i}^*$ is not a direct function of R_i^2, it is a function of β_i. It also depends on metrics related to $\sigma_{\varepsilon_i}^2$ and σ_c^2 because both $\int_{\underline{\varepsilon}_i}^{\overline{\varepsilon}_i} \int_0^{M_i(\varepsilon_i)} F(\widetilde{y}_c) d\widetilde{y}_c dG(\varepsilon_i)$ and $\int_{M_c^l}^{M_c} F(\widetilde{y}_c) d\widetilde{y}_c$ are formed from integrations of CDFs where we can study the effect of an increase in $\sigma_{\varepsilon_i}^2$ or σ_c^2 in terms of a mean-preserving spread (m.p.s.) of $G(\varepsilon_i)$ or $F(\widetilde{y}_c)$. Although variance itself is widely used as a measure of risk, m.p.s. provides a more general way to study increased risk (Rothschild & Stiglitz, 1970). Proposition 3 provides relationships between $s_{c|i}^*$ and systemic risk components, where the mathematical proof can be found in Part IV of Appendix S1.[6]

Proposition 3. Ceteris paribus, $s_{c|i}^*$ increases with a m.p.s. in $G(\varepsilon_i)$. Signing the effects of a m.p.s. in $F(\widetilde{y}_c)$ and an increase in β_i on $s_{c|i}^*$ requires additional information or conditions. However, when $\varepsilon_i \equiv 0$ then $s_{c|i}^*$ is strictly increasing in β_i.

The finding that $s_{c|i}^*$ increases with a m.p.s. in $G(\varepsilon_i)$ is intuitive because under the assumption that no single unit has an impact on county yield, an increase in idiosyncratic risk increases $\mathrm{E}(\widetilde{n}_i)$ but has no effect on $\mathrm{E}(\widetilde{n}_c)$. By Equation (8), $s_{c|i}^*$ increases as a result.

As for the effect of a m.p.s. in $F(\widetilde{y}_c)$ on $s_{c|i}^*$, intuition might suggest that a m.p.s. in $F(\widetilde{y}_c)$ increases $\mathrm{E}(\widetilde{n}_c)$ because AYP pays more frequently as county yield risk increases. However, the existence of the loss limit factor complicates the effect of a m.p.s. in $F(\widetilde{y}_c)$ on $\mathrm{E}(\widetilde{n}_c)$. For example, if the left spread of a m.p.s. in $F(\widetilde{y}_c)$ works only on the range $[0, \mu_c l]$ and the right spread of the m.p.s. only shifts some $\widetilde{y}_c$ values smaller than $\mu_c \phi_c$ to be greater than $\mu_c \phi_c$, then $\mathrm{E}(\widetilde{n}_c)$ decreases with the m.p.s. operation. In addition, $\mathrm{E}(\widetilde{n}_i)$ increases with a m.p.s. in $F(\widetilde{y}_c)$ because unit yield risk increases with an increase in the county yield risk given that idiosyncratic yield risk is unaffected by county yield risk. Thus, even if $\mathrm{E}(\widetilde{n}_c)$ increases with a m.p.s. in $F(\widetilde{y}_c)$, determining the overall effects of a m.p.s. in $F(\widetilde{y}_c)$ on $s_{c|i}^*$ requires additional information or conditions. However, because the distribution of average unit yield approaches that of average county yield, it might be reasonable to expect that on average, a m.p.s. in $F(\widetilde{y}_c)$ has similar effects on $\mathrm{E}(\widetilde{n}_i)$ and $\mathrm{E}(\widetilde{n}_c)$, and so the overall effect on $s_{c|i}^*$ would then be small. Conjecture 1 follows.

Conjecture 1. On average, the effect of a m.p.s. in $F(\widetilde{y}_c)$ on $s_{c|i}^*$ is small.

Finally, although an increase in β_i only affects $\mathrm{E}(\widetilde{n}_i)$, the effect on TRASR is ambiguous. Intuition suggests that $s_{c|i}^*$ increases with an increase in β_i because as unit yield becomes more sensitive to county yield, unit yield variation increases for a given level of county yield variation. As a result, $\mathrm{E}(\widetilde{n}_i)$ increases while $\mathrm{E}(\widetilde{n}_c)$ does not change. This is exactly the case when $\varepsilon_i \equiv 0$, that is, when there is no idiosyncratic risk. In this case, $\widetilde{y}_i - \mu_i$ is solely determined by $\widetilde{y}_c - \mu_c$ and always has the sign of $\widetilde{y}_c - \mu_c$. Given $F(\widetilde{y}_c)$, an increase in β_i thus increases the probability of receiving a YP payment but has no effect on the probability of receiving an AYP payment. But suppose instead that $\varepsilon_i \not\equiv 0$. In this case, when $\widetilde{y}_i - \mu_i < 0$, $\varepsilon_i < 0$ and $\widetilde{y}_c - \mu_c > 0$, then $\widetilde{y}_i - \mu_i$ will be less negative as β_i increases, which in turn means fewer YP indemnities (if any). As a result, $\mathrm{E}(\widetilde{n}_i)$ may decrease with an increase in β_i. However, we should expect $s_{c|i}^*$ to increase with an increase in β_i when the lower bound of ε_i, $\underline{\varepsilon}_i$, is

[6]In Part V of Appendix S1 we also present a comparative statics analysis of $s_{c|i}^*$ with respective to other variables and parameters appearing in Equation (11).

sufficiently large (less negative) such that the YP indemnity is never triggered when all of $\widetilde{y}_i - \mu_i < 0$, $\varepsilon_i < 0$ and $\widetilde{y}_c - \mu_c > 0$ apply. Equation (A3) in Appendix S1, Part III, provides us with such a condition, namely $-\underline{\varepsilon}_i \leq \mu_i(1-\phi_i)$, where $-\underline{\varepsilon}_i$ can be viewed as the worst possible idiosyncratic yield loss and $\mu_i(1-\phi_i)$ is the YP deductible. [7]

Proposition 4. Ceteris paribus, $s^*_{c|i}$ increases with an increase in β_i whenever $-\underline{\varepsilon}_i \leq \mu_i(1-\phi_i)$.

The condition in Proposition 4 is likely to hold because the YP deductible is included to remove moral hazard, which is reflected in idiosyncratic yield loss. Moreover, deep losses in crop yields are generally caused by systemic weather events, and crop insurance pays less frequently in good weather years where idiosyncratic yield risk dominates.

Proposition 4 has important implications. Because R_i^2 also increases with an increase in β_i, then an increase in β_i leads to increases in both $s^*_{c|i}$ and R_i^2 whenever $-\underline{\varepsilon}_i \leq \mu_i(1-\phi_i)$ applies. Thus, under the mild condition that the YP deductible is greater than the worst possible idiosyncratic yield loss, an increase in unit yield's sensitivity to county yield has two opposing effects on farmers' choices of AYP: farmers enjoy better risk protection from AYP; they also receive more YP payments as the increased sensitivity amplifies the effect of county yield risk on total unit yield risk. This finding contrasts with predictions given in previous studies, where both individual loss and area loss are exogenously given and only an artificial correlation level is allowed to vary. When individual loss level is fixed, an increase in yield correlation always favors AYP as area yield improves at tracking individual yield.

The following remark, a corollary to Proposition 4, sheds lights on conditions under which Proposition 4 is more likely to hold.

Remark 1. Ceteris paribus, $s^*_{c|i}$ is more likely to increase with an increase in β_i for (i) a higher μ_i, (ii) a smaller ϕ_i, and (iii) a greater value of $\underline{\varepsilon}_i$.

Claims (i) and (ii) indicate that, ceteris paribus, farmers operating high-yield land and choosing lower YP coverage levels are more likely to find their $s^*_{c|i}$ values increase with an increase in β_i. However, farmers operating productive land, as in Iowa and Illinois, generally choose high coverage levels. They also generally encounter smaller worst possible idiosyncratic yield loss as they plant on more shock-resilient lands.

Claim (iii) provides clues about the relationship between $s^*_{c|i}$ and R_i^2 across units. Because a mean-preserving contraction of $G(\varepsilon_i)$ is generally associated with an increase in $\underline{\varepsilon}_i$, Claim (iii) implies that units with less dispersed idiosyncratic yield distributions are more likely to see an increase in β_i have a positive effect on $s^*_{c|i}$. Thus, units with less dispersed idiosyncratic yield distributions and larger β_i values tend to have larger $s^*_{c|i}$ values. Because county yield risk is the same for all units in a county, Proposition 1 then implies that within a county, units with larger R_i^2 values tend to have less dispersed idiosyncratic yield distributions and larger β_i values, and thus larger $s^*_{c|i}$ values. This result indicates that farmers who enjoy good risk protection from AYP will be more likely to find current AYP subsidy rates financially unrewarding. It is then unsurprising to see low AYP take-up rates.

3 | DATA

Unit corn yield data have been obtained from the 2008 unit-level RMA records. Locations of these unit-level data are known to the county level only. The data are sequences of 4–10-year yield historical yields used to establish APH yields. Our sample consists of 213,429 unirrigated units with

[7]Note that $-\underline{\varepsilon}_i \leq \mu_i(1-\phi_i)$ is not a constraint built into LAM and does not apply to all units we investigate in this paper.

10 actual yield records from 1998 through 2007 in 580 counties where the irrigation rate does not exceed 20%.

County average corn yield data are from NASS.[8] We consider only counties in 12 major corn production states in the Midwest and Great Plains: IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, SD, and WI. To obtain sufficient observations to estimate county yield trend, we only keep counties with 50 consecutive years of observations from 1958 through 2007. To ensure that each county's unit set is representative, we also drop counties with <30 units. See Part VI in Appendix S1 for details on the data screening process.

We use data from the National Oceanic and Atmospheric Administration (NOAA) to construct two sets of county-level weather variables. The first set contains growing degree days (GDD) and stress degree days (SDD). These variables are widely used to measure heat conditions (Du et al., 2015; Schlenker & Roberts, 2006, 2009). For each county, the GDD variable, labeled as G_c, measures the 10-year (1998–2007) average growing season (May to August) accumulation of beneficial degrees in the [10°C, 30°C] range (Neild & Newman, 1987). The SDD variable, labeled as S_c, measures the 10-year average growing season accumulation of degrees exceeding 32.22°C (Schlenker & Roberts, 2009). The second climate variable set measures the relative moisture in a county. We use the Palmer's Z (PZ) index to measure the departure of monthly weather from average moisture conditions. PZ values within the range [−2, 2.5] are viewed as normal, whereas values below −2 indicate severe drought and values above 2.5 indicate severe wetness. We construct a drought variable, D_c, to measure the number of months where severe drought occurred in county c over 1998–2007, and a wetness variable, W_c, to measure the number of months where severe wetness occurred in county c over the same period. Details on constructions of weather variables are provided in Appendix S1, Part VII.

Land quality data are from the National Resource Conservation Service. County land quality, labeled as L_c, is defined as the percent of all land that is in land capability classes (LCC) I or II in that county. There are eight LCCs in total, where classes I and II are most favorable for cultivation, whereas classes III or higher have gradually more severe limitations for cultivation. Land quality has little variation across years, and we use the 2010 measure to construct L_c.

In Figure 2 we map the geographic distributions of weather and land quality variables while descriptive statistics can be found in Table A3 in Appendix S1, Part VII. Panels a and b show that GDD and SDD generally increase as one moves south. SDD also increases as one moves west. Panel c shows that counties in the Western and Eastern Corn Belts generally experienced more severe drought, whereas Panel d shows that counties in the Northern Corn Belt generally experienced more severe wetness. Counties with higher fractions of good land, as shown by Panel e, are mainly located in Iowa, Illinois, Indiana, and Western Ohio.

4 | EMPIRICAL RESULTS

In this section we empirically estimate systemic risk and calibrate $s^*_{c|i}$. Propositions and conjectures derived in the conceptual framework section are also examined.

4.1 | Measuring systemic risk

We first adapt Equation (1) to obtain empirical systemic risk estimates with

$$\widetilde{y}_{i,t} - \mu_{i,t} = \alpha_i + \beta_i\left(\widetilde{y}_{c,t} - \mu_{c,t}\right) + \varepsilon_{i,t}, \tag{12}$$

[8]We also conduct an analysis where the NASS county yield is replaced by the acre-weighted RMA county average. These results are reported in Part IX in Appendix S1 and are very similar to results reported in the main context, to follow.

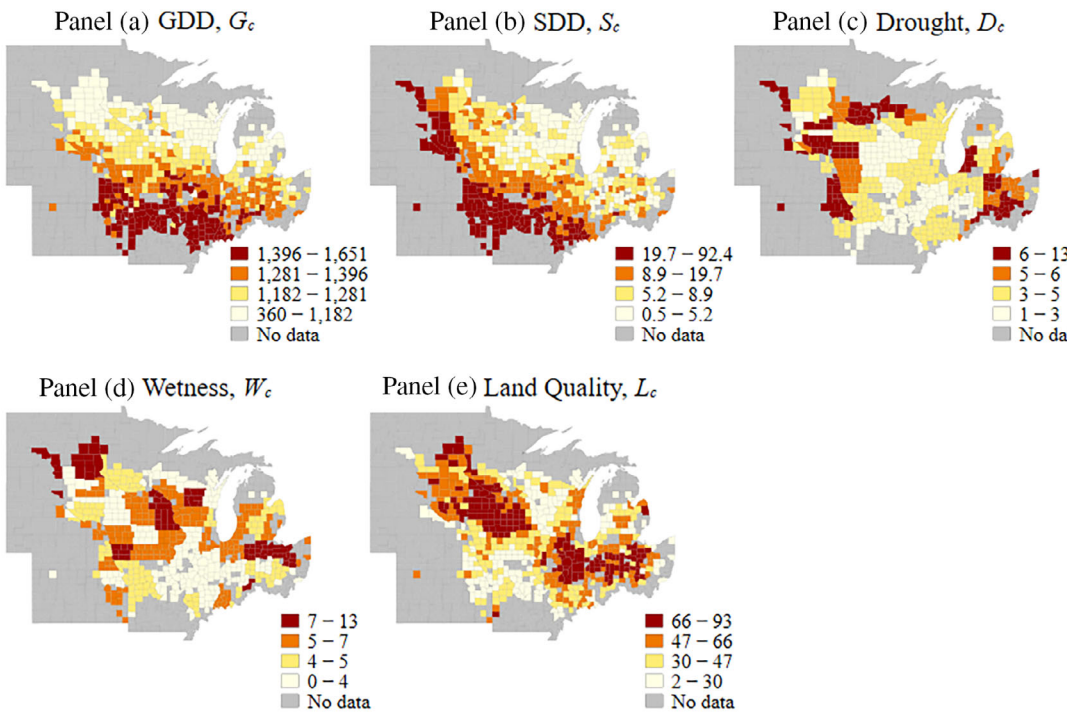

FIGURE 2 Geographic distributions of natural resource endowments *Notes*: 1. This figure plots geographic distributions of county GDD, county SDD, frequencies of severe drought and severe wetness, and the percent of county land that is in either land capability class I or class II. Numbers in legends are quartile ranges. 2. Weather variable frequencies are monthly observations over 1998–2007 out of the 120-month interval.

where α_i is the unit-specific constant term. Unit-level systemic risk, R_i^2, is then measured by the R^2 statistic when Equation (12) is estimated by ordinary least squares (OLS), that is, $R_i^2 = \beta_i^2 \mathrm{Var}(\widetilde{y}_{c,t} - \mu_{c,t}) / \mathrm{Var}(\widetilde{y}_{i,t} - \mu_{i,t})$. We follow Deng et al. (2007) by applying a log-linear model to detrend yield data as well as to establish unit APH yields and county yield expectations. We also follow RMA by adjusting up APH yields as APH yields lag true expected yields due to improved crop genetics and cultural practices (Plastina & Edwards, 2014). Part VIII in Appendix S1 provides more details on the detrending and APH adjustment methods.

We estimate Equation (12) and decompose the resulting systemic risk measure according to Equation (4). Table 1 presents the descriptive statistics for unit-level systemic risk and its three components. Mean R_i^2, at 0.46, suggests that systemic risk generally explains slightly less than half of total unit yield variation. Thus, on average, the risk reduction potential of AYP is limited, which might be an important reason for the low area insurance take-up rate. But the magnitude of systemic risk varies considerably across units, ranging from near 0 to near 1. The mean β_i estimate is 1.04 with range – 4.78 to 9.19.[9] Only about 2.6% of β_i estimates are negative. The means of σ_{ε_i} and σ_c are 18.3 and 18.2, respectively.

Figure 3 plots the geographic distributions of county average systemic risk and county averages of systemic risk components. Panel a shows that systemic risk is highest at the southern and western

[9]The 1.04 mean deviates slightly from Miranda's (1991) assertion that acre-weighted average β_i within a given county should equal to 1. This might be because our county yield data are from NASS whereas unit yield data are from RMA. The NASS county yield generally does not equal the mean of the acre-weighted RMA unit yield (Zulauf et al., 2017). In addition, by dropping units in our data screening process, we have further loosened the connection between county yield and unit yield.

