TOWARDS UNDERSTANDING CROP YIELD SYSTEMIC RISK AND ITS IMPLICATION FOR CROP INSURANCE CHOICES

By

Xuche Gong

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agriculture, Food and Resource Economics – Master of Science

2019

ABSTRACT

TOWARDS UNDERSTANDING CROP YIELD SYSTEMIC RISK AND ITS IMPLICATION FOR CROP INSURANCE CHOICES

By

Xuche Gong

Area based insurance contracts have long been offered to crop producers as an option for risk management. However, the take-up rate for such programs remains low. In this paper, utilizing RMA unit-level corn yield data and NASS county-level corn yield data, we investigate roles of systemic risk and premiums subsidies in producers' choices between area and individual insurance contracts. We find that, on average, systemic risk explains slightly more than one third of total unit yield variability. Systemic risk is high in the Southern and Western Corn Belts and its geographic distribution matches well the geographic distribution of county yield variance. Systemic risk increases with both beneficial and stressful heat accumulations, frequency of drought, and land quality. We also study the lower bound on subsidy rate for area insurance when normalized by that for individual insurance such that the expected net returns to area yield insurance equals the expected net return of individual yield insurance. We find that this lower bound is negatively correlated with systemic risk. Producers in high systemic risk counties will require fewer subsidies to possibly choose area insurance over individual insurance. Moreover, we find that were transfer maximization a producer's only concern then the current area subsidy rate might be a major deterrent for producers to choose low coverage level area insurance. Raising the area insurance subsidy rate might be a feasible option to induce more area insurance demand because the transfer-equalizing area insurance subsidy rate exceeds 100% for only a small fraction of producers.

I dedicate this work to my parents and grandparents

ACKNOWLEDGEMENTS

I would like to thank my major professors, Dr. Hongli Feng, and my co-major professor, Dr. David A. Hennessy, for their unremitting guidance and encouragement. This work cannot be done so accomplished without their work. I would also like to thank my other committee members, Dr. Robert Myers and Dr. Kyoo il Kim. They have provided me many useful comments and suggestions when I was writing the paper. I also want to thank my AFRE classmates and my Chinese friends. I cannot go through these two years without their support. Finally, I want to send my deepest gratitude to my parents for their support in my study abroad.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Since its inception in the 1930s, the Federal Crop Insurance Program (FCIP) has continuously supplied agricultural producers in the United States with insurance products to manage production risks. The program has undergone especially rapid expansion since the enactment of the Federal Crop Insurance Reform Act of 1994, which substantially increased premium subsidy levels. In 2018, crop insurance covered about 85.6% of planted acres for the ten major crops and the ratios were even higher for corn, soybeans and wheat[1], providing agricultural producers with a solid safety net. However, the high participation rate comes at a cost. According to the Risk Management Agency (RMA), FCIP has accounted for about $7.3 billion per annum direct costs over the period of 2009 through 2018, making it the most expensive agricultural commodity program in the U.S.[2] In addition, studies have found that these significant subsidy levels have led to some moral hazard problems, such as changes in farmers' input use, crop choice, and acreage decision (Quiggin et al. 1993; Smith and Goodwin 1996; Babcock and Hennessy 1996; Goodwin et al. 2004; Goodwin and Smith 2013; Yu et al 2017), and have caused substantial deadweight loss imposed on taxpayers (Lusk 2016). As a result, the federal government has long sought to reduce insurance program costs. A widely proposed option is for government to subsidizes only the systemic risk part, leaving the rest to the market (Miranda 1991; Miranda and Glauber 1997; Coble and Barnett 2008; Dismukes et al. 2010; Goodwin and Smith 2013; Congressional Budget Office 2017).

---

[1] The ten major crops are: corn, cotton extra-long staple, cotton upland, oats, rice, sorghum, soybeans, sugar beets, sugarcane, and wheat. Planted acres data are from USDA's 2008 Acreage Survey and failed acres are also included. Insurance data are from RMA's 2018 Summary of Business.

[2] Direct cost equals total premiums paid by farmers minus underwriting gains paid to Approved Insurance Providers and total indemnities paid to farmers. Data are from RMA "Direct Costs of Federal Crop Insurance Program" at https://www.rma.usda.gov/-/media/RMAweb/AboutRMA/Program-Budget/18cygovcost.ashx?la=en.

Systemic risk in crop insurance markets refers to strong correlation between individual losses and stems from the fact that yield losses are generally driven by natural disasters that affect a wide range of farms within a given region (Miranda and Glauber 1997). The existence of systemic risk justifies a government's subsidies in crop insurance market to the extent that this risk form undermines an insurers' ability to diversify risk across individuals, forces insurers to set premiums at prohibitively high levels, and eventually leads the insurance market to breakdown (Duncan and Myers 2000). However, some studies have also found that if government sufficiently refunds loss caused by systemic risk via some form of area based reinsurance contracts or commodity programs, then private insurers might be capable of dealing with the remaining idiosyncratic risks even in the presence of asymmetric information problems, as occurs in other property insurance markets (Miranda and Glauber 1997; Duncan and Myers 2000; Coble and Barnett 2008). The Congressional Budget Office (2017) has also proposed to subsidize only area-based insurance products as one option to reduce the budgetary costs of crop insurance programs. The potential for efficiently separating systemic risk from idiosyncratic risk depends on our ability to characterize and measure systemic risk.

Systemic risk is also closely related to demand for area-based insurance. Unlike individual-based insurance, area-based insurance makes indemnity payments based on easily observed area yield or revenue loss, which are generally not influenced by the insured. Thus, costs caused by information asymmetry and high administration costs can be substantially reduced (Halcrow 1949; Miranda 1991; Smith et al. 1994; Skees et al. 1997; Mahul 1999; Vercammen 2000; Barnett et al 2005; Shen and Odening 2013). As a result, the USDA introduced the first area yield insurance program and area revenue program in 1990s, and the 2014 Farm Bill further introduced several county-level revenue-based insurance plans (Goodwin and Hungerford

2014). However, demand for area-based insurance remains low. As shown by Figure 1, the share of area insurance insured acres in total insured acres was less than 5% in most years and was less than 20% even at its peak year, 2006. Experiences from area-based insurance programs and weather index programs in other countries suggest that the existence of basis risk might be an important deterrent to participating in area-based insurance programs (Elabed et al. 2013; Shen and Odening 2013; Clarke 2016; Hill et al. 2016; Jensen et al. 2018). Basis risk exists when individual yield is poor but area average yield is good. Basis risk decreases when the correlation between individual loss and area loss increases, i.e., when systemic risk increases. Thus, precise estimates of systemic risk should also help to identify whether basis risk is the major deterrence to the demand for area-based insurance in the United States.