TABLE 1 Descriptive statistics for estimates of systemic risk and systemic risk components

Variable	N	Mean	St. dev	Min	Max
R_i^2	213,429	0.46	0.25	0.00	1.00
β_i	213,429	1.04	0.59	−4.78	9.19
σ_{ε_i}	213,429	18.3	9	0.85	95.02
σ_c	580	18.2	6.32	6.64	43.5

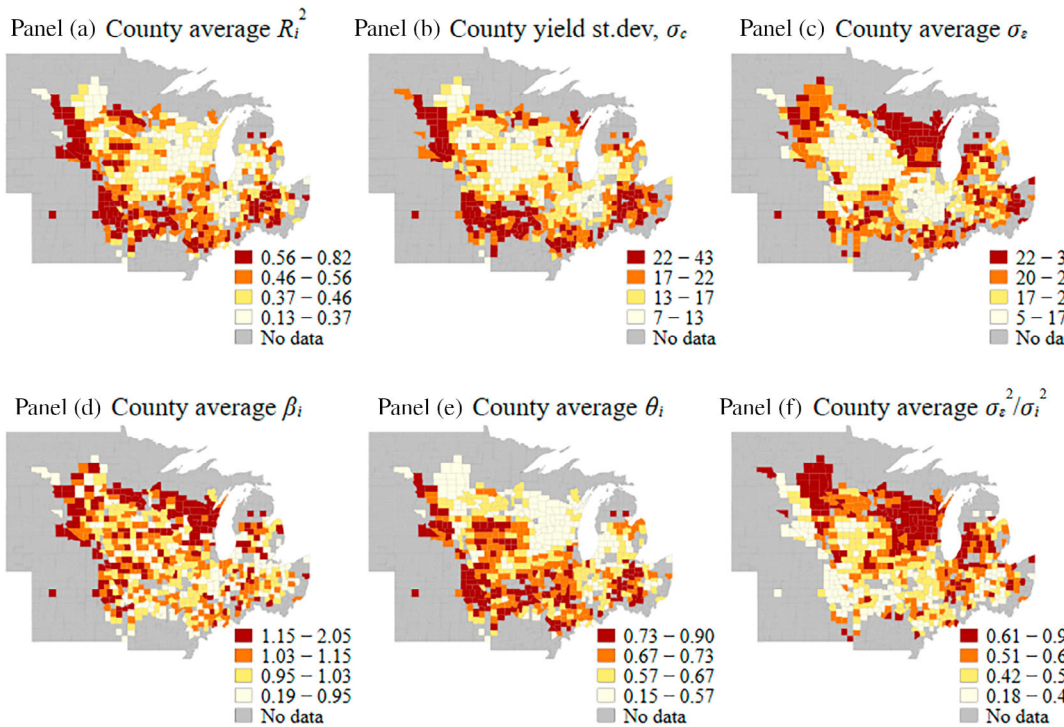

FIGURE 3 Geographic distributions of county average systemic risk and county averages of systemic risk components
Note: Panels a to f respectively plot county averages of unit-level systemic risk, county yield standard deviation, unit-level idiosyncratic yield standard deviation, unit yield's sensitivity to county yield, the correlation coefficient between unit yield and county yield, and the share of idiosyncratic yield variation in total unit yield variation. Numbers in legends are quartile ranges.

fringes of the Corn Belt, but is comparatively low in much of Iowa and counties to its northeast. Intuition, however, might suggest that AYP would work best in the latter areas, which have homogenous natural resource endowments and strong yield correlations. This incongruity arises because, as is shown in Panel b, these areas also have small county yield variances. When discussing Proposition 2, Claim (ii), we commented that AYP is ineffective in removing yield risk when county yield variance accounts for only a small part of total unit yield variance. Panels e and f of Figure 3 show that although Iowa has a relatively strong correlation between unit yield and county yield, it has a moderate share of idiosyncratic yield variation in total unit yield variation. On the contrary, counties in the southern and western fringes of the Corn Belt are characterized by both strong yield correlations and small shares of idiosyncratic yield variation. This mismatch between major corn production areas and high systemic risk areas may also contribute to area insurance's low overall take-up rate.

Panel c of Figure 3 shows that counties at the Corn Belt periphery have relatively large idiosyncratic yield variances, especially counties in Wisconsin and the Eastern Dakotas. On the contrary,

counties in Iowa and Illinois have low idiosyncratic yield variances. Panel d presents evidence that Corn Belt fringe counties have relatively large unit yield sensitivities to county yield, but the pattern is not as evident because some counties in the central part also have large yield sensitivities, whereas counties at the southern fringe have small yield sensitivities.

4.2 | The effects of farmers' reactions to yield risks on systemic risk

Chambers and Quiggin (2002) criticizes LAM for characterizing unit yield as a stochastic variable not subject to farmers' control. It is true that LAM does not explicitly model farmers' behaviors. However, as pointed out by Ramaswami and Roe (2004) on its page 422, "if the LAM is derived from the aggregation of individual technologies, then its parameters can be seen to be functions of individual choice variables. The criticism of Chambers and Quiggin will then no longer apply." An implication of the argument in Ramaswami and Roe (2004) is that different reactions to yield risks will lead to different estimates of LAM parameters and thus different estimates of systemic risk. We find evidence for this conjecture from a sample of units in counties where irrigation rates are no no less than 20%.

Figure 4 plots the respective average systemic risks of irrigated and unirrigated units with respect to county irrigation rate. The average systemic risk of unirrigated units, denoted as $\overline{R}^2_{uir}$, is above 0.5 in counties where irrigation rates are less than 30% whereas the average systemic risk of irrigated units, denoted as $\overline{R}^2_{ir}$, is below 0.2 in these counties. As county irrigation rate increases, $\overline{R}^2_{uir}$ decreases and $\overline{R}^2_{ir}$ increases. These results suggest that in less irrigated counties, yields of irrigated units are detached from yields of the more prevalent unirrigated units, leading to low systemic risk. In heavily irrigated counties, yields of non-irrigated units are detached from those of the more prevalent irrigated units. AYP is unlikely to work well for irrigated units in rainfed dominant counties and unirrigated units in irrigation-intensive counties.

4.3 | The effects of natural resource endowments on systemic risk

To study how natural resource endowments determine systemic risk, we regress each of the three systemic risk components on county weather and land quality variables. The following unit-level log-linear model is estimated by OLS,

$$\ln(X_i) = \delta_0 + \delta_1 G_c + \delta_2 S_c + \delta_3 D_c + \delta_4 W_c + \delta_5 L_c + \delta_6 APH_i + v_i, \tag{13}$$

where $X_i \in \left\{\beta_i^2, \sigma_{\varepsilon_i}^2, \sigma_c^2\right\}$ is a $n \times 1$ vector, G_c, S_c, D_c, W_c, L_c are $n \times 1$ natural resource endowment variables as defined before, v_i is a $n \times 1$ error term, and $n = 213,429$. The variable APH_i represents each unit's 2008 APH yield. We include this variable to control for land heterogeneity within a county. If land heterogeneity is correlated with natural resource endowment variables, then omitting it will lead to biased estimates. Because we do not have unit-level land quality data, we use each unit's 2008 APH yield as a proxy variable. Although weather conditions and unit land quality jointly determine APH yield, the variation remaining in APH is presumably primarily related to unit land qualities after controlling for weather variables. Another concern is that, because county yield variance and natural resource endowment variables are at county level, the unit-level regression of $\ln\left(\sigma_c^2\right)$ on natural resource endowment variables has many duplicated observations. For this reason, we also run a county-level regression of $\ln\left(\sigma_c^2\right)$ on natural resource endowment variables, that is, $\ln\left(\sigma_c^2\right)$ and all natural resource endowment variables are 580×1 vectors, whereas each observation represents a county.

Regression results for Equation (13) appear in Table 2. Column (1) shows that GDD has a significant negative effect on β_i, whereas SDD has a significantly positive effect. Thus, more excessive heat increases unit yield's sensitivity to county yield, whereas more beneficial heat decreases the

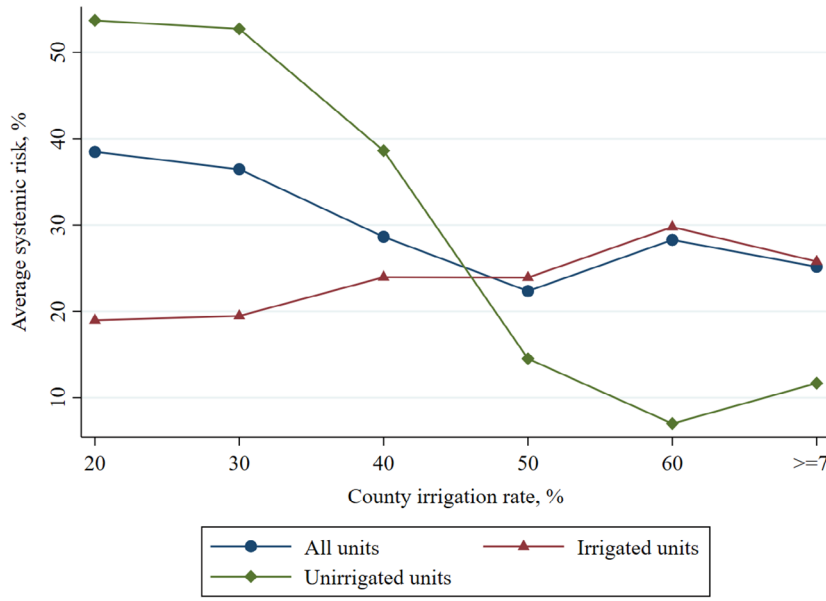

FIGURE 4 Average systemic risks of irrigated and unirrigated units with respect to county irrigation rate *Note*: Irrigation rate is at the county level and equals the share of irrigated units among all units in a county.

sensitivity. Drought occurrence significantly increases unit yield's sensitivity to county yield, whereas wetness has the converse effect. Better unit land tends to decrease unit yield's sensitivity. Overall, results in Column (1) show that β_i increases as climate conditions and land quality deteriorate. The marginally negative effect of wetness constitutes a counterexample to the previous statement, an outcome that may be because the adverse effects of severe wetness are localized and depend on whether the land unit is flood prone.[10]

Column (2) shows that GDD has a significantly positive effect on $\sigma^2_{\varepsilon_i}$, but SDD has a significantly negative effect. Thus, yield risk is less influenced by idiosyncratic factors in years with excessive heat. Drought occurrence and wetness occurrence have no significant effects on $\sigma^2_{\varepsilon_i}$. An increase in land quality significantly decreases $\sigma^2_{\varepsilon_i}$. Because better land also reduces σ^2_c, as shown by Columns (3) and (4), we may expect better land to play a buffering role in reducing the effects of extreme climate on crop yields.[11] Column (3) and column (4) also show that more heat accumulations and drought incidence tend to increase σ^2_c. Wetness incidence tends to reduce σ^2_c. County yield variability thus decreases as overall moisture conditions improve.

Column (5) reports the aggregate effects of natural resource endowments on systemic risk as obtained by the OLS regression of $-2\ln(\tau_i)$ on natural resource endowment variables. One can verify that, consistent with Equation (4), the coefficient of a natural resource endowment variable on $-2\ln(\tau_i)$ equals the sum of the coefficients of that variable on $2\ln(\beta_i)$ and $2\ln(\sigma_c)$ less the coefficient on $2\ln(\sigma_{\varepsilon_i})$. Overall, systemic risk increases with SDD and drought but decreases with wetness. Systemic risk also increases with unit land quality, suggesting that the positive effect of land quality in reducing idiosyncratic yield variance dominates its negative effects in reducing unit yield's sensitivity to county yield and county yield variance (please refer to Equation 4). However, our

[10]Specifically, and as with irrigation, flooding separates units into two groups that correlate differently with county yield. In counties where flooding affects a large area, affected units should follow county yield closely, but in counties where flooding only affects a small area, unaffected units should follow county yield closely. Overall, flooding likely reduces yield correlations. We also report regression results for Equation (13) but replace the wetness and drought variables with the county average precipitation variable, see Table A5 in Appendix S1, Part X.

[11]See Du et al. (2018) for the interaction effects between land quality and climate conditions.

TABLE 2 Regression results of systemic risk variables on natural resource endowments variables

	(1) $2\ln(\beta_i)$	(2) $2\ln(\sigma_{\varepsilon_i})$	(3) $2\ln(\sigma_c)$	(4) $2\ln(\sigma_c)$	(5) $-2\ln(\tau_i)$
$G_c/100$	-0.039^b	0.063^a	0.101^a	0.061^a	-0.001
	(0.017)	(0.012)	(0.026)	(0.022)	(0.034)
$S_c/10$	0.053^a	-0.144^a	0.110^a	0.140^a	0.307^a
	(0.019)	(0.019)	(0.035)	(0.028)	(0.044)
D_c	0.039^a	-0.007	0.093^a	0.101^a	0.139^a
	(0.009)	(0.007)	(0.014)	(0.014)	(0.017)
W_c	-0.016^c	-0.007	-0.040^a	-0.057^a	-0.050^a
	(0.008)	(0.007)	(0.014)	(0.012)	(0.016)
$L_c(10\%)$	0.001	-0.049^a	-0.024^c	-0.028^b	0.026
	(0.010)	(0.006)	(0.014)	(0.012)	(0.018)
APH_i	-0.003^a	-0.012^a	-0.007^a		0.002^b
	(0.001)	(0.001)	(0.001)		(0.001)
APH_c_CV				1.007^a	
				(0.348)	
Constant	0.514^b	7.085^a	4.948^a	4.540^a	-1.623^a
	(0.230)	(0.209)	(0.329)	(0.311)	(0.399)
Observations	213,429	213,429	213,429	580	213,429
R^2	0.012	0.111	0.441	0.351	0.076

Notes: 1. Column (3) reports OLS estimation results of $2\ln(\sigma_c)$ on natural resource endowment variables at the unit level of analysis where one observation represents one unit. Column (4) reports OLS estimation results of $2\ln(\sigma_c)$ on natural resource endowment variables at the county level where one observation represents one county. 2. The variable APH_c_CV denotes the coefficient of variation of units' 2008 APH values in county c, which captures the dispersion of land quality in a county and thus reflects land heterogeneity within a county. 3. Standard errors in parentheses are clustered at the county level; [a], [b], and [c] denote significance at 0.01, 0.05, and 0.1 levels, respectively.

investigations of systemic risk are subject to the time period for the data available. Because no catastrophic weather events for national corn production occurred during our sample years, our estimates of systemic risk and of the relationship between systemic risk and natural resource endowments are better understood as normal year outcomes.[12]

In addition to signs and significances, we also investigate the importance of each natural resource endowment variable in determining systemic risk. Following Huettner and Sunder (2012), we use the Shapley value to measure the "bargaining" power of each natural resource endowment variable in claiming explanation for the variation component of systemic risk. Their idea is to remove each explanatory variable from all possible combinations of other explanatory variables and so observe the variable's average contribution to the R^2. Column (1) of Table 3 shows that β_i is the most important determinant of systemic risk as it accounts for about 66% of systemic risk variation. County yield variance and idiosyncratic yield variance explain about 14% and 20% of systemic risk variation, respectively.

Results in Columns (2) through (4) of Table 3 show that natural resource endowment variables, drought occurrence, SDD, and unit land quality are the most important factors in explaining variations of systemic risk components. However, as shown in Column (5), the overall effect of unit land quality on systemic risk is small, but the overall effects of SDD and drought occurrence on systemic risk remain large. This is because, as shown in Table 2, the effect of unit land quality on $\sigma^2_{\varepsilon_i}$ is offset

[12]See Figure 2 at https://www.agry.purdue.edu/ext/corn/news/timeless/YieldTrends.html. Our data sit between the drought years in the 1980s and the 2012 drought year where the flood years in the early 1990s were more geographically confined and had more moderate impacts on higher ground even where they decimated crops on lower ground.

TABLE 3 Shapley values of natural resource endowment variables in explaining systemic risk variables, %

	(1) $-2\ln(\tau_i)$	(2) $2\ln(\beta_i)$	(3) $2\ln(\sigma_{\varepsilon_i})$	(4) $2\ln(\sigma_c)$	(5) $-2\ln(\tau_i)$
$2\ln(\beta_i)$	65.9				
$2\ln(\sigma_{\varepsilon_i})$	19.8				
$2\ln(\sigma_c)$	14.3				
G_c		4.7	1.5	13.3	9.2
S_c		17.6	5.9	27.6	39.5
D_c		35.9	1.8	23.9	38.2
W_c		4.2	0.5	3.7	7.1
L_c		2.8	17.7	7.2	1.8
APH_i		34.7	72.6	24.3	4.3

by its effects on β_i and σ_c^2 upon aggregation. On the contrary, the effects of SDD and drought occurrence on the three systemic risk components are concordant, and so the overall effects are larger.

4.4 | Calibrating TRASR

We turn now to calibrating $s^*_{c|i}$ and then investigating the relationship between $s^*_{c|i}$ and R_i^2. To calibrate $s^*_{c|i}$ from Equation (11), we need to know both $F(\widetilde{y}_c)$ and $G(\varepsilon_i)$.[13] Studies investigating crop yield distributions mainly adopt two distinct methodologies for estimating distributions, parametric and nonparametric. Parametric methods often assume that crop yield follows a specific distribution, such as the normal, gamma, or beta distribution (Gallagher, 1987; Harri et al., 2011; Sherrick et al., 2004). Nonparametric methods offer flexibility in capturing local yield idiosyncrasies (Goodwin & Ker, 1998; Ker & Goodwin, 2000). Because our study contains many counties and units where appropriate corn yield distribution specifications might differ by location, flexibility considerations lead us to choose nonparametric kernel density estimation. Details regarding how kernel estimations are performed at both county level and unit level are reported in Part XI of Appendix S1.

Table 4 presents descriptive statistics for coverage-level conditional $s^*_{c|i}$. We set $\rho = 1.2$ because 1.2 is the protection factor level that maximizes $\text{E}(\widetilde{n}_c)$. Because ϕ_i increases in 0.05 increments from 0.5 to 0.85 and ϕ_c increases in 0.05 increments from 0.70 to 0.9, we have $8 \times 5 = 40$ possible coverage level combinations. We focus on $s^*_{c|i}$ evaluations where $\phi_i \geq 75\%$ because 75% is the minimum coverage level chosen by most farmers in the Corn Belt (Schnitkey & Sherrick, 2014). Moreover, because some units have negative β_i estimates and thus negative $s^*_{c|i}$ values, and some units have extremely large $s^*_{c|i}$ values, we focus on the sample with positive β_i estimates and on median statistics.

Table 4 shows that the median $s^*_{c|i}$ value increases with an increase in ϕ_i and decreases with an increase in ϕ_c. When $\phi_i \geq 80\%$ and when $\phi_c \leq 80\%$ then median $s^*_{c|i}$ values generally exceed 100%, whereas when $\phi_c \geq 85\%$ then all $s^*_{c|i}$ median values are smaller than 100%. Thus, to coax more farmers into choosing AYP, low coverage level AYP subsidy rates must exceed YP subsidy rates, but high coverage level AYP subsidy rates may not exceed YP subsidy rates.

We then compare $s^*_{c|i}$ with $s_{c|i}$, which denotes the current ratio of AYP subsidy rate over YP subsidy rate, to investigate whether the current AYP subsidy rates discourage most farmers from choosing AYP over YP. If $s^*_{c|i}$ exceeds $s_{c|i}$, then a risk-averse or risk-neutral farmer will never choose AYP over YP because YP provides better risk protection and higher expected subsidy transfers. Table A7 in Appendix S1, Part XII, lists premium subsidy rates for individual and area insurance plans before and

[13]In light of Equation (8), a reasonable alternative approach to calculating TRASR values is by using premium data. In Part XIII of Appendix S1 we present such an analysis as a robustness check on methods used in the main text, to follow.