However, given the importance of estimating systemic risk in controlling crop insurance costs, comparatively few studies have investigated it. Applying a large size farm-level yield data, Barnett et al. (2005) estimated correlations between farm and county corn yields to range between 0.36 (in Michigan) and 0.82 (in Illinois). Using simulated data, Dismukes et al. (2010) reported average nationwide farm-state revenue correlations to be 0.55 for corn, 0.54 for soybeans, and 0.39 for cotton. Claassen and Just (2011) found that systemic variation explains about 48% of farm yield variation for corn in Illinois and 40% for wheat in Dakotas. However, these studies only investigated the correlation between individual yield and area yield. From the perspective of risk management, a more relevant issue is the correlation between individual loss and area loss. As the only exception, Zulauf et al. (2013) examined farm-level yield and revenue loss that is systemic with yield and revenue loss at the county, state, and nation levels. They found farm loss systemic with area loss to be large, and to decline as the geographical aggregate level increased. However, Zulauf et al. (2013) just simply calculated the ratio of the yield and revenue loss at high geographical aggregate levels over the yield and revenue loss at farm level, and did not model the relationship between farm loss and area loss.

In this article, we apply Miranda's (1991) single-factor capital market model to develop a novel theoretically grounded approach to measuring and decomposing systemic risk. We also decompose this systemic component into three components. In his innovative study, Miranda (1991) decomposed farm-level yield variability into a systemic component that is correlated with area yield and an idiosyncratic component that is uncorrelated with area yield. Miranda (1991) measured systemic risk by the beta, which measures the sensitivity of unit yield deviations from expected value to area yield deviations from expected value. This has become a workhorse procedure for crop insurance analysis, and also farm-level policy studies more generally because farm-level data are generally unavailable and the single factor stochastic structure is very useful when simulating farm-level data from county-level data (Mahul 1999; Barnett et al. 2005; Coble and Dismukes 2008; Carriquiry et al. 2008; Cooper et al. 2012).

However, although the beta captures well the co-movement between farm yield and county yield, it does not convey the relative importance of systemic risk with respective to idiosyncratic risk, which is particularly important when determining risk structure in the crop insurance market. In this article, we characterize systemic risk as the proportion of unit yield variation that can be explained by county yield related variation. This measurement, the R^2 statistic, is always bounded between 0 and 1. As we will show, the statistic is determined not only by *i*) the beta, but also *ii*) county yield variance and *iii*) a farm's idiosyncratic yield variance.

Using a large-scale unit-level dataset, we estimate county-level systemic risk for 589 counties across the Midwest. We find that, on average, systemic risk explains about one-third of unit yield variation and is larger in U.S. Southern and Western Corn Belt counties. Moreover, we find that the geographic distribution of systemic risk well matches the geographic distribution of county yield variance. This finding is consistent with a strand of literature that finds yield correlations to be higher in extreme weather years (Okhrin et al. 2013; Goodwin and Hungerford 2014; Tack and Holt 2016; Du et al. 2017), as counties with higher yield variance generally experience more adverse weather shocks.

In addition to measurement and decomposition, we extend our analysis framework to investigate how systemic risk is determined by county growing conditions. Du et al. (2017) proposed and empirically tested a model linking county-level yield-yield dependence with growing conditions. They found a substitution effect between soil and benign water availability levels, but a complementarity effect between soil and beneficial heat variables. Our work extends theirs by investigating how farm-county yield-yield correlation varies with county weather conditions and soil quality. We find that systemic risk is significantly increasing in county heat accumulations (both beneficial and stressful) and drought appearance. Land quality also has a significantly positive effect on systemic risk. Our study also investigates how growing condition variables affect systemic risk through each of the three systemic risk components and how these mediation effects reinforce or counteract each other.

A policy issue that cannot be separated from the character of systemic risk is how to set cost-effective subsidies for area-based programs when subsidized individual yield insurance contracts are also available. As shown in Figure 1, a substantial decline in the share of area insured acres occurred in 2009. This decline coincided with a reduction in area insurance premium subsidy rates and increase in enterprise unit insurance premium subsidy rates in compliance with the 2008 Farm Bill. The above observation provides anecdotal evidence on the importance of premium subsidies rates on area-based insurance demand. However, though many studies have found that crop insurance demand is sensitive to subsidy levels, most of these have focused on individual-based insurance (Coble et al., 1996; Goodwin et al. 2004; Shaik et al., 2008; O’Donoghue, 2014; Du et al., 2016).

Among the few exceptions, Deng et al. (2007) found that after considering the large premium wedge between individual yield insurance and premium subsidies, area yield insurance was preferred to individual yield insurance by cotton producers but not by soybeans producers. In a theoretical mean-variance preference model, Bulut et al. (2012) found that whenever premium rates for area insurance and individual insurance are both actuarially fair then producers should demand full individual insurance and no area insurance. If area insurance is fully subsidized, then area insurance might replace a portion of individual insurance. Although these studies have documented the importance of subsidies in determining area insurance demand, questions remains open as to whether current area insurance subsidy rates provide producers with sufficiently high compensation for their risk exposure under area-based insurance and whether raising area insurance subsidy rates is a viable option for increasing area insurance demand.

Moreover, no study has explored the relationship between systemic risk and the effective area insurance subsidy rate that will induce producers to possibly choose area insurance over individual insurance. Areas with a high systemic risk and low effective area insurance subsidy rate are ideal places to grow area insurance demand because area insurance provides growers in these regions with comparatively good risk protection and they require lower subsidy levels to participate in the program.

Recognizing that individual insurance provides better risk coverage in comparison with area insurance, in this paper the task we set ourselves is to calibrate the threshold relative area subsidy rate (TRASR) at which individual insurance and area insurance provide the same expected level of transfer the grower. Thus, below TRASR even risk-neutral growers will not choose area insurance over individual insurance. TRASR thus provides the lower bound on the ratio of area insurance subsidy rate over individual insurance subsidy rate needed to possibly induce producers to choose area insurance over individual insurance. We then compare TRASR to the ratio of the current area insurance subsidy rate over the current individual insurance subsidy rate to ascertain whether the current subsidy rate structures deter producers from choosing area insurance. We also check whether an area insurance subsidy rate no less than 100%, i.e., providing free area insurance or better, is a necessary condition for most producers to choose area insurance over individual insurance. Finally, we relate TRASR to systemic risk in order to establish whether there exists a good match between the risk protection function of area insurance and the effective area insurance subsidy rate.

We find that while the current insurance subsidy rate structure does deter most producers from choosing low coverage level area yield insurance contracts, this is not true for high coverage level area yield insurance contracts. For most producers, the minimum required area insurance subsidy rate to possibly choose area insurance over individual insurance is less than 100%, suggesting that when transfer maximization is the grower's only concern then raising area insurance subsidy rates to levels lower than 100% might be a feasible option to induce more area insurance demand. We also find a negative correlation between systemic risk and TRASR, which supports our belief that high systemic risk counties are indeed ideal areas to implement area insurance.

CHAPTER 2. CONCEPTUAL FRAMEWORK

In this section, we present the conceptual framework about how we model systemic risk and how we calibrate the TRASR. Some important propositions are also derived.