TABLE 4 Descriptive statistics for $s^*_{c|i}$ multiplied by 100 (measured in percent)

ϕ_i	ϕ_c	N	Mean	St. dev	Min	Median	Max
75%	70%	207,230	207	344	0	124	13,992
	75%	207,230	133	160	0	90	3843
	80%	207,230	92	97	0	66	2231
	85%	207,230	66	65	0	49	1463
	90%	207,230	48	46	0	37	1076
80%	70%	207,230	263	409	0	166	16,023
	75%	207,230	169	185	0	121	4074
	80%	207,230	117	111	0	89	2365
	85%	207,230	83	73	0	66	1515
	90%	207,230	61	51	0	49	1115
85%	70%	207,230	336	486	0	221	18,513
	75%	207,230	215	214	0	161	4313
	80%	207,230	148	126	0	118	2504
	85%	207,230	106	82	0	88	1569
	90%	207,230	77	57	0	66	1154

Notes: 1. Only the sample of units with $\beta_i > 0$ are considered here. 2. Units with extremely large $s^*_{c|i}$ values are most likely to be those with large idiosyncratic risk and in counties where county yield distributions have long and thin left tails.

after the 2008 Farm Bill. We choose the enterprise unit (EU) contract subsidy rate to derive $s_{c|i}$ because in recent years EU contracts have covered the largest share of insured corn acres (Bulut, 2020; Coble, 2017). Moreover, farmers may treat EU contracts as weak substitutes for AYP contracts because the former aggregate over several units in an area and so pool risks from a set of individual units.[14] Table A7 shows that $s_{c|i} < 1$ across all coverage levels except 85% in the post-2009 period, but $s_{c|i} > 1$ for all coverage levels in the pre-2009 period, indicating that the subsidy rate change in the 2008 Farm Bill favors EU contracts over AYP contracts.

Panel A of Table 5 reports the percent of units with $s^*_{c|i} > s_{c|i}$. When $\phi_c \leq 80\%$, $s^*_{c|i}$ exceeds $s_{c|i}$ for the majority of units. Thus, the current subsidy schedule does not encourage low coverage levels with AYP. In addition, federal legislation suggests that such low coverage levels are undesirable policy outcomes because they leave farmers exposed to basis risk. When $\phi_c > 80\%$, 24% to 40% of units continue to have $s^*_{c|i}$ values above their $s_{c|i}$ values, suggesting that the current AYP subsidy schedule also deters a significant fraction of farmers from choosing high coverage level AYP contracts over YP contracts. Moreover, the smaller percent of units with $s^*_{c|i} > s_{c|i}$ at high AYP coverage levels suggests that were AYP chosen then most likely the chosen level would exceed 80%. This conjecture matches with farmers' choices. Using the RMA Summary of Business (SOB) data, we find that among all corn AYP contracts sold over 1997–2019, about 86% are at $\phi_c \geq 80\%$ and 64% are at $\phi_c = 90\%$.

4.5 | TRASR, systemic risk, and area insurance demand

We turn now to investigating the relationship between $s^*_{c|i}$ and systemic risk variables together with how this relationship affects AYP demand. As Equation (11) suggests that any such relationship is nonlinear, we use Spearman's rank correlation test to check for correlations.

[14]However, our data were obtained at a time when EU share was low (pre-2009 policy change). Because yield aggregation under EU reduces idiosyncratic risk and Proposition 3 suggests that $s^*_{c|i}$ decreases with a decrease in idiosyncratic risk, our $s^*_{c|i}$ estimates may be larger than would be the case were data under the current EU share used.

TABLE 5 The percent of sample units with $s^*_{c|i} > s_{c|i}$ and the percent of sample units with $s^*_{c|i} > \bar{s}_{c|i}$

ϕ_i	ϕ_c 70%	75%	80%	85%	90%
	Panel A: Percent of units with $s^*_{c\|i} > s_{c\|i}$, post-2009 subsidy rates				
75%	**68.0**	**56.8**	46.6	33.2	24.0
80%	**75.2**	**64.8**	**54.6**	39.4	28.5
85%	**78.8**	**68.2**	**57.3**	39.9	27.3
	Panel B: Percent of units with $s^*_{c\|i} > s_{c\|i}$, pre-2009 subsidy rates				
75%	**52.7** (−15.3)	38.9 (−17.9)	28.8 (−17.8)	17.0 (−16.2)	10.5 (−13.5)
80%	**59.7** (−15.5)	45.1 (−19.6)	33.7 (−20.9)	19.6 (−19.9)	11.6 (−16.9)
85%	**63.0** (−15.8)	47.4 (−20.8)	34.6 (−22.7)	18.7 (−21.3)	10.0 (−17.2)
	Panel C: Percent of units with $s^*_{c\|i} > \bar{s}_{c\|i}$				
75%	48.2 (−4.6)	34.2 (−4.7)	21.4 (−7.5)	11.5 (−5.5)	5.5 (−5.0)
80%	**55.5** (−4.2)	40.5 (−4.7)	25.5 (−8.3)	13.3 (−6.2)	6.0 (−5.6)
85%	**57.7** (−5.3)	41.5 (−5.9)	24.6 (−9.9)	11.6 (−7.0)	4.5 (−5.5)

Notes: 1. Values in bold are >50%. 2. Values in parentheses in Panel B are pairwise differences between the percent of sample units with $s^*_{c|i} > s_{c|i}$ in the post-2009 period and the percent of sample units with $s^*_{c|i} > s_{c|i}$ in the pre-2009 period. 3. Values in parentheses in Panel C are pairwise differences between the percent of sample units with $s^*_{c|i} > s_{c|i}$ in the pre-2009 period and the percent of sample units with $s^*_{c|i} > \bar{s}_{c|i}$ in the post-2009 period.

Table 6 reports test results. Consistent with Proposition 3, Column (1) shows that $s^*_{c|i}$ is positively correlated with $\sigma^2_{\varepsilon_i}$. Greater idiosyncratic yield risk increases $s^*_{c|i}$ because it increases the expected YP payment but has no effect on the expected AYP payment. Column (2) shows that $s^*_{c|i}$ is marginally correlated with σ^2_c, which supports Conjecture 1 that the average effect of an increase in county yield risk on $s^*_{c|i}$ is small. Column (3) shows that $s^*_{c|i}$ is strongly positively correlated with β_i, revealing that an increase in unit yield's sensitivity to county yield amplifies the effect of county yield risk on unit yield risk.

Finally, Column (4) shows that $s^*_{c|i}$ has a relatively strong positive correlation with R^2_i at all coverage level combinations. The correlation coefficient ranges from 0.43 to 0.51 and monotonically increases with an increase in ϕ_c. These findings confirm our conjecture that units with higher systemic risk tend to have higher TRASR values. Thus, AYP is unable to simultaneously provide sufficient risk protection and meet farmer's transfer seeking demand, which further explains the low area insurance take-up.

We then check for the actual geographic pattern of AYP uptake. Figure 5 plots the geographic distribution of the share of acres insured under AYP in all acres insured under either AYP or YP for 2020. Most counties did not have any AYP policy records in the 2020 SOB data. Most remaining counties have very small shares of AYP insured acres, with a median value of 6.3%, whereas the share of yield-based contracts in total sold crop insurance contracts is already very low, with a median value of 3.3%. Overall, when $\phi_i = 75\%$ and $\phi_c = 90\%$, the respective means of county average systemic risk and county median $s^*_{c|i}$ values are 0.47 and 45% for counties whose AYP shares are above the 75th percentile, and 0.43 and 52% for counties whose AYP shares are below the 25th percentile. Differences in these numbers support our conjecture that AYP demand should be higher in areas with higher systemic risk and lower TRASR values. Given the small sample size, however, these differences are statistically insignificant.

5 | INCREASING AYP SUBSIDY RATES AS A POLICY OPTION TO GENERATE GREATER AYP DEMAND

As current AYP subsidy rates likely deter farmers from choosing AYP over YP, especially for farmers with high systemic risk, in this subsection we study whether raising AYP subsidy rates can increase AYP demand.

TABLE 6 Spearman's rank correlation coefficients between systemic risk variables and $s^*_{c|i}$

ϕ_i	ϕ_c	(1) $\sigma^2_{\varepsilon_i}$	(2) σ^2_c	(3) β_i	(4) R^2_i
75%	70%	0.3196*	0.0936*	0.8519*	0.4593*
	75%	0.3247*	0.0853*	0.8906*	0.4806*
	80%	0.3293*	0.0786*	0.9129*	0.4912*
	85%	0.3327*	0.0798*	0.9231*	0.4987*
	90%	0.3369*	0.0973*	0.9221*	0.5064*
80%	70%	0.3025*	0.0833*	0.8269*	0.4472*
	75%	0.3106*	0.0742*	0.8770*	0.4740*
	80%	0.3187*	0.0668*	0.9099*	0.4892*
	85%	0.3250*	0.0681*	0.9283*	0.5004*
	90%	0.3319*	0.0880*	0.9323*	0.5111*
85%	70%	0.2786*	0.0653*	0.7884*	0.4257*
	75%	0.2893*	0.0549*	0.8505*	0.4587*
	80%	0.3012*	0.0460*	0.8970*	0.4799*
	85%	0.3115*	0.0471*	0.9276*	0.4967*
	90%	0.3224*	0.0702*	0.9403*	0.5124*

Note: *Significance at 0.05 level.

We consider two subsidy schemes. We first study whether returning to pre-2009 relative subsidy rates could generate adequate transfers to farmers. As shown in Figure 1, there was a notable demand for area insurance before the 2008 Farm Bill adjusted down the relative subsidy rates of area insurance contracts over EU contracts, making the pre-2009 relative subsidy rate a reference point for any subsidy schemes that aim to increase area insurance demand. We then investigate whether offering free AYP contracts could increase AYP demand. The area insurance literature claims that area insurance plans are superior to individual plans in reducing information and transaction costs. If so, then completely replacing individual insurance plans with fully subsidized area insurance plans may help reduce the overall costs of FCIP.

Panel B of Table 5 reports the units with $s^*_{c|i} > s_{c|i}$ as a percent of all units in the pre-2009 period. Compared with the post-2009 period, the pairwise percent of units with $s^*_{c|i} > s_{c|i}$ is 14–23 percentage points lower in the pre-2009 period. We also find that for units with systemic risk estimates above the 75th percentile, the magnitude of decline increases to more than 30 percentage points. Thus, the relative subsidy rate change arising from the 2008 Farm Bill significantly increased the share of units for which the net returns from buying AYP are lower than the net returns from buying YP. Units enjoying relatively good risk protection from AYP are most affected. Returning to the pre-2009 relative subsidy rate may help resuscitate AYP demand.

Panel C reports the units with $s^*_{c|i} > \bar{s}_{c|i}$ as a percent of all units, where $\bar{s}_{c|i}$ denotes the 100% AYP subsidy rate over the post-2009 YP subsidy rate. Results show that units with $s^*_{c|i} > \bar{s}_{c|i}$ amount to < 6% of all units when $\phi_c = 90\%$. Thus, fully subsidizing AYP contracts at the highest coverage level meets almost all farmers' minimum subsidy transfer requirements for choosing AYP over YP, and so may generate a sizable demand for AYP. However, among all units, the percent of units with $s^*_{c|i} > \bar{s}_{c|i}$ is only 4–10 percentage points lower than the percent of units for which $s^*_{c|i} > s_{c|i}$ in the pre-2009 period, suggesting that the AYP demand under full AYP subsidy rates is unlikely to significantly surpass its pre-2009 counterpart.

It is worth noting that the above results are not based on a strict counterfactual analysis and thus should be interpreted cautiously. Other factors might have also contributed to the decrease in AYP

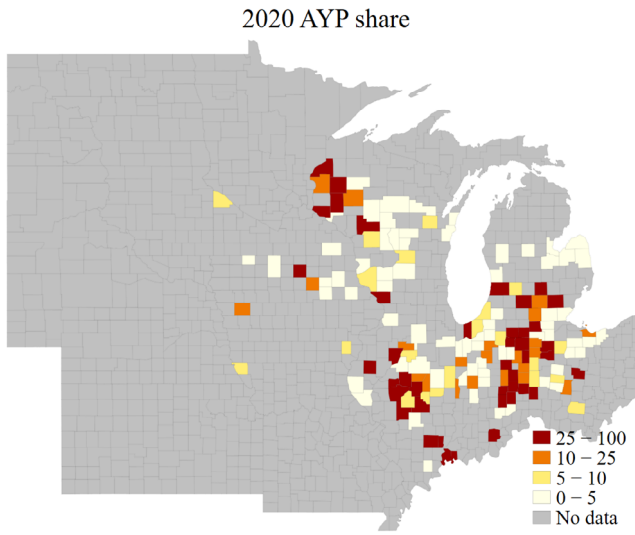

FIGURE 5 Geographic distribution of AYP share in 2020 *Notes*: 1. Data Source: Summary of Business, 2020, RMA. 2. AYP share is defined as the share of AYP insured acres in total acres insured under either AYP or YP in a county. Numbers in legends are in %. 3. Three counties had AYP shares >80%. In these counties, fewer than 10 yield insurance contracts were sold in 2020 and farms that bought AYP contracts were more representative of the area than farms that bought YP contracts.

demand. For example, Figure 1 shows that the area insurance insured acre share peaked in 2006 and declined thereafter. The subsidy rate change in 2009 substantially accelerated this downward trend. One explanation for the pre-2009 shift downward is that some area insurance buyers who had suffered losses but had not been paid ceased buying area insurance; that is, area insurance buyers realized basis risk and then abandoned area insurance. If so, then raising AYP subsidy rates is unlikely to re-engage this group. Similarly, new AYP buyers will be exposed to basis risk and may switch back from AYP to YP after they experience adverse outcomes. So AYP demand may first increase but then decrease after a rise in subsidy rate. In addition, AYP does not provide several benefits that YP provides, such as prevented planting and replanting payments (Barnett et al., 2005). These constraints may further limit farmers' willingness to choose AYP.[15]

Besides these effectiveness concerns, raising AYP subsidy rates to increase AYP demand presents several other concerns. First, the federal government has long sought to reduce the FCIP cost, while raising AYP subsidy rates will increase AYP subsidy costs. This would be of less concern were AYP more cost efficient than YP in reducing the overall FCIP cost. However, to our best knowledge, no studies have systematically investigated the actual cost efficiency of area insurance plans over individual plans, as applied in practice. In contrast, some current program features provide good reasons to be skeptical about this efficiency claim. For example, RMA requires area insurance buyers to report their acreage and production data as individual insurance buyers do. The data collection costs of area insurance plans are thus comparable with those of individual insurance plans. Second, expanding area insurance programs will likely diminish the risk protection efficiency of crop insurance programs and undermine FCIP's role in deterring ad hoc disaster aid. Innes (2003) and Bulut (2017) have shown that individual insurance can effectively deter disaster aid expectations. However, although ad hoc disaster aid and area insurance payments are both generally triggered by catastrophic events, heterogeneity in basis risk may cause some area insurance contract buyers to suffer

[15]The advents of the Supplemental Coverage Option (SCO) in 2015 and the Enhanced Coverage Option (ECO) in 2021 may further reduce farmers' willingness to choose AYP over YP because AYP buyers are not eligible to buy SCO or ECO. AYP subsidy rates need to further increase to equalize expected net returns from purchasing AYP contracts and from purchasing YP-SCO, YP-ECO, or YP-SCO-ECO contracts.

more than others and so call for ad hoc disaster aid.[16] Finally, free area insurance contracts may bias farmers' optimal decisions (see Proposition 3 on p. 326 of Innes, 2003) and lead to a lemon-like problem in the individual insurance market. Proposition A1 in Appendix S1 shows that when the relationship between unit yield and county yield follows LAM, then TRASR decreases with an increase in the expected unit yield. The analysis validates the proposition for historical data. Free AYP contracts thus are likely to drive high-yield (and usually low-risk) farms out of YP. To the extent that premium rates on high-risk land for a given county are set lower than the actuarially fair rate (Maisashvili et al., 2019; Price et al., 2019; Ramirez & Shonkwiler, 2017), the remaining YP contracts in the book of business are more likely to give rise to net losses.

6 | SYSTEMIC YIELD RISK VERSUS SYSTEMIC REVENUE RISK

Our study focuses on systemic yield risk and thus on yield insurance instead of systemic revenue risk and revenue insurance. Revenue insurance is the most popular crop insurance contract form among U.S. row crop farmers (Schnitkey et al., 2020). Its payment can be triggered by a decline in either price or yield, thus protecting farmers against both yield and price risks. As crop price risk needs to be accounted for, systemic revenue risk is a more complicated issue where systemic yield risk is just one piece. Another major piece is how prices relate to the systemic and idiosyncratic components of yield variability. Introducing price variation will require a very different set of tools, perhaps including copula methods, in order to appropriately model correlation structures (Goodwin & Hungerford, 2015). Systemic revenue risk will likely exceed price-scaled systemic yield risk, as price variation is the same for both individual revenue and area revenue. However, price and yield should to some extent move in opposing directions due to market forces. This market force effect will strengthen whenever individual yield correlates strongly with aggregate yield. We refer to aggregate yield as national yield or global yield because the effect of county yield on price variation is negligible. Thus, although farmers face the same (or quite similar) price risks, their systemic revenue risks likely differ from their systemic yield risks due to different correlation levels with the aggregation yield.

Systemic revenue risk is also likely to differ across revenue contract forms. Revenue insurance contracts with Harvest Price Exclusion (HPE) pay indemnities whenever the product of the realized yield and the realized price falls below the product of the guaranteed yield and the projected price, and thus incorporates the natural hedge. Revenue contracts without HPE, which is by far the most popular insurance contract form, allow farmers to use harvest price in determining indemnity payment whenever the harvest price exceeds the projected price. Thus, HPE is likely to fundamentally alter the systemic component of revenue risk. The magnitude, geographic distribution, and decomposition of systemic revenue risk needs study both in its own right and in comparison with yield risk. We leave these questions for further studies.

7 | CONCLUDING REMARKS

Systemic risk partly justifies government intervention in crop insurance markets and the introduction of area insurance plans, but the existing relevant literature was not intended to build up our understanding of systemic risk from its fundamental sources. Emphasizing natural resources endowments, in this paper we have modeled, measured, and decomposed systemic risk in corn yield across the Greater Midwest. These we have done both in concept and by implementing on unit-level RMA yield data. We find that systemic risk explains a little less than half of unit yield risk on average,

[16]It is noteworthy that recent ad hoc payments were more often triggered by non-yield shocks, such as the Market Facilitation Program and the Coronavirus Food Assistance Program, see Zulauf et al. (2020) for more information.

revealing a moderate level of co-movement between unit yield and area yield. This finding suggests limited risk management effectiveness of AYP. We also investigate whether AYP has provided sufficient premium subsidies to compete with YP. We develop a new concept, the threshold relative area subsidy rate, or TRASR, the relative subsidy rate of AYP over YP at which AYP's expected net return equals that of YP. Our calibrated TRASR values indicate that current AYP subsidy rates discourage farmers from choosing AYP over YP, especially at low AYP coverage levels. We also find that TRASR is positively correlated with systemic risk at the unit-level of analysis. Thus, in the presence of subsidized YP contracts, AYP contracts are typically unable to simultaneously meet farmers' risk protection and transfer-seeking demands.