2..1 Modeling Systemic Risk

Our focus is on yield risk and so we assume throughout that price is non-random. In order to avoid unnecessary notation, we set output price equal to 1 and ignore it henceforth. Following Miranda's (1991) one factor capital market model, we apply the following model to characterize the relationship between unit yield and county yield,

(1) $$\tilde{y}_i = \mu_i + \beta_i(\tilde{y}_c - \mu_c) + \theta_i\varepsilon_i.$$

Here, $\tilde{y}_i$ and $\tilde{y}_c$ are individual yield and county yield variables, respectively, while $\mu_i = \mathrm{E}(\tilde{y}_i)$, $\mu_c = \mathrm{E}(\tilde{y}_c)$, $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{Var}(\varepsilon_i) = \sigma_\varepsilon^2 = 1$, and $\mathrm{Cov}(\tilde{y}_c, \varepsilon_i) = 0$.

By way of equation (1), we decompose the unit yield deviation from expectation into a systemic component, $\beta_i(\tilde{y}_c - \mu_c)$, which is correlated with county yield, and an idiosyncratic part, $\theta_i\varepsilon_i$, which is uncorrelated with county yield. The coefficient β_i measures the sensitivity of unit yield deviations from expectation to county yield deviations from expectation. Since $\beta_i \leq 0$ is uncommon for crop production, we assume that $\beta_i > 0$ throughout the article. Also, since $\mathrm{E}(\varepsilon_i) = 0$ and $\sigma_\varepsilon^2 = 1$, we can set $\theta_i \geq 0$ without loss of generality. Now $\theta_i^2 = \mathrm{Var}(\theta_i\varepsilon_i)$ measures the idiosyncratic part of unit yield variance while $\sigma_i^2 = \mathrm{Var}(\tilde{y}_i - \mu_i) = \beta_i^2\mathrm{Var}(\tilde{y}_c - \mu_c) + \theta_i^2\mathrm{Var}(\varepsilon_i) = \beta_i^2\sigma_c^2 + \theta_i^2$, where $\sigma_i^2 = \mathrm{Var}(\tilde{y}_i)$ and $\sigma_c^2 = \mathrm{Var}(\tilde{y}_c)$. Thus, by assuming no correlation between county yield and a unit's idiosyncratic yield, unit yield variance can also be decomposed into two uncorrelated parts: the product of the square of unit yield's sensitivity to county yield and county yield variance; and unit's idiosyncratic yield variance. If

$\theta_i = 0$, then there is no idiosyncratic risk and unit yield is completely determined by county yield in all moments.

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

$$(2) \qquad R_i^2 = \frac{\mathrm{Var}[\beta_i(\tilde{y}_c - \mu_c)]}{\mathrm{Var}(\tilde{y}_i - \mu_i)} = \frac{\beta_i^2 \mathrm{Var}(\tilde{y}_c)}{\mathrm{Var}(\tilde{y}_i)} = \frac{\beta_i^2 \sigma_c^2}{\beta_i^2 \sigma_c^2 + \theta_i^2} = \frac{1}{1 + \theta_i^2/\beta_i^2 \sigma_c^2}.$$

Unlike β_i, which is widely used as measure of farm-level systemic risk in crop insurance and farm-level policy studies (Miranda 1991; Mahul 1999; Coble et al. 2000; Barnett et al. 2005; Coble and Dismukes 2008; Carriquiry et al. 2008; Cooper et al. 2012), R_i^2 is bounded between 0 and 1, and provides a straightforward measure of how important systemic risk relative to idiosyncratic risk. Values $R_i^2 > 0.5$ suggest that systemic risk is the major risk source faced by the producer so that area insurance might have the potential to reduce more than half of total risks.

Equation (2) also shows that systemic risk can be decomposed into three more fundamental components: *i*) idiosyncratic yield variance, θ_i^2; *ii*) county yield variance, σ_c^2; and *iii*) the square of unit yield's sensitivity to county yield, β_i^2. For a given insurance unit, the magnitude of systemic risk is jointly determined by these three components, where Proposition 1 provides simple inferences that can be extracted from the equation.

Proposition 1. *Ceteris paribus*, systemic risk is *i*) increasing in county yield variance, σ_c^2, and also in the square of unit yield's sensitivity to county yield, β_i^2, *ii*) decreasing in a unit's idiosyncratic yield variance, θ_i^2.[3]

[3] Equation (2) also conveys that systemic risk equals 0 when the unit yield is uncorrelated with county yield or when county yield has no variability. It equals 1 whenever no idiosyncratic risk occurs. As these these extreme cases are unlikely to happen in real life, they are omitted in our discussion.

Proposition 1 provides some indicators for which counties and insurance units are likely to have high systemic risk. The most discernible clue is that, *ceteris paribus*, systemic risk will be higher in counties with larger county yield variance. Since county yield data can be easily accessed from National Agricultural Statistical Service (NASS), we can readily identify these counties. Meanwhile, although idiosyncratic yield variance data are generally not available, we might expect that units in regions where heterogeneous approaches to production are taken are more likely to display large idiosyncratic yield variance and so low systemic risk. In particular, there is reason to believe that producers who adopt innovative production practices will have a lower systemic risk so that subsidized area-based insurance will discourage technology adoption.

2.2. Systemic Risk and County Growing Conditions

A strand of literature has found that yield correlations are higher in extreme weather years (Okhrin et al. 2013; Goodwin and Hungerford 2014; Tack and Holt 2016; Du et al. 2017). Du et al (2017) has also found that yield on better-quality land is more resilient to bad weather. In this subsection we check whether systemic risk also vary with county growing conditions.

Rather than directly model systemic risk as a function of county growing condition variables, labelled as Z_c, we allow each of the three components of systemic risk to be a function of Z_c. The aggregate effect of growing conditions on systemic risk can be obtained as follows. Letting $\tau_i^2 = \theta_i^2/\beta_i^2\sigma_c^2$, 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 yields

(3)
$$\begin{aligned}\ln\left(\frac{R_i^2}{1-R_i^2}\right) = \ln\left(\frac{1}{\tau_i^2}\right) = -\ln(\tau_i^2) = -\ln\left(\frac{\theta_i^2(Z_c)}{\beta_i^2(Z_c)\sigma_c^2(Z_c)}\right)\\ = 2\ln[\sigma_c(Z_c)] + 2\ln[\beta_i(Z_c)] - 2\ln[\theta_i(Z_c)],\end{aligned}$$

which shows that the logistic transformation of R_i^2 can be taken to be the sum of three functions of Z_c. Since the logistic transformation is monotonic, then the effect of Z_c on R_i^2 shares its sign with the effect of Z_c on $\ln[R_i^2/(1-R_i^2)]$.

Equation (3) also shows that effects of county growing conditions on systemic risk pass through σ_c, β_i, and θ_i. Thus, for a given county growing condition variable, if it has the same effect on σ_c and β_i, then these two effects are concordant with each other. However, were a growing condition to have the same effects on θ_i and either one or other of σ_c and β_i then these two effects would offset.

2.3. A Brief Introduction of Individual Insurance Contract and Area Insurance Contract

Before introducing our definition of TRASR, we first outline how individual-based insurance and area-based insurance work. For simplification, we only study TRASR for yield-based insurance programs. Multiple Peril Crop Insurance (MPCI) is picked to represent the individual yield based insurance while Area Yield Protection (AYP) is picked to represent the area yield based insurance because they are the major individual yield insurance program and area yield insurance programs currently implemented in the United States. A similar analysis can be conducted for revenue-based insurance programs, but empirics would be more demanding because a stochastic price variable correlated with both farm and area average yields would need to be accounted for. All premium, subsidy and indemnity variables are measured in bushels per acre (bu./ac).

Producer i in county c with random yield $\tilde{y}_i$ who chooses to purchase MPCI will receive indemnity payments in the form

$$\tilde{n}_i = \max(\phi_i \bar{y}_i - \tilde{y}_i, 0), \tag{4}$$

where $\tilde{n}_i$ is the realized MPCI indemnity payments, $\bar{y}_i$ is an insurer's reference 'expected' unit yield, or the guaranteed unit yield, as established by the RMA, and ϕ_i is the MPCI coverage level chosen by the producer with $\phi_i \in \{0.5, \dots, 0.85\}$ where evaluations are in 5% increments. Thus, MPCI would pay producers indemnities when realized individual yield is lower than the policy protection amount, $\phi_i \bar{y}_i$. The lower the realized individual yield, the higher the indemnity amount.

Similarly, AYP pays indemnities when county average yield is lower than policy protected county yield level, but its indemnity function takes a more complicated form,

$$\tilde{n}_c = \rho \max\left[\min\left(\frac{\bar{y}_c \phi_c - \tilde{y}_c}{\phi_c - l}, \bar{y}_c\right), 0\right], \tag{5}$$

where $\tilde{n}_c$ is the realized AYP indemnity payment, $\bar{y}_c$ is the guaranteed county average yield, $\tilde{y}_c$ is the realized county average yield, ϕ_c is the AYP coverage level with $\phi_c \in \{0.7, \dots, 0.9\}$ where again evaluations are in 5% increments. A protection factor, ρ, with $\rho \in [0.8, 1.2]$, is introduced to allow the producer to adjust the amount of AYP indemnities. This is important because county yield losses do not perfectly match individual yield losses, see β_i in equation (1), and the protection factor allows a grower's choice of AYP indemnities to better match expected individual losses. A loss limit factor, l, with fixed value 0.18, is also introduced to scale up AYP payments where the $\min(\cdot,\cdot)$ function ensures that no additional indemnity is paid whenever the realized county yield is below 18% of the expected county yield. Overall, empirical data show that AYP generally pays more indemnities than MPCI under the same guaranteed yields, coverage choices, and realized yields. This is so because the indemnity amount is scaled up by $\rho/(\phi_c - l)$, but it also charges higher initial premiums as a result (Sherrick and Schnitkey 2016).

2.4. Calibrating Threshold Relative Area Subsidy Rate

In this subsection we illustrate the importance of premium subsidy in inducing producers to choose AYP over MPCI. We will also derive our model for TRASR.

Since MPCI provides better risk protection than AYP, under the actuarially fair premium assumption, risk averse growers will always choose MPCI over AYP whenever no premium subsidy is provided. To see this, let π_i and π_c denote, respectively, MPCI and AYP premium rates, while letting s_i and s_c denote respective MPCI and AYP subsidy rates. Under the actuarial fairness assumption, i.e., $\pi_i = \mathrm{E}(\tilde{n}_i)$ and $\pi_c = \mathrm{E}(\tilde{n}_c)$, the expected net returns from purchasing MPCI and AYP are

(6) $$\mathrm{E}(\tilde{y}_i^{net}) = \mathrm{E}[\tilde{y}_i + \tilde{n}_i - (1 - s_i)\pi_i] = \mathrm{E}(\tilde{y}_i) + s_i\mathrm{E}(\tilde{n}_i),$$

and

(7) $$\mathrm{E}(\tilde{y}_i^{net}) = \mathrm{E}[\tilde{y}_i + \tilde{n}_c - (1 - s_c)\pi_c] = \mathrm{E}(\tilde{y}_i) + s_c\mathrm{E}(\tilde{n}_c),$$

respectively. Thus, when no premium subsidy is offered, i.e., when $s_i = s_c = 0$, then the expected net returns from purchasing MPCI and AYP are the same and equal $\mathrm{E}(\tilde{y}_i)$. Risk averse growers then will never choose AYP over MPCI as MPCI provides better risk protection.

Only when premium subsidies are introduced and the condition $s_c\mathrm{E}(\tilde{n}_c) > s_i\mathrm{E}(\tilde{n}_i)$ is satisfied, might there be a positive probability that risk averse growers will choose AYP over MPCI. The relative subsidy rate

(8) $$\hat{s} = s_c/s_i = \mathrm{E}(\tilde{n}_i)/\mathrm{E}(\tilde{n}_c)$$

provides the lower bound of the relative subsidy rate that is required to make a risk-neutral producer indifferent between choosing AYP and MPCI. We call $\hat{s}$ the threshold relative area subsidy rate (TRASR), below which even risk-neutral growers will always choose MPCI over AYP. A risk-neutral producer with TRASR exceeding 1 would require the AYP subsidy rate to surpass the MPCI subsidy rate in order to be indifferent between these two insurance contracts.

By substituting in MPCI and AYP indemnity functions, we can develop an explicit form for TRASR. First note that by assuming the yield expectations established by RMA perfectly matches the actual unit yield expectation and county yield expectation, i.e., $\bar{y}_i = \mu_i$ and $\bar{y}_c = \mu_c$, equations (1) and (4) then jointly imply $\tilde{n}_i = \max[\beta_i(\mu_c - \tilde{y}_c) - \mu_i(1 - \phi_i) - \theta_i\varepsilon_i, 0]$, which presents MPCI indemnities as a function of county yield and idiosyncratic yield. Now the MPCI indemnity function can be rewritten as

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

where we use $\mathrm{M}_i(\varepsilon_i) = \mu_c - [\mu_i(1-\phi_i) + \theta_i\varepsilon_i]/\beta_i$ to denote the trigger value of $\tilde{y}_c$ below which MPCI pays strictly positive indemnities. Here, β_i can be roughly treated as the scaling factor for the MPCI indemnity function as it scales the difference between county yield expectation and the realized county yield.

Similarly, AYP indemnity function can be rewritten as

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

where we use $\mathrm{M}_c = \mu_c\phi_c$ to denote the trigger value for $\tilde{y}_c$ below which AYP pay strictly positive indemnities and we use $\mathrm{M}_c^l = \mu_c l$ to denote the lower bound on $\tilde{y}_c$ below which AYP pays its maximum indemnity level. We further use $\alpha_c = \rho/(\phi_c - l)$ to denote the inverse of the slope of the second segment of AYP indemnity function, which can also be treated as the scaling factor for AYP indemnity payments. Since $0.8 \leq \rho \leq 1.2$, $0.7 \leq \phi_c \leq 0.9$ and $l = 0.18$ it follows that $\alpha_c \geq 0.8/0.72$ and so $\alpha_c > 1$ always hold.