We also briefly evaluate the policy option of increasing AYP subsidy rate to generate greater AYP demand. We find that even free AYP contracts are unlikely to support an AYP demand that significantly exceeds the pre-2009 level. It is generally acknowledged that basis risk deters farmers from choosing AYP contracts and discourages AYP buyers from renewing that contract form. Our study of systemic risk helps quantify the extent of such deterrence. Further inquiries are needed into the sensitivity of AYP demand to subsidy rates as well as into implications for budgetary costs and risk protection efficiencies of FCIP.

Climate change may alter future systemic risk patterns and area insurance demand. Our analysis finds that systemic risk significantly increases with drought occurrence and more excessive heat events. These events are projected to be more frequent with the changing climate. Notwithstanding, our result also shows that systemic risk decreases with more excessive precipitation events, which are also projected to occur with higher frequency. This impact may, however, be ameliorated or even negated because century-old U.S. midwestern drainage systems are midway through major upgrades that commenced around 1990 (Castellano et al., 2019). Looking forward, area-specific systemic risk patterns and area insurance demand can be investigated by linking downscaled climate and weather projection data to statistical yield models that account for intracounty, weather-conditioned yield distributions. Additional opportunities will arise whenever quality estimates of annual yield information at point locations become available because such data can be connected to soil maps. Findings in Jin et al. (2017), Li et al. (2019), Jiang et al. (2020), and elsewhere provide evidence that remote sensing and related inference methods may be close to making this connection.

It is also worth noting that switching from individual insurance to area insurance may change farmers' production behavior. Previous studies generally agree that area insurance helps remove moral hazard issues and thus should not alter farmers' production decisions. One exception is Chambers and Quiggin (2002), who argue that farmers would adjust their production behavior to align their own risks with area risk and thus to take advantage of the income-smoothing properties of AYP contracts. Although we do not model farmers' behavior, our decomposition of systemic risk and our study of how irrigation affects systemic risk suggest that farmers who adopt the prevalent production practice in a county bear higher systemic risk and should enjoy better risk protection from AYP than those who adopt less common production practices. One implication of this inference is that AYP may discourage farmers from choosing a technology new to an area and yet in time encourage laggards to adopt so that their yield realizations become better aligned with those of other producers in the area. Moreover, farmers' tendency to align their own production choices with prevalent production practices in the presence of AYP is likely to increase systemic yield risk and promote lock-in of Pareto-dominated production practices induced by networking effects.[17] Insurance policy design would benefit from a deeper understanding of the interactions among systemic risk, AYP, and related production choices.

ACKNOWLEDGMENTS

This paper was initiated while all were at Michigan State University. We gratefully acknowledge support from the Elton R. Smith Foundation, Michigan State University. This paper is also partially

[17]See Cowan and Gunby (1996), Holmes and Lee (2012), and Arora et al. (2021) for discussions on network lock-in in agriculture.

funded by a grant from the National Institute of Food and Agriculture (NIFA) of USDA, Award No. 2021-67023-34928. We are most grateful for valuable feedback provided by the editor, Timothy Richards, and three anonymous referees. Open access funding provided by the Iowa State University Library.

REFERENCES

Annan, Francis, and Wolfram Schlenker. 2015. "Federal Crop Insurance and the Disincentive to Adapt to Extreme Heat." *American Economic Review* 105(5): 262–6. https://doi.org/10.1257/aer.p20151031.

Arora, Gaurav, Hongli Feng, David A. Hennessy, Charles R. Loesch, and Susan Kvas. 2021. "The Impact of Production Network Economies on Spatially-Contiguous Conservation–Theoretical Model with Evidence from the U.S. Prairie Pothole Region." *Journal of Environmental Economics and Management* 107(5): 102442. https://doi.org/10.1016/j.jeem.2021.102442.

Awondo, S., and G. Datta. 2018. "Performance of Farm Level vs Area Level Crop Insurance." Paper presented at 30th International Conference of Agricultural Economists. No. 277265. Vancouver, British Columbia, July 28–August 2, 2018.

Barnett, Barry J., J. Roy Black, Hu Yingyao, and Jerry R. Skees. 2005. "Is Area Yield Insurance Competitive with Farm Yield Insurance?" *Journal of Agricultural and Resource Economics* 30(2): 285–301.

Bulut, Harun. 2017. "Managing Catastrophic Risk in Agriculture through ex Ante Subsidized Insurance or ex Post Disaster Aid." *Journal of Agricultural and Resource Economics* 42(3): 406–26.

Bulut, Harun. 2020. "The Impact of Enterprise Unit Policy Change on the Quantity Demanded for Crop Insurance." *Agricultural Finance Review* 80(4): 507–27. https://doi.org/10.1108/AFR-08-2019-0090.

Bulut, Harun, and Keith J. Collins. 2014. "Designing Farm Supplemental Revenue Coverage Options on Top of Crop Insurance Coverage." *Agricultural Finance Review* 74(3): 397–426. https://doi.org/10.1108/AFR-08-2013-0032.

Carriquiry, Miguel A., Bruce Babcock, and Chad E. Hart. 2008. "Using a Farmer's Beta for Improved Estimation of Expected Yields." *Journal of Agricultural and Resource Economics* 33(1): 52–68.

Castellano, Michael J., Sotirios v. Archontoulis, Matthew J. Helmers, Hanna J. Poffenbarger, and Johan Six. 2019. "Sustainable Intensification of Agricultural Drainage." *Nature Sustainability* 2(10): 914–21. https://doi.org/10.1038/s41893-019-0393-0.

Chambers, Robert G., and John Quiggin. 2002. "Optimal Producer Behavior in the Presence of Area-Yield Crop Insurance." *American Journal of Agricultural Economics* 84(2): 320–34. https://doi.org/10.1111/1467-8276.00300.

Claassen, Roger, and Richard E. Just. 2011. "Heterogeneity and Distributional Form of Farm-Level Yields." *American Journal of Agricultural Economics* 93(1): 144–60. https://doi.org/10.1093/ajae/aaq111.

Clarke, Daniel J. 2016. "A Theory of Rational Demand for Index Insurance." *American Economic Journal: Microeconomics* 8(1): 283–306. https://doi.org/10.1257/mic.20140103.

Coble, Keith H. 2017. "The Use of Enterprise Units in Crop Insurance." Agricultural Economics Blog, Mississippi State University. June 23, 2017. https://blogs.extension.msstate.edu/agecon/2017/07/23/the-use-of-enterprise-units-in-crop-insurance/.

Coble, Keith H., and Robert Dismukes. 2008. "Distributional and Risk Reduction Effects of Commodity Revenue Program Design." *Review of Agricultural Economics* 30(3): 543–53. https://doi.org/10.1111/j.1467-9353.2008.00429.x.

Cowan, Robin, and Philip Gunby. 1996. "Sprayed to Death: Path Dependence, Lock-in and Pest Control Strategies." *Economic Journal* 106(436): 521. https://doi.org/10.2307/2235561.

Deng, Xiaohui, Barry J. Barnett, and Dmitry V. Vedenov. 2007. "Is there a Viable Market for Area-Based Crop Insurance?" *American Journal of Agricultural Economics* 89(2): 508–19. https://doi.org/10.1111/j.1467-8276.2007.00975.x.

Du, Xiaodong, Hongli Feng, and David A. Hennessy. 2017. "Rationality of Choices in Subsidized Crop Insurance Markets." *American Journal of Agricultural Economics* 99(3): 732–56. https://doi.org/10.1093/ajae/aaw035.

Du, Xiaodong, David A. Hennessy, Hongli Feng, and Gaurav Arora. 2018. "Land Resilience and Tail Dependence among Crop Yield Distributions." *American Journal of Agricultural Economics* 100(3): 809–28. https://doi.org/10.1093/ajae/aax082.

Du, Xiaodong, Cindy L. Yu, David A. Hennessy, and Ruiqing Miao. 2015. "Geography of Crop Yield Skewness." *Agricultural Economics* 46(4): 463–73. https://doi.org/10.1111/agec.12174.

Feng, Hongli, Du Xiaodong, and David A. Hennessy. 2020. "Depressed Demand for Crop Insurance Contracts, and a Rationale Based on Third Generation Prospect Theory." *Agricultural Economics* 51(1): 59–73. https://doi.org/10.1111/agec.12541.

Gallagher, Paul. 1987. "U.S. Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances." *American Journal of Agricultural Economics* 69(4): 796–803. https://doi.org/10.2307/1242190.

Goodwin, Barry K., and Ashley Hungerford. 2015. "Copula-Based Models of Systemic Risk in U.S. Agriculture: Implications for Crop Insurance and Reinsurance Contracts." *American Journal of Agricultural Economics* 97(3): 879–96. https://doi.org/10.1093/ajae/aau086.

Goodwin, Barry K., and Alan P. Ker. 1998. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts." *American Journal of Agricultural Economics* 80(1): 139–53. https://doi.org/10.2307/3180276.

Goodwin, Barry K., and Vincent H. Smith. 2013. "What Harm Is Done by Subsidizing Crop Insurance?" *American Journal of Agricultural Economics* 95(2): 489–97. https://doi.org/10.1093/ajae/aas092.

Harri, Ardian, Keith H. Coble, Alan P. Ker, and Barry J. Goodwin. 2011. "Relaxing Heteroscedasticity Assumptions in Area-Yield Crop Insurance Rating." *American Journal of Agricultural Economics* 93(3): 707–17. https://doi.org/10.1093/ajae/aar009.

Hayes, Dermot J., Sergio H. Lence, and Chuck Mason. 2003. "Could the Government Manage its Exposure to Crop Reinsurance Risk?" *Agricultural Finance Review* 63(2): 127–42. https://doi.org/10.1108/00215040380001145.

Hill, Ruth Vargas, Miguel Robles, and Francisco Ceballos. 2016. "Demand for a Simple Weather Insurance Product in India: Theory and Evidence." *American Journal of Agricultural Economics* 98(4): 1250–70. https://doi.org/10.1093/ajae/aaw031.

Holmes, Thomas J., and Sanghoon Lee. 2012. "Economies of Density Versus Natural Advantage: Crop Choice on the Back Forty." *Review of Economics and Statistics* 94(1): 1–19. https://doi.org/10.1162/REST_a_00149.

Huettner, Frank, and Marco Sunder. 2012. "Axiomatic Arguments for Decomposing Goodness of Fit According to Shapley and Owen Values." *Electronic Journal of Statistics* 6: 1239–50. https://doi.org/10.1214/12-EJS710.

Innes, Robert. 2003. "Crop Insurance in a Political Economy: An Alternative Perspective on Agricultural Policy." *American Journal of Agricultural Economics* 85(2): 318–35. https://doi.org/10.1111/1467-8276.00122.

Jiang, Hao, Hu Hao, Renhai Zhong, Xu Jinfan, Xu Jialu, Jingfeng Huang, Shaowen Wang, Yibin Ying, and Tao Lin. 2020. "A Deep Learning Approach to Conflating Heterogeneous Geospatial Data for Corn Yield Estimation: A Case Study of the US Corn Belt at the County Level." *Global Change Biology* 26(3): 1754–66. https://doi.org/10.1111/gcb.14885.

Jin, Zhenong, George Azzari, and David B. Lobell. 2017. "Improving the Accuracy of Satellite-Based High-Resolution Yield Estimation: A Test of Multiple Scalable Approaches." *Agricultural and Forest Meteorology* 247(12): 207–20. https://doi.org/10.1016/j.agrformet.2017.08.001.

Keller, James B., and Tina L. Saitone. 2022. "Basis Risk in the Pasture, Rangeland, and Forage Insurance Program: Evidence from California." *American Journal of Agricultural Economics* 104(4): 1203–23. https://doi.org/10.1111/ajae.12282.

Ker, Alan P., and Barry K. Goodwin. 2000. "Nonparametric Estimation of Crop Insurance Rates Revisited." *American Journal of Agricultural Economics* 82(2): 463–78. https://doi.org/10.1111/0002-9092.00039.

Lampe, Immanuel, and Daniel Würtenberger. 2020. "Loss Aversion and the Demand for Index Insurance." *Journal of Economic Behavior and Organization* 180(12): 678–93. https://doi.org/10.1016/j.jebo.2019.10.019.

Li, Yan, Kaiyu Guan, Yu Albert, Bin Peng, Lei Zhao, Bo Li, and Jian Peng. 2019. "Toward Building a Transparent Statistical Model for Improving Crop Yield Prediction: Modeling Rainfed Corn in the U.S." *Field Crops Research* 234(3): 55–65. https://doi.org/10.1016/j.fcr.2019.02.005.

Liu, Elaine M. 2013. "Time to Change What to Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China." *Review of Economics and Statistics* 95(4): 1386–403. https://doi.org/10.1162/REST_a_00295.

Maisashvili, Aleksandre, Henry Bryant, George Knapek, and James Marc Raulston. 2019. "Are Crop Insurance Premium-Implied Yield Distributions Valid?" *Agricultural Finance Review* 79(4): 467–90. https://doi.org/10.1108/AFR-09-2018-0077.

Miranda, Mario J. 1991. "Area-Yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73(2): 233–42. https://doi.org/10.2307/1242708.

Miranda, Mario J., and Joseph W. Glauber. 1997. "Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets." *American Journal of Agricultural Economics* 79(1): 206–15. https://doi.org/10.2307/1243954.

Neild, Ralph E., and James E. Newman. 1987. "Growing Season Characteristics and Requirements in the Corn Belt." Cooperative Extension Service, Purdue University. https://www.extension.purdue.edu/extmedia/NCH/NCH-40.html.

Plastina, Alejandro, and William Edwards. 2014. "Trend-Adjusted Actual Production History (APH)." Ag Decision Maker, Extension and Outreach, Iowa State University. https://www.extension.iastate.edu/agdm/crops/html/a1-56.html.

Plastina, Alejandro, and William Edwards. 2017. "Proven Yields and Insurance Units for Crop Insurance." Ag Decision Maker, Extension and Outreach, Iowa State University. https://www.extension.iastate.edu/agdm/crops/html/a1-55.html.

Price, Michael J., Cindy L. Yu, David A. Hennessy, and Du. Xiaodong. 2019. "Are Actuarial Crop Insurance Rates Fair?: An Analysis Using a Penalized Bivariate B-Spline Method." *Journal of the Royal Statistical Society. Series C: Applied Statistics* 68(5): 1207–32. https://doi.org/10.1111/rssc.12363.

Ramaswami, Bharat, and Terry L. Roe. 2004. "Aggregation in Area-Yield Crop Insurance: The Linear Additive Model." *American Journal of Agricultural Economics* 86(2): 420–31. https://doi.org/10.1111/j.0092-5853.2004.00588.x.

Ramirez, Octavio A., and J. Scott Shonkwiler. 2017. "A Probabilistic Model of the Crop Insurance Purchase Decision." *Journal of Agricultural and Resource Economics* 42(1): 10–26.

Ray, Deepak K., James S. Gerber, Graham K. MacDonald, and Paul C. West. 2015. "Climate Variation Explains a Third of Global Crop Yield Variability." *Nature Communications* 6(1): 5989. https://doi.org/10.1038/ncomms6989.

Rothschild, Michael, and Joseph Stiglitz. 1970. "Increasing Risk: I. A Definition." *Journal of Economic Theory* 2(3): 225–43.

Schlenker, Wolfram, and Michael J. Roberts. 2006. "Nonlinear Effects of Weather on Corn Yields." *Review of Agricultural Economics* 28(3): 391–8. https://doi.org/10.1111/j.1467-9353.2006.00304.x.

Schlenker, Wolfram, and Michael J. Roberts. 2009. "Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change." *Proceedings of the National Academy of Sciences* 106(37): 15594–8. https://doi.org/10.1073/pnas.0906865106.

Schnitkey, Gary, Nick Paulson, Krista Swanson, and Carl Zulauf. 2020. "Revenue Protection: The Most Used Crop Insurance Product." Farmdoc Daily, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. November 17, 2020. https://farmdocdaily.illinois.edu/2020/11/revenue-protection-the-most-used-crop-insurance-product.html.

Schnitkey, Gary, and Bruce Sherrick. 2014. "Coverage Levels on Crop Insurance and the SCO Alternative." Farmdoc Daily, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. April 29, 2014. https://farmdocdaily.illinois.edu/2014/04/coverage-levels-on-crop-insurance-and-sco.html.

Sherrick, Bruce J., Fabio C. Zanini, Gary D. Schnitkey, and Scott H. Irwin. 2004. "Crop Insurance Valuation under Alternative Yield Distributions." *American Journal of Agricultural Economics* 86(2): 406–19. https://doi.org/10.1111/j.0092-5853.2004.00587.x.

Skees, Jerry R., J. Roy Black, and Barry J. Barnett. 1997. "Designing and Rating an Area Yield Crop Insurance Contract." *American Journal of Agricultural Economics* 79(2): 430–8. https://doi.org/10.2307/1244141.

Stigler, Matthieu, and David Lobell. 2021. "On the Benefits of Index Insurance in US Agriculture: A Large-Scale Analysis Using Satellite Data." *arXiv preprint* arXiv:2011.12544. http://arxiv.org/abs/2011.12544.

Tack, Jesse B., Keith H. Coble, and Barry Barnett. 2018. "Warming Temperatures Will Likely Induce Higher Premium Rates and Government Outlays for the U.S. Crop Insurance Program." *Agricultural Economics* 49(5): 635–47. https://doi.org/10.1111/agec.12448.

Tack, Jesse B., and Matthew T. Holt. 2016. "The Influence of Weather Extremes on the Spatial Correlation of Corn Yields." *Climatic Change* 134(1–2): 299–309. https://doi.org/10.1007/s10584-015-1538-4.

Wang, H. Holly, and Hao Zhang. 2003. "On the Possibility of a Private Crop Insurance Market: A Spatial Statistics Approach." *Journal of Risk and Insurance* 70(1): 111–24. https://doi.org/10.1111/1539-6975.00051.

Zulauf, Carl, Vecdi Demircan, Gary Schnitkey, Art Barnaby, Gregg Ibendahl, and Kevin Herbel. 2013. "Examining Contemporaneous Farm and County Losses Using Farm Level Data." Paper presented at the Agricultural and Applied Economics Association's 2013 Crop Insurance and the Farm Bill Symposium, Louisville, KY. October, 8–9, 2013.