With some simple transformations, the expected indemnities for MPCI and AYP can be expressed as

(11) $$\mathrm{E}(\tilde{n}_i) = \beta_i \iint_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i),$$

and

(12) $$\mathrm{E}(\tilde{n}_c) = \alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c,$$

respectively, where $F(\cdot)$ is the cumulative density function of $\tilde{y}_c$ and $G(\cdot)$ is the cumulative density function of ε_i. Thus, TRASR can be rewritten as

$$(13) \qquad \hat{s} = \frac{\mathrm{E}(\tilde{n}_i)}{\mathrm{E}(\tilde{n}_c)} = \frac{\beta_i \int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)}{\alpha_c \int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c},$$

which presents TRASR as a function of two random variables, $\tilde{y}_c$ and ε_i, and a set of parameters. We will study this relationship by both formal analysis and numerical simulation in the sections that follow. In particular, we seek to understand how TRASR is affected by systemic risk. A natural log version of equation (13) is adopted to simplify the analysis, i.e.,

$$(14) \qquad \begin{aligned} \ln(\hat{s}) = \ln(\beta_i) + \ln\left[\int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)\right] - \ln(\alpha_c) \\ - \ln\left[\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c\right]. \end{aligned}$$

Thus, TRASR can be roughly decomposed into four parts: MPCI's indemnity scaling factor, unscaled MPCI's expected indemnity payment, AYP's indemnity scaling factor, and unscaled AYP's expected payments.

2.5. TRASR and Systemic Risk

A comparison of (14) and (3) is in order because terms can be matched in an informative way. Although TRASR is not a direct function of systemic risk, it is a function of unit yield's sensitivity to county yield, β_i, unit idiosyncratic yield standard deviation, θ_i, and metrics related to county yield variance.[4] Differentiating $\ln(\hat{s})$ with respect to θ_i yields

$$(15) \qquad \frac{d\ln(\hat{s})}{d\theta_i} = -\frac{\int F\big(\mathrm{M}_i(\varepsilon_i)\big)\varepsilon_i dG(\varepsilon_i)}{\int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)} > 0.$$

In order to demonstrate the inequality note that $F(\mathrm{M}_i(\varepsilon_i))$ is a decreasing function of ε_i so that $F(\mathrm{M}_i)\varepsilon_i$ puts larger weights on negative values of ε_i and smaller weights on positive

[4] Both $\int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)$ and $\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c$ can be seen as integrations of cumulative distribution functions and so can be seen as a generalized measure of a variable's variance where an increase represents a mean-preserving spread. In that light, the first expression can be taken to represent the variability of unit yield and the second can be taken to represent the variability of county yield.

values of ε_i. By $E(\varepsilon_i) = 0$ and the covariance inequality we then have

$\int F\big(\mathrm{M}_i(\varepsilon_i)\big)\varepsilon_i dG(\varepsilon_i) < 0$ and $d\ln(\hat{s})/d\theta_i > 0$.

Inequality (15) asserts that TRASR is increasing in a unit's idiosyncratic yield standard deviation. This is because in deriving inequality (15) we have held $G(\varepsilon_i)$ fixed. Thus, by equation (1), *ceteris paribus*, increase in θ_i only increases the absolute value of $\tilde{y}_i - \mu_i$, or the absolute values of losses (when $\tilde{y}_i - \mu_i < 0$) and gains (when $\tilde{y}_i - \mu_i > 0$) incurred by the producer, but has no effect on the probability of incurring losses and gains. Since larger losses lead to more indemnity payments while larger gains have no effect on indemnity payments, then the expected MPCI indemnity payment will increase. Moreover, since a change in a unit's idiosyncratic yield standard deviation has no effect on the expected AYP indemnity payment, TRASR will then increase.

Inequality (15) also reveals the fact that area-based insurance might be costly to implement in areas where unit yield does not match well with county yield. Producers in these areas are still exposing to risks under the area-based insurance program and thus would require a high subsidy rate to compensate the risk exposure.

Similarly, differentiating $\ln(\hat{s})$ with respect to β_i yields

$$(16)\qquad \frac{d\ln(\hat{s})}{d\beta_i} = \frac{1}{\beta_i} + \frac{\mu_i(1-\phi_i)\int F\big(\mathrm{M}_i(\varepsilon_i)\big)dG(\varepsilon_i)}{\beta_i^2 \int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)} + \frac{\theta_i \int F\big(\mathrm{M}_i(\varepsilon_i)\big)\varepsilon_i dG(\varepsilon_i)}{\beta_i^2 \int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)}.$$

The sign of equation (16) is undetermined because the first two terms on the right-hand side are positive but the third is negative as $\int F\big(\mathrm{M}_i(\varepsilon_i)\big)\varepsilon_i dG(\varepsilon_i) < 0$. An increase in β_i increases MPCI's indemnity scaling factor but has ambiguous effect on unscaled MPCI indemnity payment. By equation (9), on the one hand, changes in β_i might change the indemnity amount received by the producer where the sign of this change depends on the sign of $\mu_c - \tilde{y}_c$. On the other hand, changes in β_i might also change the trigger value, $\mathrm{M}_i(\varepsilon_i)$, thus changing the probability of receiving indemnities. Overall, the sign of equation (16) depends on both the distribution of county yield and the distribution of a unit's idiosyncratic yield. But when $\theta_i = 0$, then $d\ln(\hat{s})/d\beta_i > 0$, i.e., when there is no idiosyncratic risk, then TRASR is increasing in unit yield sensitivity to county yield.

Equation (14) conveys limited information on how TRASR depends on county yield standard deviation as σ_c does not enter the $\hat{s}$ function, directly at least. This is because in presenting the function of $\hat{s}$, we have fixed the county yield distribution, $F(\tilde{y}_c)$, where the distribution captures the variable's moments. Thus, it is difficult to study the relationship between TRASR and county yield standard deviation in this framework. But a simple example can help illustrate how TRASR changes with county yield standard deviation that are caused by changes in the scale of losses but not by changes in the probability of incurring losses or gains.

Ceteris paribus, by equation (9) and (10), one-unit reduction in $\tilde{y}_c$ will increase MPCI indemnity payments by β_i and increase AYP indemnity payments by α_c. As loss probabilities do not change, then when $\alpha_c > \beta_i$, $\mathrm{E}(\tilde{n}_c)$ will increase by more than $\mathrm{E}(\tilde{n}_i)$ and TRASR will decrease. Thus, when holding loss probabilities fixed, the effect of changes in county yield standard deviation on TRASR depends on the relative size of AYP's indemnity scaling factor and MPCI's indemnity scaling factor. Given that previous studies estimated a clustering of β_i around 1 and AYP participants generally pick $\rho = 1.2$, for most producers $\alpha_c > \beta_i$ is more reasonable than $\alpha_c \leq \beta_i$. Thus, for most producers, TRASR is likely to decrease with an increase in county yield variance when holding county yield loss probability fixed.

Proposition 2. *Ceteris paribus*, TRASR is increasing in unit's idiosyncratic yield standard deviation. While the effect of a unit yield's sensitivity to TRASR depends on both the distribution of county yield and the distribution of a unit's idiosyncratic yield, when no idiosyncratic risk occurs then TRASR is strictly increasing in unit yield' sensitivity to county yield. The effect of county yield standard deviation on TRASR depends on the relative size of AYP's indemnity scaling factor and MPCI's indemnity scaling factor. When AYP's indemnity scaling factor is larger than MPCI's indemnity scaling factor then TRSAR is decreasing in county yield standard deviation.

Proposition 2 establishes that the relationship between TRASR and systemic risk depends on a large set of parameters, including the distributions of county yield and unit's idiosyncratic yield. But it asserts that TRASR is higher for producers with a large idiosyncratic risk and for producers in counties with lower county yield variance. Since systemic risk is decreasing in a unit's idiosyncratic yield variance and increasing in county yield variance, there might exist a negative correlation between TRASR and systemic risk. This relationship has an important policy implication. That is, AYP has the potential to simultaneously provide good risk protection and save subsidy costs in counties where systemic risk is high. By calibrating with appropriate data, we can easily target these counties.

2.6. TRASR and Yield Expectations

We now turn to analyze how yield expectations affect TRASR. Differentiating $\hat{s}$ with respect to μ_i provides

$$\frac{d\ln(\hat{s})}{d\mu_i} = -\frac{(1-\phi_i)\int F\big(\mathrm{M}_i(\varepsilon_i)\big)dG(\varepsilon_i)}{\beta_i \int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)} < 0, \tag{17}$$

i.e., TRASR is decreasing in individual yield expectation. Please note the revealing presence of $(1-\phi_i)$ in the expression. Relationship (17) may appear at first to be counterintuitive because for a given distribution of unit yield, an increase in mean unit yield should increase both the probability of receiving MPCI indemnities and the amount of MPCI indemnities received, while have no effect on AYP indemnities. Thus, expected MPCI indemnity payments should increase and TRASR should increase as well, not decrease.

The conflict arises because expected yield in our model is not exogenously given but is determined by yield distributions. To see this, by equation (1), when holding county yield, county expected yield, and a unit's idiosyncratic yield constant, an one unit increase in unit expected yield will also increase realized unit yield by one unit. Now, the MPCI indemnities will decrease by $1 - \phi_i$ (or 0), and the probability of receiving MPCI indemnities will also decrease because the unit yield distribution has shifted to the right by one unit while the indemnity trigger value only shifted to the right by $\phi_i < 1$ units. Thus, an increase in the unit yield expectation will decrease the expected MPCI indemnity payment while having no effect on AYP indemnity payments. As a result, TRASR will decrease with an increase in unit expected yield.

The above analysis also reveals a bias induced by proportional coverage in the current crop insurance system. Since crop insurance pays indemnities based on shortfalls from the guaranteed yield which is proportional to expected yield, producers with higher expected yield might need to experience a large yield loss, which generally has a smaller probability of occurrence on better land, in order to receive indemnities, especially when they choose a low coverage level. Thus, productive producers who have high yield expectations would prefer to either not insure or insure at high coverage levels, i.e., to plunge (Tobin 1958). This analysis might help explain the observation that producers in Central Corn Belt where yields are high generally choose high coverage levels. Insurance incentives would be very different where absolute yield shortfalls insured.

Remark. Plunging behavior is more likely for choices on better quality land than on worse quality land.

Similarly, taking the first derivative of $\ln(\hat{s})$ with respect to μ_c, we get

(18) $$\frac{d\ln(\hat{s})}{d\mu_c} = \frac{\int F\big(\mathrm{M}_i(\varepsilon_i)\big)dG(\varepsilon_i)}{\int\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c\, dG(\varepsilon_i)} + \frac{lF(\mathrm{M}_c^l)}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} - \frac{\phi_c F(\mathrm{M}_c)}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c}.$$

The sign of equation (18) is undetermined because the first two right-hand terms are positive while the third is negative. This is because, similar to the effect of changes in unit yield expectation, changes in county expected yield also change the distribution of county yield and thus have multiple effects on the expected AYP indemnity payments. By equation (1), holding unit yield, unit expected yield, and unit idiosyncratic yield fixed, a one unit increase in μ_c increases $\tilde{y}_c$ by one unit. By equation (10), then the maximum indemnity level $\mu_c\rho$ increases and the range, $\tilde{y}_c \leq \mu_c\rho$, within which the producer receives the maximum AYP indemnity level, is also enlarged. The increase in μ_c also widens the range $\mu_c l < \tilde{y}_c < \mu_c\phi_c$ within which AYP pays positive indemnities, but the indemnity level now decreases by $\alpha_c(1-\phi_c)$. Thus, an increase in county expected yield expectation has multiple effects on expected AYP indemnity payments that might offset each other in signs and its effect on TRASR is undetermined.

Proposition 3. *Ceteris paribus*, TRASR is decreasing in individual unit expected yield because an increase in individual unit expected yield decreases both the probability of receiving MPCI indemnity payments and the amount of MPCI indemnities received given that a payment is made. The effect of an increase in county yield expectation on TRASR is ambiguous.

Proposition 3 asserts that TRASR is lower for producers with higher expected yield, i.e., for producers with higher historical yields as yield expectation is largely determined by yield history. Thus, productive producers may choose AYP over MPCI at a reasonable subsidy rate.

2.7. TRASR and Coverage Levels and Protection Factor

Under the actuarially fair premium assumption, it is obvious that an increase in MPCI coverage level will increase TRASR because it increases expected MPCI indemnity payment and has no effect on expected AYP indemnity payments. But the relationship between AYP coverage level and TRASR is not so clear.

Taking the first-order derivative of $\ln(\hat{s})$ with respect to ϕ_c, we get

$$\frac{d\ln(\hat{s})}{d\phi_c} = -\frac{F(\mathrm{M}_c)\mu_c}{\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c} + \frac{\rho}{\alpha_c(\phi_c - l)^2}. \tag{19}$$

The first right-hand term in equation (19) is negative while the second is positive. Thus, the sign of equation (19) is undetermined. This is because, by equation (14), an increase in ϕ_c has a positive effect on the AYP's indemnity scaling factor related part but has a negative effect on the unscaled AYP indemnity payments part. The overall effect is then determined by the relative size of these two effects.

Finally, an increase in protection factor will decrease TRASR because it increases AYP indemnity scaling factor and then increases expected AYP indemnity payments.

Proposition 4. Ceteris Paribus, TRASR is increasing in MPCI coverage level and is decreasing in AYP protection factor. The relationship between TRASR and AYP coverage level is unclear.

Proposition 4 asserts that TRASR is higher for producers who are willing to choose higher MPCI coverage levels because now MPCI provides more indemnity payments. Given the fact that average coverage levels have increased over the past decade (Schnitkey and Sherrick 2014), AYP might need to set a high subsidy rate in order to attract producers to choose AYP over MPCI. The negative relationship between TRASR and the AYP protection factor suggests that allowing producers to choose a protection factor larger than the current maximum level might induce more producers to choose AYP over MPCI. However, as Miranda (1991) has discussed, setting the protection level too high would be politically infeasible and would raise the level and variability of total indemnity outlays.