Zulauf, Carl, Gary Schnitkey, Jonathan Coppess, Nick Paulson, and Krista Swanson. 2020. "Ad Hoc Payments: A Leading Indicator of Farm Policy Change." Farmdoc Daily, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. July 29, 2020. https://farmdocdaily.illinois.edu/2020/07/ad-hoc-payments-a-leading-indicator-of-farm-policy-change.html.

Zulauf, Carl, Gary Schnitkey, Nick Paulson, and Jonathan Coppess. 2017. "Comparing NASS and RMA County Yields for Corn." Farmdoc Daily, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. November 2, 2017. https://farmdocdaily.illinois.edu/2017/11/comparing-nass-and-rma-county-yields-for-corn.html.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Gong, Xuche, David A. Hennessy, and Hongli Feng. 2022. "Systemic Risk, Relative Subsidy Rates, and Area Yield Insurance Choice." *American Journal of Agricultural Economics* 1–26. https://doi.org/10.1111/ajae.12342

AJAE appendix for Systemic Risk, Relative Subsidy Rates and Area Yield Insurance Choice

Xuche Gong, David A. Hennessy, Hongli Feng

July 15, 2022

The material contained herein is supplementary to the article named in the title and published in the *American Journal of Agricultural Economics*.

Part I. Proof that $\boldsymbol{\beta_i = \theta_i \sigma_i / \sigma_c}$

Multiply both sides of equation (1) by $\tilde{y}_c$ and take expectations to obtain

$$\begin{aligned} \mathrm{E}(\tilde{y}_i \tilde{y}_c) &= \mu_i \mathrm{E}(\tilde{y}_c) + \beta_i \mathrm{E}[(\tilde{y}_c - \mu_c)\tilde{y}_c] + \mathrm{E}(\varepsilon_i \tilde{y}_c) \\ &= \mu_i \mathrm{E}(\tilde{y}_c) + \beta_i \mathrm{E}(\tilde{y}_c^2) - \beta_i \mu_c \mathrm{E}(\tilde{y}_c) \qquad \text{(A1)} \\ &= \mu_i \mu_c + \beta_i \mathrm{E}(\tilde{y}_c^2) - \beta_i \mu_c^2, \end{aligned}$$

where the second equation holds because $\mathrm{E}(\varepsilon_i \tilde{y}_c) = 0$. Since $\mathrm{E}(\tilde{y}_i \tilde{y}_c) = \mathrm{Cov}(\tilde{y}_i, \tilde{y}_c) + \mathrm{E}(\tilde{y}_i)\mathrm{E}(\tilde{y}_c) = \mathrm{Cov}(\tilde{y}_i, \tilde{y}_c) + \mu_i \mu_c$ and $\mathrm{E}(\tilde{y}_c^2) = \mathrm{Var}(\tilde{y}_c) + \mathrm{E}(\tilde{y}_c)\mathrm{E}(\tilde{y}_c) = \mathrm{Var}(\tilde{y}_c) + \mu_c^2$, equation (A1) can be rewritten as $\mathrm{Cov}(\tilde{y}_i, \tilde{y}_c) + \mu_i \mu_c = \mu_i \mu_c + \beta_i[\mathrm{Var}(\tilde{y}_c) + \mu_c^2] - \beta_i \mu_c^2$. With some rearrangements, we obtain $\mathrm{Cov}(\tilde{y}_i, \tilde{y}_c) = \beta_i \mathrm{Var}(\tilde{y}_c)$. Thus, $\beta_i = \mathrm{Cov}(\tilde{y}_i, \tilde{y}_c)/\mathrm{Var}(\tilde{y}_c) = \theta_i \sigma_i \sigma_c / \sigma_c^2 = \theta_i \sigma_i / \sigma_c$.

Part II. How TRASR divides subsidy rate space

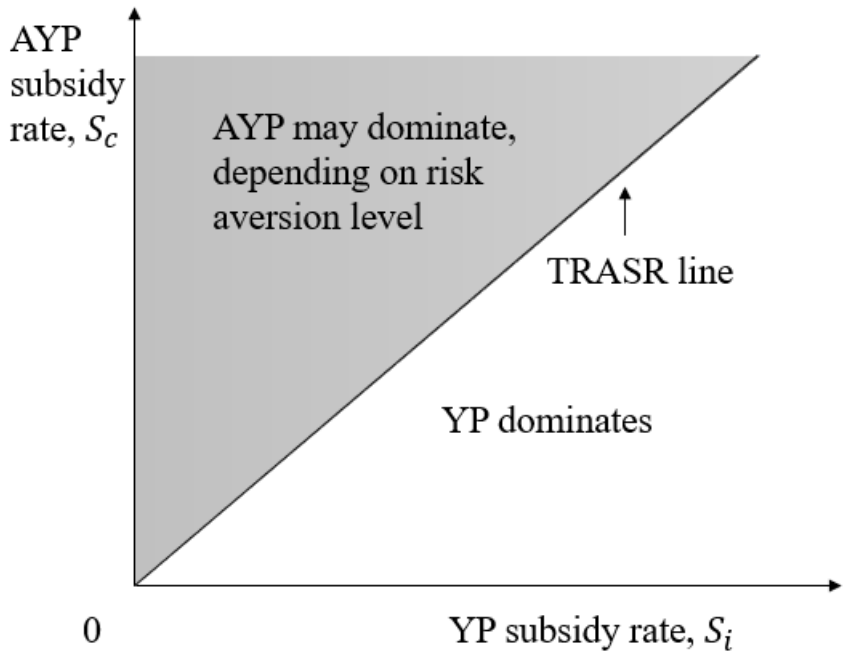

Figure A1. How TRASR divides the subsidy rate space of AYP and YP

Part III. Proof for equation (11)

We assume that $\tilde{y}_c$ is continuously distributed over the range $[0, y_c^u]$. This assumption leads to

the conclusion that $F(0) = 0$. We check for zero-value yield in the NASS county corn yield data. Among the 53,932 county-year observations between 1910 and 2019 for our 580 sample counties, no one has a zero-value yield. Moreover, among the total available 177,621 county-year observations from 1910 to 2019 in the NASS county corn yield dataset, only 0.28% of observations have zero-value yields. Thus, the condition $F(0) = 0$ should hold in most cases, especially for our sample counties.

Since $\tilde{y}_i \in [0, y_i^u]$, $\tilde{y}_c \in [0, y_c^u]$ and $\beta_i > 0$, then by equation (1), $\varepsilon_i \in [-\mu_i - \beta_i(y_c^u - \mu_c), y_i^u - \mu_i + \beta_i\mu_c]$. Let $\bar{\varepsilon}_i = y_i^u - \mu_i + \beta_i\mu_c$ and $\underline{\varepsilon}_i = -\mu_i - \beta_i(y_c^u - \mu_c)$, and given $\mathrm{M}_i(\varepsilon_i) = \mu_c - [\mu_i(1-\phi_i) + \varepsilon_i]/\beta_i$, $\mathrm{M}_c = \mu_c\phi_c$, $\mathrm{M}_c^l = \mu_c l$, and $\alpha_c = \rho/(\phi_c - l)$, taking expected values in equation (9) yields

$$
\begin{aligned}
\mathrm{E}(\tilde{n}_i) &= \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \left\{ \int_0^{\mathrm{M}_i(\varepsilon_i)} \beta_i[\mathrm{M}_i(\varepsilon_i) - \tilde{y}_c] dF(\tilde{y}_c) + \int_{\mathrm{M}_i(\varepsilon_i)}^{y_c^u} 0 dF(\tilde{y}_c) \right\} dG(\varepsilon_i) \\
&= \beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \left\{ \int_0^{\mathrm{M}_i(\varepsilon_i)} [\mathrm{M}_i(\varepsilon_i) - \tilde{y}_c] dF(\tilde{y}_c) \right\} dG(\varepsilon_i) \\
&= \beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \left\{ \int_0^{\mathrm{M}_i(\varepsilon_i)} \mathrm{M}_i(\varepsilon_i) dF(\tilde{y}_c) - \int_0^{\mathrm{M}_i(\varepsilon_i)} \tilde{y}_c dF(\tilde{y}_c) \right\} dG(\varepsilon_i) \\
&= \beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \left\{ \mathrm{M}_i(\varepsilon_i) F(\tilde{y}_c)\Big|_{\tilde{y}_c=0}^{\tilde{y}_c=\mathrm{M}_i(\varepsilon_i)} - \left[\tilde{y}_c F(\tilde{y}_c)\Big|_{\tilde{y}_c=0}^{\tilde{y}_c=\mathrm{M}_i(\varepsilon_i)} - \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \right] \right\} dG(\varepsilon_i) \\
&= \beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \left\{ \left[\mathrm{M}_i(\varepsilon_i) F\big(\mathrm{M}_i(\varepsilon_i)\big) - \mathrm{M}_i(\varepsilon_i) F(0) \right] - \left[\mathrm{M}_i(\varepsilon_i) F\big(\mathrm{M}_i(\varepsilon_i)\big) - 0F(0) \right] \right. \\
&\qquad \left. + \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \right\} dG(\varepsilon_i) \\
&= \beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i).
\end{aligned}
$$

Similarly, taking expected values in equation (10) yields

$$
\begin{aligned}
\mathrm{E}(\tilde{n}_c) &= \int_0^{\mathrm{M}_c^l} \mu_c \rho dF(\tilde{y}_c) + \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} (\mathrm{M}_c - \tilde{y}_c) dF(\tilde{y}_c) + \int_{\mathrm{M}_c}^{y_c^u} 0 dF(\tilde{y}_c) \\
&= \int_0^{\mathrm{M}_c^l} \mu_c \rho dF(\tilde{y}_c) + \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \mathrm{M}_c dF(\tilde{y}_c) - \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c) \\
&= \mu_c \rho F(\tilde{y}_c)|_{\tilde{y}_c=0}^{\tilde{y}_c=\mathrm{M}_c^l} + \alpha_c \mathrm{M}_c F(\tilde{y}_c)|_{\tilde{y}_c=\mathrm{M}_c^l}^{\tilde{y}_c=\mathrm{M}_c} - \alpha_c \left[\tilde{y}_c F(\tilde{y}_c)|_{\tilde{y}_c=\mathrm{M}_c^l}^{\tilde{y}_c=\mathrm{M}_c} - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c \right] \\
&= [\mu_c \rho F(\mathrm{M}_c^l) - \mu_c \rho F(0)] + \alpha_c [\mathrm{M}_c F(\mathrm{M}_c) - \mathrm{M}_c F(\mathrm{M}_c^l)] - \alpha_c [\mathrm{M}_c F(\mathrm{M}_c) - \mathrm{M}_c^l F(\mathrm{M}_c^l)] \\
&\qquad + \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c \\
&= \mu_c \rho F(\mathrm{M}_c^l) - \alpha_c (\mathrm{M}_c - \mathrm{M}_c^l) F(\mathrm{M}_c^l) + \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c \\
&= \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c,
\end{aligned}
$$

where the last equation holds because $\alpha_c(\mathrm{M}_c - \mathrm{M}_c^l)F(\mathrm{M}_c^l) = \rho(\mu_c\phi_c - \mu_c l)F(\mathrm{M}_c^l)/(\phi_c - l) = \mu_c \rho F(\mathrm{M}_c^l)$. Equation (11) then follows as

$$
s_{c|i}^* = \frac{\mathrm{E}(\tilde{n}_i)}{\mathrm{E}(\tilde{n}_c)} = \frac{\beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)}{\alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c}.
$$

Part IV. Comparative statics of TRASR with respect to systemic risk components

In this section we study the comparative statics of TRASR with respect to systemic risk components. We first study the effect of a mean-preserving spread (m.p.s.) of $G(\varepsilon_i)$ on $s_{c|i}^*$. Letting $H(\varepsilon_i) = \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c$, we have $dH(\varepsilon_i)/d\varepsilon_i = -F(\mathrm{M}_i(\varepsilon_i))/\beta_i < 0$ and $d^2H(\varepsilon_i)/d\varepsilon_i^2 = f(\mathrm{M}_i(\varepsilon_i))/\beta_i^2 > 0$, where $f(\cdot)$ is the probability density function (PDF) of $\tilde{y}_c$. Therefore $H(\varepsilon_i)$ is convex in ε_i and a m.p.s. of $G(\varepsilon_i)$ should increase the expected value of $H(\varepsilon_i)$. Thus, for any $G^m(\varepsilon_i)$ that is a m.p.s. of $G(\varepsilon_i)$, we have $\int_{\underline{\varepsilon}_i^m}^{\bar{\varepsilon}_i^m} H(\varepsilon_i) dG^m(\varepsilon_i) > \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} H(\varepsilon_i) dG(\varepsilon_i)$, where $\bar{\varepsilon}_i^m$ and $\underline{\varepsilon}_i^m$ are the respective upper and lower bounds for ε_i after the

m.p.s operation with $\bar{\varepsilon}_i^m \geq \bar{\varepsilon}_i$ and $\underline{\varepsilon}_i^m \leq \underline{\varepsilon}_i$. Assuming no single unit has an effect on county yield,[1] then $s_{c|i}^*|_{G^m(\varepsilon_i)} > s_{c|i}^*|_{G(\varepsilon_i)}$, i.e., $s_{c|i}^*$ increases with a m.p.s. of $G(\varepsilon_i)$.

A similar procedure can be performed to investigate the effect of a m.p.s. of $F(\tilde{y}_c)$ on $s_{c|i}^*$. The result is, however, ambiguous. Letting $F^m(\tilde{y}_c)$ be any m.p.s. of $F(\tilde{y}_c)$, then we have $\int_0^{\mathrm{M}_i(\varepsilon_i)} F^m(\tilde{y}_c)d\tilde{y}_c > \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c$, $\int_0^{\mathrm{M}_c^l} F^m(\tilde{y}_c)d\tilde{y}_c > \int_0^{\mathrm{M}_c^l} F(\tilde{y}_c)d\tilde{y}_c$ and $\int_0^{\mathrm{M}_c} F^m(\tilde{y}_c)d\tilde{y}_c > \int_0^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c$. However, we are unable to compare the value of $\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F^m(\tilde{y}_c)d\tilde{y}_c$ with that of $\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c$ unless we know exactly how $\tilde{y}_c$ is distributed within the interval $[0, \mathrm{M}_c]$ both before and after the m.p.s. operation. In addition, since unit yield risk is a function of county yield risk, the YP indemnity may also change with a m.p.s. of the county yield distribution, which further complicates the comparison of $s_{c|i}^*$ values before and after the m.p.s. operation.

Finally, we investigate the effect of an increase in β_i on $s_{c|i}^*$. We use a log version of equation (11) to simplify the analysis,

$$\text{(A2)} \qquad \ln(s_{c|i}^*) = \ln(\beta_i) + \ln\left[\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i}\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)\right] - \ln(\alpha_c) - \ln\left[\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c\right].$$

Differentiating $\ln(s_{c|i}^*)$ with respect to β_i yields

$$\text{(A3)} \qquad \frac{d\ln(s_{c|i}^*)}{d\beta_i} = \frac{1}{\beta_i} + \frac{\mu_i(1-\phi_i)\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\, dG(\varepsilon_i)}{\beta_i^2\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i}\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)} + \frac{\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i\, dG(\varepsilon_i)}{\beta_i^2\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i}\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)}.$$

Terms related to the partial effects of β_i on $\bar{\varepsilon}_i$ and $\underline{\varepsilon}_i$ are missing in equation (A3) because in taking the first-order partial derivative of $\ln(s_{c|i}^*)$ with respect to β_i, we hold fixed the distribution of ε_i. Thus, $\bar{\varepsilon}_i$ and $\underline{\varepsilon}_i$ can be viewed as constants with respect to β_i. To see this, recall that $\mathrm{M}_i(\varepsilon_i) = \mu_c - [\mu_i(1-\phi_i) + \varepsilon_i]/\beta_i$. Every parameter shown in $\mathrm{M}_i(\varepsilon_i)$ other than β_i

is held fixed in the comparative statics analysis.

The first two right-hand terms of equation (A3) are obviously positive but the third term is negative. This is because $F[\mathrm{M}_i(\varepsilon_i)]$ is a positive and decreasing function of ε_i, so $F(\mathrm{M}_i)\varepsilon_i$ overweighs negative ε_i values when compared with positive ε_i values. Then by $\mathrm{E}(\varepsilon_i) = 0$, we obtain $\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i \, dG(\varepsilon_i) < 0$. However, when $\varepsilon_i \equiv 0$ then $d\ln(s_{c|i}^*)/d\beta_i > 0$. Moreover, although we cannot sign equation (A3) when $\varepsilon_i \not\equiv 0$, a sufficient condition for $s_{c|i}^*$ to be increasing in β_i is that $\mu_i(1-\phi_i)\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i \, dG(\varepsilon_i) + \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i \, dG(\varepsilon_i) \geq 0$, because $\beta_i^2 \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i) > 0$ always hold. The condition is met whenever $\mu_i(1-\phi_i) + \varepsilon_i \geq 0$ for all values of ε_i, or whenever $-\underline{\varepsilon}_i \leq \mu_i(1-\phi_i)$, because then $\mu_i(1-\phi_i)\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i \, dG(\varepsilon_i) + \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)]\varepsilon_i \, dG(\varepsilon_i) = \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)][\mu_i(1-\phi_i) + \varepsilon_i] \, dG(\varepsilon_i) \geq 0$.

Part V. Comparative statics of TRASR with respective to non-systemic risk variables and parameters

1. TRASR and yield expectations

We now turn to analyzing how unit yield expectation affects $s_{c|i}^*$. Differentiating ln $(s_{c|i}^*)$ with respect to μ_i yields

$$\text{(A4)} \qquad \frac{d\ln(s_{c|i}^*)}{d\mu_i} = -\frac{(1-\phi_i)\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} F[\mathrm{M}_i(\varepsilon_i)] \, dG(\varepsilon_i)}{\beta_i \int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)} < 0,$$

i.e., $s_{c|i}^*$ is decreasing in the expected unit yield. With similar arguments in the comparative statics analysis of $s_{c|i}^*$ with respect to β_i, we can show that in taking the first-order partial derivative of $s_{c|i}^*$ with respect to μ_i, $\bar{\varepsilon}_i$ and $\underline{\varepsilon}_i$ can be viewed as constants with respect to μ_i and thus terms related to the partial effects of μ_i on $\bar{\varepsilon}_i$ and $\underline{\varepsilon}_i$ are missing in relationship (A4).