CHAPTER 3. DATA AND VARIABLES

Unit corn yield data have been obtained from the 2008 unit-level RMA records. These data are four-to-ten-year yield historical yield data used to establish Actual Production History (APH) that is used to set up unit-level yield expectations. The historical yields are continuous unless the crop being insured is not planted in a certain year (Edwards 2011). If at least four successive yield records are not available, 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 period 1998 through 2007 because *i*) including units with transition yield records will introduce artificial correlation between county yield and unit yield and thus bias our estimate of systemic risk, *ii*) including years before 1998 will result in a year sample that is only comprised of units that did not plant in some years between 1998 and 2007 and will also result in few unit observations in some years that might not well capture the temporal systemic risk in some counties.[5] We also drop units that adopt the irrigation practice and units in counties where the irrigation rate[6] exceeds 20%[7] because systematic difference might exist between irrigated land and non-irrigated land, and also between counties where irrigation rate is high and where irrigation rate is low. Table 11 in Appendix B summarizes observation losses after each data screening step and after merging with county yield data, weather data, and land quality data.

County average corn yield data are from National Agricultural Statistical Service (NASS). Only counties in twelve traditional major corn production states in the Midwest and Great

---

5 Only keep yield records in the most recent ten previous years are also the common practice in studies adopting RMA unit-level data, see Deng et al. (2007) and Claassen and Just (2011).

6 Using RMA unit-level data, county irrigation rate is defined as the ratio of the number of units that adopt irrigation practice over the total number of insurance units in that county.

7 Dropping counties whose irrigation rate exceeds 30% yields similar results.

Plains are kept in this study: Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin. These states accounted for more than 88% of total corn production in the United States in 2018.[8] For the purpose of getting enough observations to estimate county yield trend, we drop counties without 50 successive yearly yield records over 1958-2007. To make our study county representative, we further drop counties with less than 30 insurance units.

We construct two sets of weather variables to reflect county-level weather conditions. The first set contains growing degree days (GDD) and stressful degree days (SDD), which are widely used in the literature to measure heat conditions (Schlenker et al. 2006; Schlenker and Roberts 2009; Du et al. 2015). Heat data are from National Oceanic and Atmospheric Administration (NOAA) and are recorded at station level. 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 in day d in 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. For each county, GDD, labeled as G_c, is defined as the ten-year average of total beneficial degrees in the range [10 ℃, 30 ℃] over the growing season (Neild and Newman 1990) while SDD, labeled as S_c, is defined as the ten-year average of the sum of excess degrees that are greater than 32.22 ℃ over the growing season (Schlenker and Roberts, 2009). Specifically,

$$G_c = \frac{1}{10}\sum_{t\in Y_t}\sum_{d\in M_t}\left[0.5\left(\min\left(\max\left(T_{c,d,t}^{Max}, T^l\right), T^h\right) + \min\left(\left(\max\left(T_{c,d,t}^{Min}, T^l\right), T^h\right)\right) - T^l\right)\right], \tag{20}$$

[8] Production data are from Crop Production 2018 Summary, USDA, 2019. https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropan19.pdf

and

$$S_c = \frac{1}{10}\sum_{t\in Y_t}\sum_{d\in M_t}\left[0.5\left(\max(T^{Max}_{c,d,t}, T^k) + \max\left(T^{Min}_{c,d,t}, T^k\right)\right) - T^k\right], \tag{21}$$

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 ($M_t = \{May, June, July, August\}$) in year t, and Y_t denotes the set of our sample weather data years ($Y_t = \{1998, 1979, \ldots, 2007\}$).

The second weather variable set measures the relative moisture in a county. We use the Palmer's Z (PZ) index from NOAA to measure drought and excess moisture (Xu et al. 2013). PZ index measures the departure of monthly weather from the average moisture condition in a climate division level. PZ index within the range [-2, 2.5] is viewed as normal while below indicates severe drought and above 2.5 indicates severe wetness (Karl 1986; NOAA 2014). Since different parts of a county might be covered by different climate divisions, to transfer the climate division data into county level, we first calculate the ratio of county acres covered by each intersect climate division to get a weight metrics. We then time the monthly climate division PZ index with the associated weight and sum the products across all intersect climate divisions to get the monthly county-level PZ index, $P_{c,m,t}$. We define a drought variable,

$$D_c = \sum_{t\in Y_t}\sum_{m\in M_t} I_t\left(P_{c,m,t} < -2\right), \tag{22}$$

and a wetness variable,

$$W_c = \sum_{t\in Y_t}\sum_{m\in M_t} I_t\left(P_{c,m,t} > 2.5\right), \tag{23}$$

where m denotes month and $I_t = 1$ whenever the condition in the parentheses is satisfied and $I_t = 0$ otherwise. Thus, D_c measures the frequency of severe short-term drought occurred in county c over 1998-2007, and W_c measures the frequency of severe short-term wetness occurred in county c in the same time period.

Land quality data are from National Resource Conservation Service (NRCS). County land quality, labeled as L_c, is defined as the fraction 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 while classes higher than II at least have plural severe limitations for cultivation.[9] Since there is little variation in land capability across years, we use the 2010 land capability as our measure of land quality for all sample years.

Table 1 presents the descriptive statistics for our variables. The mean of yearly unit yield is 149 bushels per acre and the standard deviation is 39.4 bushels per acre. Mean yearly county average yield is lower than mean yearly unit yield, suggesting that our unit sample might have over-weighted the number of high productivity farms. County yield also has a smaller standard deviation, as expected. The mean of GDD and SDD are 1,290 and 14.8, respectively, suggesting that on average, the accumulation of beneficial degrees is plentiful for corn growth and the appearance of excessive heat is rare. The mean of monthly severe drought occurrence is 4.72 and the mean of monthly severe wetness occurrence is 5.53. Thus, severe drought and severe wetness do not occur often in our sample counties over the sample period. The large mean for severe wetness also suggests that there was an oversupply of moisture through 1998 to 2007. Mean fraction of county land in LCC I or LCC II is 47.3%, reflecting the fact that a large fraction of our sample counties' land is favorable for cultivation.

In addition to the descriptive statistics, in Figure 2 we also map the geographic distribution of weather variables and the land quality variable. Panels A and B show that GDD and SDD generally increase as one moves south. This trend is consistent with the fact that temperature increases as one moves south. SDD also increases as one moves west. Panel C shows that

[9] The classification scheme follows USDA's description at https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/?cid=nrcs143_014040

counties in Western and Eastern Corn Belt generally experienced more severe drought than other counties while panel D shows that counties in the north part of Corn Belt generally experienced more severe wetness than counties in the south part of Corn Belt. Counties with a large fraction of good land, as shown by Panel E, are mainly located in the Northwestern and Southeastern Corn Belt states, especially in Iowa, Illinois, Indiana, and Western Ohio.