Relationship (A4) may appear at first to be counterintuitive because an increase in μ_i

increases the probability of receiving a YP payment but has no effect on the probability of receiving an AYP payment. This line of argument suggests that $s^*_{c|i}$ should increase with an increase in μ_i. However, this reasoning process allows μ_i to change but holds the distribution of $\tilde{y}_i$ fixed. Since $\mu_i = \mathrm{E}(\tilde{y}_i) = \int_0^{y_i^u} \tilde{y}_i \, dI(\tilde{y}_i)$, where $I(\cdot)$ is the CDF of $\tilde{y}_i$, then any change in μ_i should involve changes in the unit yield distribution.

Since ε_i and β_i are held fixed in taking the first-order partial derivative of $s^*_{c|i}$ with respect to μ_i, under the assumption that no single unit is sufficiently large to affect county yield, then an a-unit increase in μ_i can only result from an a-unit rightward shift in the unit yield distribution. To see this, recall that $\tilde{y}_i = \mu_i + \beta_i(\tilde{y}_c - \mu_c) + \varepsilon_i$, then $\tilde{y}_i + a = \mu_i + a + \beta_i(\tilde{y}_c - \mu_c) + \varepsilon_i$ must hold whenever $\beta_i(\tilde{y}_c - \mu_c)$ and ε_i are held fixed. In this case, the probability that yield is $\tilde{y}_i + a$ under a shift is the same as the probability that yield is $\tilde{y}_i$ under the initial distribution. Thus, the unit yield distribution just translates a units rightward and there is no redistribution of probabilities among possible yield realizations. However, the yield guarantee, $\phi_i \mu_i$, only shifts rightward by $\phi_i a < a$ units. Consequently, the probability of receiving YP indemnities decreases and the expected YP indemnity also decreases. TRASR decreases as a result.

The above reasoning process reveals that when TRASR is constructed from LAM, the comparative statics analysis requires an increase in μ_i to have no effect on the idiosyncratic risk and systemic risk components of unit yield. This constraint is removed when the TRASR index is directly constructed from equation (8), where LAM is not involved. The LAM-free TRASR index can either increase or decrease with an increase in μ_i, because different parts of the unit yield distribution can respond differently to an increase in μ_i.

The simple example in Table A1 illustrates this point. There are five possible yields, from 1 to 5, and an original yield probability distribution given by probability distribution 1. When the YP coverage level is 80%, then the expected YP indemnity under probability distribution 1 is

0.228. Now, suppose the probability distribution changes from distribution 1 to distribution 2, where there is a probability shift of 0.1 from 4 to 5. The resulting expected YP indemnity increases to 0.26. Similarly, suppose the probability distribution changes from distribution 1 to distribution 3, where there is a probability shift of 0.1 from 1 to 2. Then the resulting expected YP indemnity decreases to 0.16. Thus, whether the expected YP payment increases or decreases with an increase in μ_i depends on the specific distribution shift. Under the assumption that no single unit is sufficient large to affect county yield, the LAM-free TRASR value can either increase or decrease with an increase in μ_i.

Table A1. Expected YP indemnities under different yield distribution, $\boldsymbol{\phi_i = 80\%}$

Possible yields	1	2	3	4	5
Probability distribution 1	0.1	0.3	0.3	0.2	0.1
Expected yield, μ_i			2.9		
Yield guarantee, $\phi_i\mu_i$			2.32		
Expected indemnity, $\mathrm{E}(\tilde{n}_i)$			0.228		
Probability distribution 2	0.1	0.3	0.3	0.1	0.2
Expected yield, μ_i			3		
Yield guarantee, $\phi_i\mu_i$			2.4		
Expected indemnity, $\mathrm{E}(\tilde{n}_i)$			0.26		
Probability distribution 3	0	0.4	0.3	0.2	0.1
Expected yield, μ_i			3		
Yield guarantee, $\phi_i\mu_i$			2.4		
Expected indemnity, $\mathrm{E}(\tilde{n}_i)$			0.16		

Given Table A1's results, we test whether the comparative static result that TRASR decreases with an increase in μ_i holds in our data. We do so in two ways.

First, we check the correlation coefficient between TRASR and μ_i for each county. Units in the same county face the same county yield distribution but have different values of β_i, ε_i, and μ_i, and different unit yield distributions. As a result, different values of μ_i in a county sample are generally associated with different values of β_i and ε_i, and different unit yield distributions. More importantly, within each county, the yield distributions associated with greater μ_i values

are not identical to rightward shifts of the yield distributions associated with smaller μ_i values. Thus, testing the correlation coefficient between TRASR and μ_i within each county can reveal that, after allowing the unit yield distribution to vary with an increase in μ_i in the way that the data permit, how TRASR will change with an increase in μ_i. The testing results show that the correlation coefficients between TRASR and μ_i are negative for more than 85% of counties, indicating that the comparative static result that TRASR decreases with an increase in μ_i widely holds for the actual data available to us.

Second, we use another method to calculate $\mathrm{E}(\tilde{n}_i)$. For each unit, we first calculate the yield difference between the coverage level scaled 2008 APH yield and each of the ten previous year's actual yields. Each unit thus has ten yield difference values. Then for each unit we replace its negative yield difference values with zeros and average the resulting ten values to obtain the unit's $\mathrm{E}(\tilde{n}_i)$. This approach is completely free of LAM, while the correlation coefficients between $\mathrm{E}(\tilde{n}_i)$ and 2008 APH yield (at the full sample level) are from -0.41 to -0.53. Thus, even if LAM is not involved in constructing $\mathrm{E}(\tilde{n}_i)$ and TRASR, the comparative static result that TRASR decreases with an increase in μ_i still applies to our dataset.

As for the comparative statics analysis of $s^*_{c|i}$ with respect to μ_c, first note that LAM is not involved in the process of deriving $\mathrm{E}(\tilde{n}_c)$. As a result, taking the first-order partial derivative of $\mathrm{E}(\tilde{n}_c)$ with respect to μ_c ignores the condition that any change in the county yield expectation should involve changes in the county yield distribution, leading us to conclude that $\mathrm{E}(\tilde{n}_c)$ increases with an increase in μ_c as $d\ln[\mathrm{E}(\tilde{n}_c)]/d\mu_c = F(\mathrm{M}_c)\phi_c - F(\mathrm{M}_c^l)l > 0$. LAM, however, does not model the relationship between possible county yield realizations and the county yield expectation and thus provides no information on how the county yield distribution should change with an increase in μ_c. In fact, and similar to the example shown in Table A1, we can show that $\mathrm{E}(\tilde{n}_c)$ can either increase or decrease with an increase in μ_c whenever different

parts of the county yield distribution can respond differently to changes in μ_c. Thus, signing the effects of an increase in μ_c on $\mathrm{E}(\tilde{n}_c)$ and $s^*_{c|i}$ requires additional information or conditions on yield distributions and parameter values.

Proposition A1. *Ceteris paribus*, $s^*_{c|i}$ decreases with an increase in μ_i. The effects of an increase in μ_c on $s^*_{c|i}$ cannot be determined without additional information or conditions.

2. TRASR and coverage levels

Under actuarially fair premiums, it is obvious that an increase in ϕ_i will increase $s^*_{c|i}$ because it increases $\mathrm{E}(\tilde{n}_i)$ while having no effect on $\mathrm{E}(\tilde{n}_c)$. In addition, taking the first-order partial derivative of $\ln(s^*_{c|i})$ with respect to ϕ_c yields

$$\text{(A5)} \qquad \frac{d\ln(s^*_{c|i})}{d\phi_c} = \frac{1}{\phi_c - l} - \frac{F(\mathrm{M}_c)\mu_c}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} = \frac{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c}(\mathrm{M}_c^l - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} < 0,$$

where the second equality holds because

$$\begin{aligned}
\frac{1}{\phi_c - l} - \frac{F(\mathrm{M}_c)\mu_c}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} &= \frac{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c - F(\mathrm{M}_c)\mu_c(\phi_c - l)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\left[F(\tilde{y}_c)\big|_{\tilde{y}_c=\mathrm{M}_c^l}^{\tilde{y}_c=\mathrm{M}_c} - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c)\right] - [F(\mathrm{M}_c)\mathrm{M}_c - F(\mathrm{M}_c)\mathrm{M}_c^l]}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c)\mathrm{M}_c - F(\mathrm{M}_c^l)\mathrm{M}_c^l - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c) - F(\mathrm{M}_c)\mathrm{M}_c + F(\mathrm{M}_c)\mathrm{M}_c^l}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c)\mathrm{M}_c^l - F(\mathrm{M}_c^l)\mathrm{M}_c^l - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\int_0^{\mathrm{M}_c} \mathrm{M}_c^l dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c^l} \mathrm{M}_c^l dF(\tilde{y}_c) - \left[\int_0^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c^l} \tilde{y}_c dF(\tilde{y}_c)\right]}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\int_0^{\mathrm{M}_c}(\mathrm{M}_c^l - \tilde{y}_c)dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c^l}(\mathrm{M}_c^l - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c}(\mathrm{M}_c^l - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c}.
\end{aligned}$$

An increase in ϕ_c has a negative effect on AYP's indemnity scaling factor, α_c, but a positive effect on the unscaled AYP indemnity payment. The latter effect dominates the former effect, leading to an increase in $\mathrm{E}(\tilde{n}_c)$ and thus a decrease in $s^*_{c|i}$.

Proposition A2. *Ceteris paribus*, $s^*_{c|i}$ is increasing in ϕ_i and decreasing in ϕ_c.

3. TRASR, protection factor, and loss limit factor

Taking the first-order partial derivative of $\ln(s^*_{c|i})$ with respect to $\ln(\rho)$ yields

$$\frac{d\ln(s^*_{c|i})}{d\ln(\rho)} = -\frac{d\ln(\alpha_c)}{d\ln(\rho)} = -1. \tag{A6}$$

Thus, a one-percent increment in ρ decreases $s^*_{c|i}$ by one percent. Since TRASR is the relative subsidy rate of AYP over YP at which the expected net return of AYP equals that of YP, then to keep $s^*_{c|i}$ unchanged, s_c should decrease by one percent with a one-percent increment in ρ whenever s_i is held fixed. Thus, allowing farmers to choose a higher AYP protection factor level and increasing the AYP subsidy rate both increase the expected AYP payment, and they do so in the same way. A related point emerges from noting that farmers are best informed about their farm yield distributions while the distribution of county-level yield is common knowledge. It is likely therefore that, compared with increasing AYP subsidy rates, allowing farmers to freely choose the AYP protection factor may better encourage farmers to choose AYP contracts. However, as Miranda (1991) has noted, allowing farmers to freely choose the protection level would be politically infeasible while also raising the expected level and variability of total indemnity outlays.

If we assume that the value of the loss limit factor, l, can also be chosen by farmers, then we can obtain some insights into the relationships between $s^*_{c|i}$ and l. The first-order partial derivative of $\ln(s^*_{c|i})$ with respect to l is

(A7) $$\frac{d\ln\left(s_{c|i}^{*}\right)}{dl} = -\frac{1}{\phi_c - l} + \frac{F(\mathrm{M}_c^l)\mu_c}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} = \frac{-\int_{\mathrm{M}_c^l}^{\mathrm{M}_c}(\mathrm{M}_c - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} < 0,$$

where the second equality holds because

$$\begin{aligned}
-\frac{1}{\phi_c - l} + \frac{F(\mathrm{M}_c^l)\mu_c}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} &= \frac{F(\mathrm{M}_c^l)\mu_c(\phi_c - l) - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c^l)\mathrm{M}_c - F(\mathrm{M}_c^l)\mathrm{M}_c^l - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c^l)\mathrm{M}_c - F(\mathrm{M}_c^l)\mathrm{M}_c^l - \left[F(\tilde{y}_c)\big|_{\tilde{y}_c=\mathrm{M}_c^l}^{\tilde{y}_c=\mathrm{M}_c} - \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c)\right]}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c^l)\mathrm{M}_c - F(\mathrm{M}_c^l)\mathrm{M}_c^l - [F(\mathrm{M}_c)\mathrm{M}_c - F(\mathrm{M}_c^l)\mathrm{M}_c^l] + \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{F(\mathrm{M}_c^l)\mathrm{M}_c - F(\mathrm{M}_c)\mathrm{M}_c + \left[\int_0^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c^l} \tilde{y}_c dF(\tilde{y}_c)\right]}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\int_0^{\mathrm{M}_c^l} \mathrm{M}_c dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c} \mathrm{M}_c dF(\tilde{y}_c) + \int_0^{\mathrm{M}_c} \tilde{y}_c dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c^l} \tilde{y}_c dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{\int_0^{\mathrm{M}_c^l}(\mathrm{M}_c - \tilde{y}_c)dF(\tilde{y}_c) - \int_0^{\mathrm{M}_c}(\mathrm{M}_c - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} \\
&= \frac{-\int_{\mathrm{M}_c^l}^{\mathrm{M}_c}(\mathrm{M}_c - \tilde{y}_c)dF(\tilde{y}_c)}{(\phi_c - l)\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c}.
\end{aligned}$$

Thus, $s_{c|i}^{*}$ decreases with an increase in l because an increase in the loss limit factor (non-strictly) scales up the AYP indemnity payment at each $\tilde{y}_c$ value that triggers an AYP payment, but has no effect on the YP payment.

Proposition A3. *Ceteris paribus*, a one-percent increment in ρ decreases $s_{c|i}^{*}$ by one percent, while $s_{c|i}^{*}$ also decreases with an increase in l.

Part VI. Data screening process

Our unit-level data are sequences of four-to-ten-year historical yields that RMA uses to establish

Actual Production History (APH) yields. The historical yields are continuous unless the insured crop is not planted in a certain year. If at least four successive yield records are unavailable, a transition yield proportional to the ten-year average county yield is substituted in for each missing year. In our study, we only keep units with ten actual yield records within the 1998-2007 period because *i*) including units with transition yield records will introduce artificial correlation between county yield and unit yield and bias our estimate of systemic risk, *ii*) including years before 1998 will both admit unit observations where the crop was not grown in some years between 1998 and 2007 and also result in few unit observations in years that might not capture well the temporal systemic risk in some counties.[2] We also drop units under the irrigation practice and also counties where irrigation rates exceeded 20% because a systematic difference might exist between irrigated land and non-irrigated land, and also between high irrigation counties and other counties.[3] Table A2 summarizes observation losses for each data screening step as well as merging with county yield data, climate data, and land quality data.

Table A2. Observation losses for each data screening and merging step

	Number of counties remaining	Number of units remaining
Initial observations	1,031	1,470,015
Drop units with non-actual historical yields	1,003	545,548
Drop units with historical years before 1998	980	266,956
Drop irrigated units and counties whose irrigation rates exceeded 20%	847	226,754
Drop counties with less than 50-year data records between 1958 and 2007	712	223,422
Drop counties with less than 30 units	618	222,026
Merge with climate and land quality data	580	213,429

Part VII. Constructions of county climate variables

To construct GDD and SDD, we first define *i*) the daily maximum temperature (in degrees Celsius), $T_{c,d,t}^{Max}$, for county c as the mean of the highest temperatures recorded by all weather stations within that county on day d and year t, and *ii*) the daily minimum temperature, $T_{c,d,t}^{Min}$,

for county c as the mean of the lowest temperatures recorded by all weather stations within that county on day d and year t. For each county, GDD, labeled as G_c, is defined as

$$G_c = \frac{1}{10}\sum_{t\in Y_t}\sum_{d\in M_t}\left[0.5\left(\min\left(\max\left(T_{c,d,t}^{Max}, T^l\right), T^h\right) + \min\left(\max\left(T_{c,d,t}^{Min}, T^l\right), T^h\right) - T^l\right)\right], \tag{A8}$$

while SDD, labeled as S_c, is defined as

$$S_c = \frac{1}{10}\sum_{t\in Y_t}\sum_{d\in M_t}\left[0.5\left(\max(T_{c,d,t}^{Max}, T^k) + \max\left(T_{c,d,t}^{Min}, T^k\right)\right) - T^k\right] \tag{A9}$$

where $T^l = 10, T^h = 30, T^k = 32.2$, c denotes county, M_t denotes the set of days in the growing season ($M_t =$ days in $\{May, June, July, August\}$) in year t, and Y_t denotes the set of our sample weather data years ($Y_t = \{1998, 1979, \ldots, 2007\}$).

To construct county-level moisture variables, we need first to transfer the climate division data into county-level data because Palmer's Z indices are at the climate division level and different parts of a county might be covered by different climate divisions. We do so by calculating the ratio of county acres covered by each climate division intersect to obtain a weighting metric. We then multiply the monthly climate division PZ index by the associated weight and sum the products across all intersect climate divisions to obtain the monthly county-level PZ index, $PZ_{c,m,t}$. We define drought and wetness variables, as

$$D_c = \sum_{t\in Y_t}\sum_{d\in M_t} I_{c,m,t}(PZ_{c,m,t} < -2) \text{ and}$$
$$W_c = \sum_{t\in Y_t}\sum_{d\in M_t} I_{c,m,t}(PZ_{c,m,t} > 2.5), \tag{A10}$$

respectively, where m denotes month, $I_{c,m,t} = 1$ whenever the condition in the parentheses is satisfied and $I_{c,m,t} = 0$ otherwise. Thus, D_c measures the frequency of severe short-term drought occurrences in county c over 1998-2007, and W_c measures the frequency of severe short-term

wetness occurring in county c over the same time period.

Table A3 presents descriptive statistics for yield, climate, and land quality variables. The mean yearly unit yield is 149 bu./ac. and standard deviation is 39.4 bu./ac. The mean yearly county average yield is lower than the mean yearly unit yield, suggesting that our unit sample might have overweighted the number of high productivity farms. County yield also has a smaller standard deviation, as expected. The means of GDD and SDD are 1,290 and 14.7, respectively. The mean monthly severe drought incidence is 4.72 (out of 120) and the mean monthly severe wetness incidence is 5.53, suggesting that severe drought and severe wetness did not occur often in our sample counties over the sample period. The mean fraction of county land in categories LCC I or LCC II is 47.4%, showing that land is favorable for cultivation in much of the region under scrutiny.

Table A3. Descriptive statistics for yield and natural resource endowment variables

Variable	N	Mean	St. Dev	Min	Max
Unit yield (bu./ac.)	2,134,290	149	39.4	0	374
County yield (bu./ac.)	5,800	137	28.9	27	204
G_c, growing degree days	580	1,290	152	360	1,651
S_c, stress degree days	580	14.7	14.1	0.49	92.4
D_c (months in decade)	580	4.72	1.89	1	13
W_c (months in decade)	580	5.53	2.04	0	13
L_c (percent)	580	47.4	22.5	2.20	93.5

Notes: 1. The mean unit yield is the average of all unit-year yields over the 2,134,290 yearly unit observations, 1998-2007.
2. The mean county yield is the average of all county-year yields over the 5,800 yearly county observations, 1998-2007.
3. The means of county-level natural resource endowment variables are the averages over the 580 county observations.

Part VIII. Detrending yield data and adjusting APH yield

To estimate equation (12), we need to first obtain the yearly unit yield expectation, $\mu_{i,t}$, and the

yearly county yield expectation, $\mu_{c,t}$. It is commonly assumed that crop yield follows a time trend. As in Deng, Barnett and Vedenov (2007), we employ a log-linear model to estimate county yield trend for each county, namely

(A11) $$\ln(\tilde{y}_{c,t}) = \lambda_{c,0} + \lambda_{c,1}t' + v_{c,t},$$

where $t' = 2008 - t$ and $\lambda_{c,1}$ captures the backward trend in percent yield change starting from 2008. The NASS county yield data from 1958 to 2007 are used to estimate equation (A11) with estimates $\hat{\lambda}_{c,0}$ and $\hat{\lambda}_{c,1}$ while the county yield expectation for county *c* in year *t* is then calculated as $\mu_{c,t} = \hat{y}_{c,t} = \exp[\hat{\lambda}_{c,0} + \hat{\lambda}_{c,1}t']$.

RMA sets each unit's yearly reference yield expectation by use of the unit's APH yield. However, in our data, only a unit's 2008 APH yield is available. We proceed by assuming that the unit yield expectation inherits its time trend from the county yield expectation. Thus, the APH yields for 1998 to 2007 are then calculated as

(A12) $$APH_{i,t} = APH_{i,2008} * \exp[\hat{\lambda}_{c,1}t'],$$

where $t \in \{1998, 1999, \dots, 2007\}$.

A concern with using the APH yield as the expected unit yield is that the APH yield lags the true expected yield due to improved crop genetics and cultural practices (Plastina and Edwards 2014). To correct for this downward bias, RMA introduced trend-adjusted APH yields in 2012. The trend adjustment factor estimated for each crop in each county is equal to an estimate of the rate of the annual increase in the NASS county yield. This factor is then used to scale up past actual yield records. We use trend-adjusted APH yields as our individual yield expectations. Empirical results using adjusted and unadjusted APH yields are close.

Specifically, the county-year trend adjustment factor, $TA_{c,t}$, is estimated as

(A13) $$TA_{c,t} = (\hat{y}_{c,t} - \tilde{y}_{c,1958})/(t - 1958),$$

where $\hat{y}_{c,t}$ is the yield prediction for county c in year t as estimated by equation (A11), and

$\tilde{y}_{c,1958}$ is the 1958 actual yield for county c. After obtaining $TA_{c,t}$, the yield adjustment for the year t' before year t then equals $(t - t') \times TA_{c,t}$. To illustrate, suppose we seek to adjust the 2007 APH yield for unit i in county c with $TA_{c,2007} = 2$. Then the 2006, 2005, …, and 1997 actual yields for unit i will be adjusted up by 2, 4, … , and 20 bushels per acre, respectively. The 2007 APH yield would be adjusted up by $(2 + 4 + \cdots + 20)/10 = 11$ bushels per acre. The trend-adjusted APH yield formula can then be summarized as

$$TA\_APH_{i,t} = APH_{i,t} + \frac{1}{10}\sum_{j=1}^{10} j \times TA_{c,t}, \tag{A14}$$

and the trend-adjusted APH yield is then used as the expected unit yield in equation (11), i.e., $\mu_{i,t} = TA\_APH_{i,t}$.

Part IX. Systemic risk estimates using the RMA county average yield data

The 2018 Farm Bill has made the decision to base Agriculture Risk Coverage-County (ARC-CO) yields on the RMA data instead of the NASS county data, suggesting a more important role for the RMA data in the design of future county-based programs. Here we conduct an analysis where the NASS county yield is replaced by the acre-weighted RMA county average yield.

One problem with the RMA data is that it does not contain a unit's yield for the survey year. It does contain up to ten, possibly non-consecutive, prior year yield records. We use the previous year's yield record to calculate the previous year's average yield. For example, for the 2000 data we keep 1999 yield records and calculate its acre-weighted average yield as the 1999 RMA county average yield. Only actual yield type records are used for calculation and this screening procedure leaves us with 8,700 county-year observations for our 580 sample counties from 1993 to 2007, the time period for which our data are available.

Using the same detrending model described by equation (A11), we obtain the RMA county

yield expectations and unit yield predictions based on the RMA county yield trend. Figure A2 plots the RMA county average yield, the NASS county yield, the RMA expected county yield, and the NASS expected county yield from 1998 to 2007. We find that the RMA county average yield closely follows the NASS county yield while the RMA expected county yield and the NASS expected county yield have similar trends.

Based on this knowledge, we construct the 1958 RMA county yield as $\tilde{y}_{RMA,1958} = \frac{\tilde{y}_{NASS,1958}}{\tilde{y}_{NASS,1993}} * \tilde{y}_{RMA,1993}$ and use $\tilde{y}_{RMA,1958}$ to construct the RMA APH adjustment factor by the method reported in Part VIII in SM. Results for the RMA systemic risk estimates using the APH-adjusted RMA county data are reported in Table A4 and Figure A3. All results are very similar to their NASS counterparts. On average, the RMA systemic risk estimates are slightly greater than the NASS systemic risk estimates because unit-level data are more closely correlated with county-level data in the RMA dataset than in the NASS dataset.

Since each county in our RMA data only has yield records from 1993 to 2007, kernel estimations over such a short time period would be noisy even when augmenting with neighborhood counties' yield information. TRASR results would then also be noisy. However, since the RMA county yield and the RMA systemic risk estimates align closely with their NASS counterparts, we should expect that TRASR values estimated with the RMA county yield data also align closely with TRASR values estimated with the NASS county yield data.

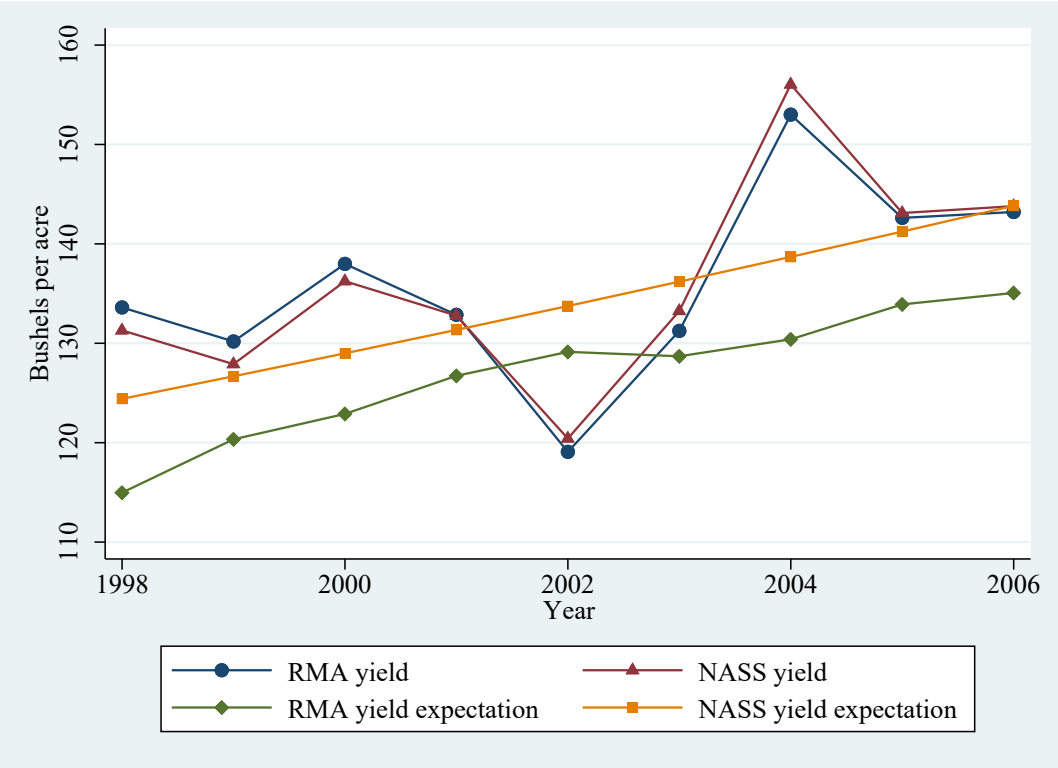

Figure A2. The RMA county average yield, the NASS county yield, the RMA expected county yield and the NASS expected county yield, 1998-2007.

Table A4. Descriptive statistics for estimates of systemic risk and systemic risk components using the RMA county yield data (N=213,429)

Variable	Mean	St. Dev	Min	Median	Max
Systemic risk, R_i^2	0.48	0.25	0.00	0.51	0.99
unit yield sensitivity, β_i	1.08	0.58	-4.77	1.05	6.69
Unit idiosyncratic revenue St. Dev, σ_{ε_i}	18.09	8.83	1.08	16.01	93.56
County revenue standard deviation, σ_c	16.62	6.02	5.87	15.70	42.41

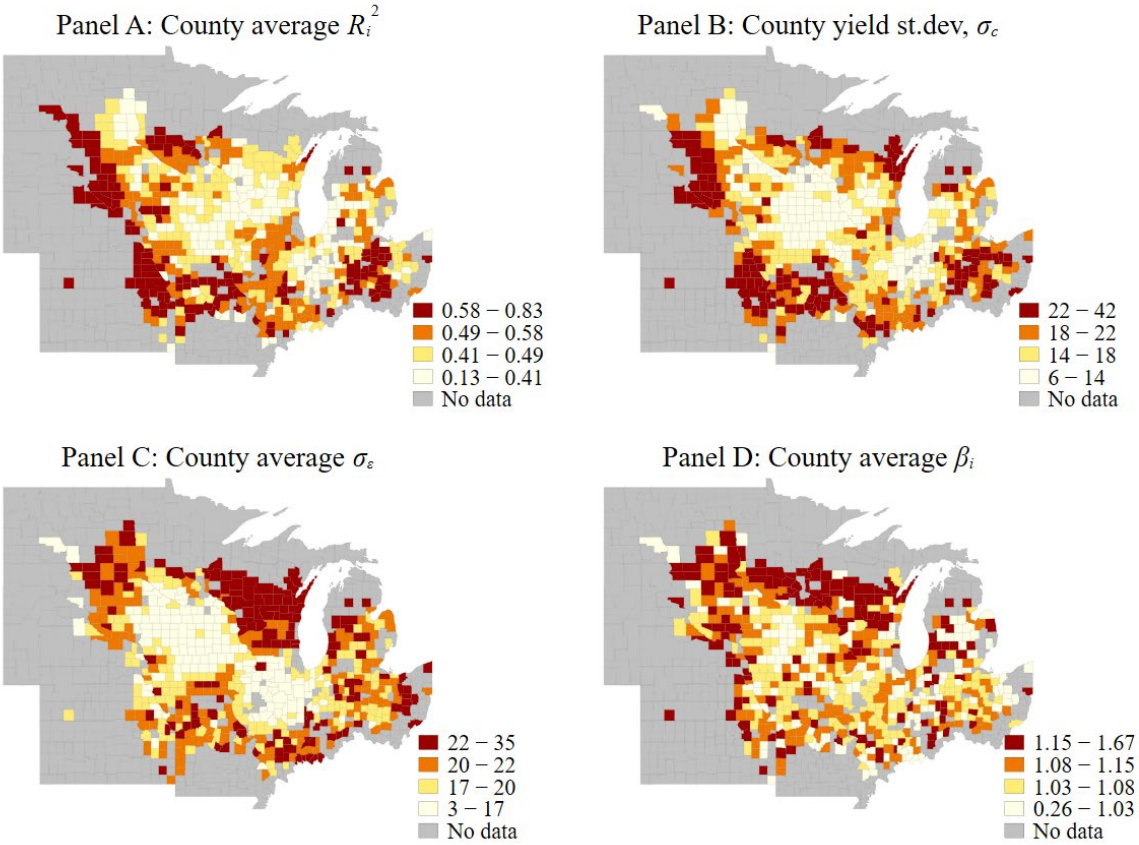

Figure A3. Geographic distributions of county average systemic risk and county average systemic risk components using the RMA county yield data

Note: Panels A to D respectively plot county averages of unit-level systemic risk, county yield standard deviation, unit-level idiosyncratic yield standard deviation and unit yield's sensitivity to county yield. Numbers in legends are quartile ranges.

Part X. Replacing GDD and SDD with the county average precipitation variable

Table A5. Regression results of systemic risk variables on natural resource endowments variables, with GDD and SDD replaced by the county average precipitation variable

	(1) $2\ln(\beta_i)$	(2) $2\ln(\sigma_{\varepsilon_i})$	(3) $2\ln(\sigma_c)$	(4) $2\ln(\sigma_c)$	(5) $-2\ln(\tau_i)$
$G_c/100$	-0.047[a]	0.065[a]	0.113[a]	0.104[a]	0.001
	(0.017)	(0.012)	(0.026)	(0.023)	(0.034)
$S_c/10$	0.082[a]	-0.144[a]	0.142[a]	0.139[a]	0.368[a]
	(0.020)	(0.020)	(0.036)	(0.027)	(0.047)
P_c (10 mm)	-0.002	0.001	-0.024[a]	-0.033[a]	-0.027[a]
	(0.003)	(0.002)	(0.005)	(0.004)	(0.006)
L_c (10 %)	0.006	-0.047[a]	-0.028[c]	-0.046[a]	0.024
	(0.011)	(0.006)	(0.015)	(0.012)	(0.020)
APH_i	-0.003[a]	-0.012[a]	-0.007[a]		0.002
	(0.001)	(0.001)	(0.001)		(0.001)
APH_c_CV				0.638[b]	
				(0.299)	
Constant	0.807[a]	6.935[a]	6.041[a]	5.653[a]	-0.090
	(0.280)	(0.212)	(0.331)	(0.287)	(0.442)
Obs.	213,320	213,320	213,320	579	213,320
R-squared	0.009	0.111	0.401	0.318	0.059

Notes: 1. The variable P_c denotes the ten-year average of county precipitation level from 1998 to 2007 in county c. To calculate P_c, we first calculate the daily precipitation level for county c as the average of daily precipitation levels recorded by all stations in county c in a given day. We then calculate the yearly precipitation level for county c by summing daily precipitation levels over the growing season for each year. Finally, we average county c's ten yearly precipitation levels to obtain P_c.
2. The variable APH_c_CV denotes the coefficient of variation of units' 2008 APH values in county c, which captures the dispersion of land quality in a county and thus reflects land heterogeneity within a county.
3. Standard errors in parenthesis are clustered at the county level; [a], [b], and [c] denote significance at 0.01, 0.05, and 0.1 levels, respectively.

Part XI. Kernel density estimation

Since kernel density estimation requires a stationary data series, before estimating a distribution we need to detrend and then retrend county yield data to a reference year: 2008 in our case. Detrending follows equation (A11) while the retrended yield for county *c* in year *t* is given by

(A15) $$\breve{y}_{c,t} = \frac{\tilde{y}_{c,t}}{\hat{y}_{c,t}} \times \hat{y}_{c,2008} \quad \forall t \in \{1958, 1959, \ldots, 2007\}$$

where $\hat{y}_{c,2008}$ and $\hat{y}_{c,t}$ are county yield predictions obtained from equation (A11), and $\tilde{y}_{c,t}$ is the observed county yield in year t.

The kernel density function for county c is given by

(A16) $$f(x_c) = \frac{1}{\sum_{t=1958}^{2007} w_{c,t}} \sum_{t=1958}^{2007} \frac{w_{c,t}}{h_c} K\left(\frac{x_c - \breve{y}_{c,t}}{h_c}\right),$$

where x_c is a specific point whose density is to be evaluated, the $\breve{y}_{c,t}$s are retrended county yields located within a pre-selected bandwidth centered at x_c, h_c is the bandwidth parameter, $K(\cdot)$ is the kernel function, and the $w_{c,t}$'s are the associated sample weights.

Researchers have reached a general consensus that the kernel function choice is less important than the bandwidth choice in kernel estimation (Parzen 1962; Tapia and Thompson 1978; Newton 1988). Thus, we choose the Epanechnikov kernel function because it is most efficient in minimizing the mean integrated squared error, which is the most common optimality criterion used for selecting bandwidth. We follow Goodwin and Ker (1998) by adopting Silverman's modified rule-of-thumb method to select the bandwidth parameter, the formula of which is given by

(A17) $$h_c = \frac{0.9}{n^{0.2}} \min\left(\sigma_{\breve{y}_c}, \frac{IQR_{\breve{y}_c}}{1.349}\right),$$

where $\sigma_{\breve{y}_c}$ is the standard deviation of $\breve{y}_{c,t}$ and $IQR_{\breve{y}_c}$ is the interquartile range of $\breve{y}_{c,t}$.

Since we only have fifty-year observations for each county, kernel estimation may be imprecise. To expand the data pool, we follow Goodwin and Ker (1998) by using yield information from adjacent counties (counties having a positive length boundary with the central county). To qualify for the calibration, the central county must have at least three adjacent counties with no missing yield records over 1958-2007. This screen ensures at least 200 yearly

yield records to estimate the central county's yield distribution. The requirement leaves us with a sample comprised of 207,230 units in 571 counties. Each yield record in adjacent counties is assigned a weight of $1/[(2N+1)\times 50]$ while each yield record in the central county is assigned a weight of $(N+1)/[(2N+1)\times 50]$, where N is the number of adjacent counties.

After estimating the density, $\hat{f}(x)$, for each point, we then turn to evaluating the cumulative density at each point, $\hat{F}(x)$. We do so by applying the trapezoid rule, in the form

$$\text{(A18)} \qquad \hat{F}(x) = \int_0^x \hat{f}(\tilde{y}_c)d\tilde{y}_c \approx \sum_{k=1}^{K} \frac{[\hat{f}(\tilde{y}_{c,k-1}) + \hat{f}(\tilde{y}_{c,k})]}{2}\Delta\tilde{y}_{c,k},$$

where $\Delta\tilde{y}_{c,k} = \tilde{y}_{c,k} - \tilde{y}_{c,k-1}$ and K is the total number of density-evaluation points at or below x. The two integrals arising in the TRASR estimate, $\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c$ and $\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c$, are also evaluated by the trapezoid rule in this way.

As for the estimation of each unit's idiosyncratic yield distribution, $G(\varepsilon_i)$, since we only have ten estimates of idiosyncratic yield for each unit and are unable to access unit adjacency information, we are unable to perform unit-level kernel estimations. Rather, within each county we pool units with similar 2008 APH yields into a group and then perform group-level nonparametric kernel estimations. By doing so, we are assuming that units with similar 2008 APH yields in the same county also have similar idiosyncratic yield distributions.[4]

Empirically, within each county the 2008 APH yields are grouped in tens. That is, units with the smallest to the tenth smallest 2008 APH yields are assigned to group 1, those with the eleventh to twentieth smallest 2008 APH yields are assigned to group 2, and so on. Whenever the last group contains fewer than ten APH values, these units are merged with the second last group. This grouping method generates 2,426 unique APH groups in total. As shown by Table A6, 50 counties have only one APH group and 70 have only two APH groups. Units in these counties are more likely to be assigned into APH groups where the idiosyncratic yield

distributions of some units differ considerably from those of other units. However, as Figure A4 shows, since counties with fewer APH groups tend to have smaller APH ranges (the Pearson's correlation coefficient between the number of APH groups and the 2008 APH range is 0.39), bias caused by pooling units with different idiosyncratic yield distributions into the same APH group should be moderate. Moreover, among the 2,426 APH groups, only 7.9% have fewer than 200 unit-year observations. Thus, the nonparametric kernel density estimation for most APH groups is largely free from imprecision concerns.

As with county yield kernel density estimation, we also choose the Epanechnikov kernel function and adopt Silverman's modified rule-of-thumb method to select the bandwidth parameters. After obtaining the density function, the cumulative density function, $G(\varepsilon_i)$, and the integral $\int_{\underline{\varepsilon}_i}^{\bar{\varepsilon}_i} \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)$, are evaluated by the trapezoid rule.

Table A6. Number of counties with different number of APH groups

Number of APH groups	Number of counties	Percent of counties with associated APH groups
1	50	8.8
2	70	12.2
3	63	11.0
4	93	16.3
5	134	23.5
6	125	21.9
7	34	5.9
8	1	0.2
9	1	0.2

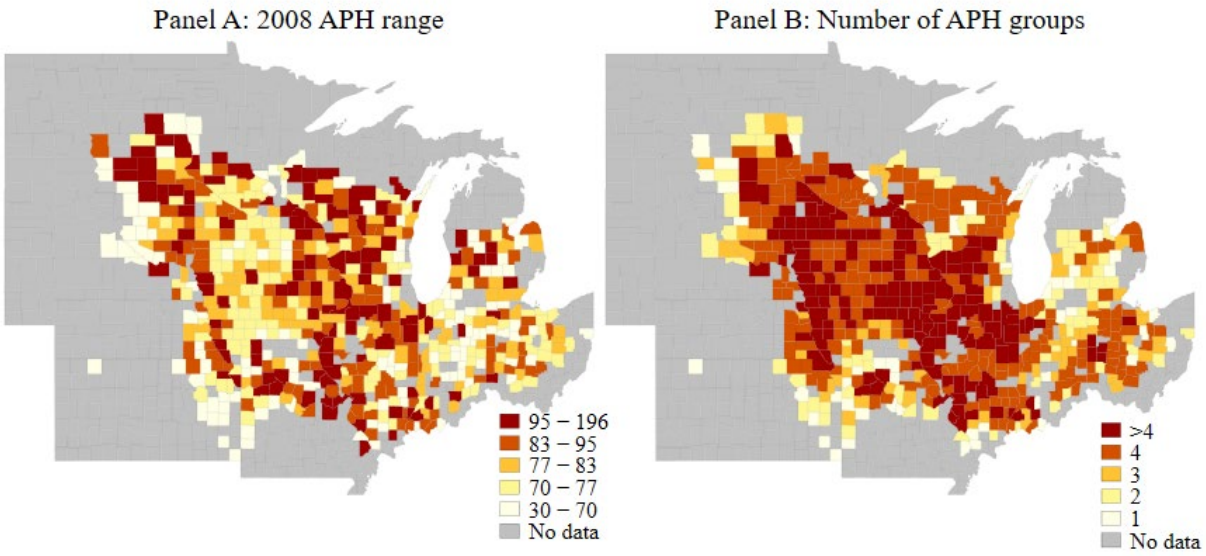

Figure A4. Geographic distributions of 2008 APH range and number of APH groups
Note: The 2008 APH range equals the maximum 2008 unit APH yield minus the minimum 2008 unit APH yield in a county. Numbers in Panel A's legend are quintile ranges.

Part XII. Premium subsidy rates for individual and area insurance plans

Table A7. Premium subsidy rates and relative subsidy rates for individual and area insurance products

Insurance product	Coverage Level (%)									
	CAT	50	55	60	65	70	75	80	85	90
	Panel A: Post-2009 subsidy rates									
Basic and Optional Units	100	67	64	64	59	59	55	48	38	n/a
Enterprise Units (EU)	n/a	80	80	80	80	80	77	68	53	n/a
Area Yield Plans (AYP)	n/a	n/a	n/a	n/a	n/a	59	59	55	55	51
Whole Farm Units	n/a	80	80	80	80	80	80	71	56	n/a
AYP/EU, $s_{c\|i}$	**n/a**	**n/a**	**n/a**	**n/a**	**n/a**	**0.74**	**0.77**	**0.81**	**1.04**	**n/a**
100%/EU, $\bar{s}_{c\|i}$	**n/a**	**n/a**	**n/a**	**n/a**	**n/a**	**1.25**	**1.30**	**1.47**	**1.89**	**n/a**
	Panel B: Pre-2009 subsidy rates									
Basic and Optional Units	100	67	64	64	59	59	55	48	38	n/a
Enterprise Units (EU)	n/a	67	64	64	59	59	55	48	38	n/a
Area Yield Plans (AYP)	n/a	n/a	n/a	n/a	67	64	64	59	59	55

AYP/EU, $s_{c\|i}$	**n/a**	**n/a**	**n/a**	**n/a**	**n/a**	**1.08**	**1.16**	**1.23**	**1.55**	**n/a**
100%/EU, $\bar{s}_{c\|i}$	**n/a**	**n/a**	**n/a**	**n/a**	**n/a**	**1.69**	**1.82**	**2.08**	**2.63**	**n/a**

Notes: 1. Post-2009 subsidy rates are as reported in Shields (2015).
2. Area Yield Plans in the pre-2009 period is Group Risk Income Protection (GRIP) and the pre-2009 subsidy rates are from Wang, Hanson, and Black (2003).
3. Pre-2009 EU subsidy rates are from Johansson, Worth, and Cooper (2014).
4. The second last row in each panel reports the relative subsidy rates of AYP plans over YP plans for the same coverage levels under which both insurance plans are available. The last row in each panel reports the 100% AYP subsidy rate over the EU subsidy rate for available AYP coverage levels.
5. Whole farm unit plans were not available in the pre-2009 period.

Part XIII. An alternative approach to calibrating TRASR values

An alternative to the method used in the main text to calibrate TRASR values is to use the RMA premium data to directly calculate TRASR values by way of equation (8). In this subsection we do so as a robustness check.

Assuming the RMA premiums are actuarially fair, in light of equation (8), TRASR levels can be calculated as

(A19) $$s^*_{c|i} = \mathrm{E}(\tilde{n}_i)/\mathrm{E}(\tilde{n}_c) = \pi_i/\pi_c.$$

That is, the TRASR value for unit i in county c equals unit i's per-acre YP premium over county c's per-acre AYP premium. Since the RMA premium data are at the county level, we can only calculate the county-level TRASR values using this dataset. To do so, for each year and each county, we first calculate the per-acre YP premium as the ratio of total YP premiums over total YP insured acres and the per-acre AYP premium as the ratio of total AYP premiums over total AYP insured acres. For each county, the yearly TRASR value is then obtained as the ratio of the per-acre YP premium over the per-acre AYP premium in that year. The yearly TRASR values are then averaged over 2002 (the first year that the RMA coverage-level-conditional premium data are available) to 2007 (the last year that our RMA unit yield data cover) to arrive at the final TRASR value for each county.

Panel A in Table A8 reports summary statistics for county TRASR values obtained from the RMA premium data. As a comparison, in panel B we also list summary statistics for county mean TRASR values obtained from our RMA unit yield data. Results show that mean TRASR values obtained by the RMA premium data are generally significantly larger than their yield data counterparts. However, panels A and B are in other ways largely consistent. For example, in both panels, all statistics decrease with an increase in the YP coverage level and increase with an increase in the AYP coverage level. A recent body of research has called into question the accuracy of RMA's unit-level rate-setting procedures (Ramirez and Shonkwiler 2017; Price et al. 2019; Maisashvili et al. 2019), so moderate discrepancies between the two approaches are not surprising.

Panels A through C of Figure A5 below depict the geographic distributions of county TRASR values obtained from the RMA premium data while panels D through F depict the geographic distributions of county mean TRASR values obtained from our RMA unit yield data. Although some differences exist, the general patterns are similar for each YP coverage level pair. Counties in the northern and eastern fringes of the Corn Belt generally have higher TRASR values. For each YP coverage level pair, the Pearson's correlation test and the Spearman's rank correlation test both report significantly positive correlation coefficients between the two measures of county TRASR values. The coefficient values are about 0.4, confirming that TRASR values obtained from these two different approaches display similar geographic distribution patterns.

Table A8. Descriptive statistics for county TRASR values

ϕ_i	ϕ_c	N	Mean	St.Dev	Min	Median	Max
Panel A: County TRASR values obtained from the RMA premium data							
	70%	200	486	296	96	416	1,622
	75%	131	336	225	76	263	1,512
75%	80%	201	252	156	86	202	933
	85%	241	171	101	59	141	783
	90%	351	118	60	39	101	413
	70%	156	587	366	131	505	2,508
	75%	106	372	201	103	322	1,451
80%	80%	165	296	162	123	253	1,081
	85%	198	197	96	81	172	738
	90%	280	139	59	49	125	428
	70%	124	696	390	176	609	2,136
	75%	89	479	252	134	418	1,562
85%	80%	126	375	212	163	303	1,155
	85%	154	237	103	82	213	649
	90%	216	166	67	63	154	477
Panel B: County mean TRASR values obtained from the RMA unit yield data							
	70%	200	304	279	63	202	1917
	75%	131	173	126	48	135	896
75%	80%	201	120	73	42	97	483
	85%	241	77	37	31	67	278
	90%	351	54	24	22	48	169
	70%	156	413	653	78	253	7285
	75%	106	202	125	65	167	1009
80%	80%	165	146	83	55	120	560
	85%	198	94	41	41	84	328
	90%	280	66	25	33	60	184
	70%	124	434	364	99	301	2159
	75%	89	261	158	87	220	1206
85%	80%	126	181	97	73	151	670
	85%	154	119	44	55	108	351
	90%	216	82	25	44	78	220

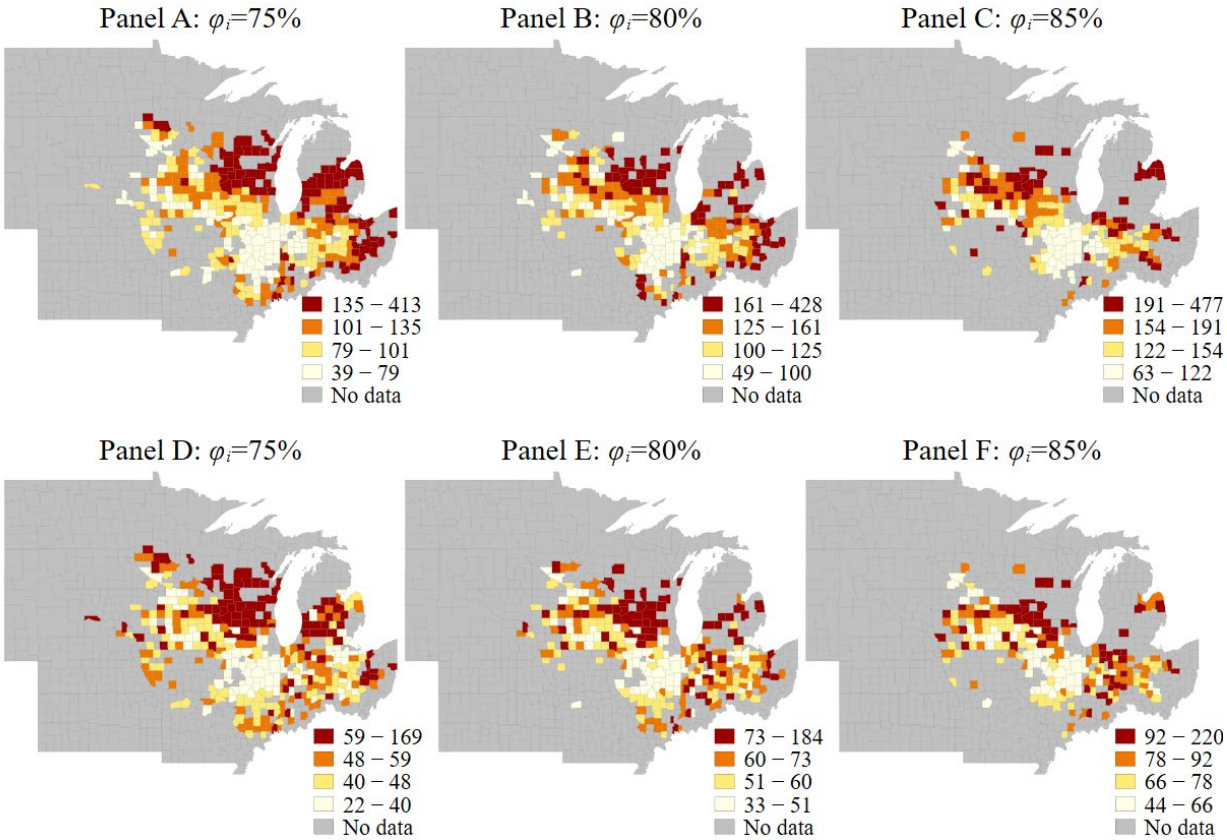

Figure A5. Geographic distributions of county TRASR values, conditional on $\phi_c = 90\%$
Notes: Panel A to C depict the geographic distributions of county TRASR values obtained by the RMA premium data and Panel D to F depict the geographic distributions of county mean TRASR values obtained from the RMA unit yield data. For each coverage level combination, only counties appearing in both datasets are depicted and numbers in legends are quartile ranges in %.

Footnotes

---

[1] This assumption is reasonable as we find that the maximum acre share taken by a unit in a county is only 6.2% in any sample year, while acre shares of 99% of unit-year observations are less than 1.1%.

[2] In studies that use the RMA unit-level data, keeping yield records only in the immediately preceding ten years is a common practice, see Deng, Barnett and Vedenov (2007) and Claassen and Just (2011).

[3] Using the RMA unit-level data, the county irrigation rate is defined as the ratio of insured units that irrigate over the total number of insured units in that county. Dropping counties whose irrigation rate exceeded 30% generates similar results.

[4] In projecting individual rate from county rate, RMA adopts a yield ratio (Ry) to reflect unit-level risk heterogeneity. Coble et al. (2010), p. 36, rationalize this choice as follows: "The use of Ry implies that an individual's premium rate can be reasonably based on the magnitude of his/her own historical rate yield relative to the county reference yield." Since Ry = (unit APH yield) / (county reference yield), and the county reference yield is the same for all units in the same county, units with similar APH yields in the same county should have similar premium rates and should be treated as having similar risks. As area risk is common across all units in the same county, the remaining idiosyncratic risks should be similar across units with similar APH yields in the same county. Thus, it is reasonable to include units with similar APH yields in the same idiosyncratic risk group when performing unit-level kernel estimations.

References

Claassen, Roger, and Richard E. Just. 2011. "Heterogeneity and Distributional Form of Farm-Level Yields." *American Journal of Agricultural Economics* 93(1): 144–160. https://doi.org/10.1093/ajae/aaq111.

Coble, Keith H, Thomas O Knight, Barry K Goodwin, Mary Frances Miller, Roderick M Rejesus, and George Duffield. 2010. "A Comprehensive Review of the RMA APH and COMBO Rating Methodology: Final Report." Prepared by Sumaria systems for the risk Management agency. https://www.researchgate.net/publication/228651550.

Deng, Xiaohui, Barry J. Barnett, and Dmitry v. Vedenov. 2007. "Is There a Viable Market for Area-Based Crop Insurance?" *American Journal of Agricultural Economics* 89(2): 508–519. https://doi.org/10.1111/j.1467-8276.2007.00975.x.

Goodwin, Barry K., and Alan P. Ker. 1998. "Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts." *American Journal of Agricultural Economics* 80(1): 139–153. https://doi.org/10.2307/3180276.

Johansson, Robert C., Tom Worth, and Joe Cooper. 2014. "Farm-level choice of crop insurance coverage level: A preliminary assessment". Crop Insurance and the 2014 Farm Bill Symposium, Louisville, KY, October 8-9, 2014.

Maisashvili, Aleksandre, Henry Bryant, George Knapek, and James Marc Raulston. 2019. "Are Crop Insurance Premium-Implied Yield Distributions Valid?" *Agricultural Finance Review* 79(4): 467–490. https://doi.org/10.1108/AFR-09-2018-0077.

Miranda, Mario J. 1991. "Area-Yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73(2): 233–242. https://doi.org/10.2307/1242708.

Newton, H. Joseph. 1988. *Timeslab: A Time Series Analysis Laboratory*. Wadsworth Publishing Company.

Parzen, Emanuel. 1962. "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics* 33(3): 1065-1076. https://doi.org/10.1214/aoms/1177704472

Plastina, Alejandro, and William Edwards. 2014. "Trend-Adjusted Actual Production History (APH)." Ag Decision Maker, Extension and Outreach, Iowa State University. September 2014. https://www.extension.iastate.edu/agdm/crops/html/a1-56.html.

Price, Michael J., Cindy L. Yu, David A. Hennessy, and Xiaodong Du. 2019. "Are Actuarial Crop Insurance Rates Fair?: An Analysis Using a Penalized Bivariate B-Spline Method." *Journal of the Royal Statistical Society. Series C: Applied Statistics* 68(5): 1207–1232. https://doi.org/10.1111/rssc.12363.

Ramirez, Octavio A, and J Scott Shonkwiler. 2017. "A Probabilistic Model of the Crop Insurance Purchase Decision." *Journal of Agricultural and Resource Economics* 42(1): 10–26.

Shields, Dennis A. 2015. "Federal Crop Insurance: Background." CRS Report for Congress, Congressional Research Service, 7-5700, R40532. Washington, DC.

Tapia, Richard A, and James R Thompson. 1978. *Nonparametric Probability Density Estimation*. Johns Hopkins University Press.

Wang, H. Holly, Steven D. Hanson, and J. Roy Black. 2003. "Efficiency costs of subsidy rules for crop insurance." *Journal of Agricultural and Resource Economics,* 28(1): 116-137.