CHAPTER 4. EMPIRICAL RESULTS

4.1. Measuring Systemic Risk

Following equation (1), we establish the following empirical model to estimate systemic risk,

(24) $$\tilde{y}_{i,t} - \mu_{i,t} = \alpha_i + \beta_i(\tilde{y}_{c,t} - \mu_{c,t}) + \theta_i \varepsilon_{i,t},$$

where a constant term α_i is introduced to ensure that the error term has zero expectation. Unit-level systemic risk is estimated by the R-squared of OLS estimation, i.e., $R_i^2 = \beta_c^2 \mathrm{Var}(\tilde{y}_{c,t} - \mu_{c,t})/\mathrm{Var}(\tilde{y}_{i,t} - \mu_{i,t})$. However, since each insurance unit has only ten-year yield observations, OLS estimation with such a short time period would be highly imprecise. We instead estimate the county-level systemic risk by pooling together all units within a given county. The empirical model is now given by

(25) $$\tilde{y}_{i,t} - \mu_{i,t} = \alpha_c + \beta_c(\tilde{y}_{c,t} - \mu_{c,t}) + \theta_i \varepsilon_{i,t},$$

where α_c and β_c denote the county-level constant term and county-level unit yield's sensitivity to county yield, respectively. County-level systemic risk, labelled as R_c^2, is then measured by the R-squared of the OLS estimation for equation (25), i.e. $R_c^2 = \beta_c^2 \mathrm{Var}(\tilde{y}_{c,t} - \mu_{c,t})/\mathrm{Var}(\tilde{y}_{i,t} - \mu_{i,t})$.

To estimate equation (25), however, we need to first estimate 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 some time trend that establishes yield expectation. Following Deng et al. (2007), we employ the following log-linear model to estimate county yield trend for each county,

(26) $$\log(\tilde{y}_{c,t}) = \lambda_{c,0} + \lambda_{c,1}(2008 - t) + v_{c,t},$$

where $\lambda_{c,1}$ captures the inverse trend in percent yield change starting from 2008, or the difference in log yield expectation between year *t* and year 2008. NASS county yield data from 1958 to 2007 are used to estimate equation (26) and county yield expectation for county *c* in year *t* is then calculated as $\mu_{c,t} = \hat{y}_{c,t} = \exp[\hat{\lambda}_{c,0} + \hat{\lambda}_{c,1}(2008 - t)]$.

For each insured unit, RMA sets up its yearly reference yield expectation by use of the unit's APH yields, which equals the average of the last ten years' actual yields. However, in our data, only 2008 APH yields are available. We proceed by assuming that unit yield expectation shares the same time trend with county yield expectation. Thus, the APHs for 1998 to 2007 are then calculated as

$$APH_{i,t} = APH_{i,2008} * \exp\left[\hat{\lambda}_{c,1}(2008 - t)\right], \quad (27)$$

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

A concern with using APH yield as unit yield expectation is that the APH yield lags the true expected yield due to improved crop genetics and cultural practices (Edwards, 2012). To correct for this downward bias, RMA introduced the trend-adjusted APH yield in 2012. Basically, a trend adjustment factor is estimated for each crop and each county, which is equal to the estimated annual increase in NASS county yield. The trend adjusted factor is then used to scale up past actual yield records. Following this practice, we also calculate the trend-adjusted APH. Specially, the yearly county specific trend-adjustment factor, $TA_{c,t}$, is estimated as

$$TA_{c,t} = (\hat{y}_{c,t} - \tilde{y}_{c,1958})/(t - 1958), \quad (28)$$

where $\hat{y}_{c,t}$ is the yield prediction for county c in year t estimated by equation (26), 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' that is ahead of year t then equals $(t - t') \times TA_{c,t}$. To illustrate, suppose we want to adjust the 2007 APH for unit i in county c and the trend adjustment factor for county c in year 2007 is $TA_{c,2007} = 2$. Then the 2006, 2005, …, and 1997 actual yield of unit i will be adjusted up by 2, 4, …, and 20 bushels per acre, respectively. The 2007 APH would be adjusted up by $(2 + 4 + \cdots + 20)/10 = 11$ bushels per acre. The formula of APH with trend-adjusted can be summarized as

$$TA\_APH_{i,t} = APH_{i,t} + \frac{1}{10}\sum_{j=1}^{10} j \times TA_{c,t}, \tag{29}$$

and the trend-adjusted APH is then used as the expected unit yield in eqn. (4) before, i.e., $\mu_{i,t} = TA\_APH_{i,t}$.

We estimated model (25) and decomposed the resulting systemic risk measure according to (3). Table 2 presents the descriptive statistics for county systemic risk and its three components. Mean county systemic risk is 0.37, suggesting that in general, systemic risk explains slightly more than one-third of total unit-level yield variability. But the magnitude of systemic risk varies considerably across counties as it ranges from 0.02 to 0.78. The mean of county-level unit yield's sensitivity to county yield is 1.05 and it ranges from 0.19 to 2.05. This fact violates Miranda's (1991) assertion that acre-weighted average β_i within a given county should equal to 1. At least two reasons may contribute to this violation. First, our county yield data are from NASS whereas unit yield data are from RMA. NASS county yield generally does not equal the mean of acre-weighted RMA unit yield (Zulauf et al. 2017), perhaps because of differences in survey methodologies used or because not all land is insured. Second, we have dropped many units in our data screening process thus losing the connection between county yield and unit yield. The means of unit idiosyncratic yield standard deviation and county yield variance standard deviation are 23.4 and 19.4, respectively, suggesting that while idiosyncratic risk and systemic risk both contribute a significant amount of variability to unit yield, idiosyncratic risk contributions are the more important.

Figure 3 maps the geographic distributions of county systemic risk and its three components. Panel A shows that the systemic risk is generally high at the Corn Belt's southern and western fringes. It also shows that systemic risk is low in Indiana but is high in some counties in Southern Minnesota. This distribution generally lines up well with the distribution of county yield standard deviations (Panel B), suggesting that county yield variance may be the most important component determining systemic risk. This finding has important implications for policy implementation because county yield variance data are easily accessible. Program designers can easily select out counties where AYP has the potential to effectively reduce yield risk by finding out counties with large yield variance.

Panel C of Figure 3 shows that counties at the Corn Belt periphery have relatively large idiosyncratic yield standard deviations, especially counties in Southern Wisconsin and the Eastern Dakotas. On the contrary, counties in Iowa and Illinois generally have low unit's idiosyncratic yield standard deviation.[10] Panel D presents evidence that Corn Belt fringe counties have relatively large unit yield's sensitivity to county yield, but the pattern is not as evident because some counties in the central part also have large unit yield's sensitivity to county yield while counties in the southern fringe have low unit yield's sensitivity to county yield. The Moran's I statistic[11] of panel D is only 0.017, though statistically significant, while the Moran's I statistics for panels A through C generally exceed 0.03. Thus, the distribution of unit yield's sensitivity to county yield is more likely to be spatially independent when compared with the distribution of systemic risk itself and the distributions of the other two systemic risk components.

---

[10] With high mean yields and low yield variability in the Central Corn Belt, Remark 1 suggests that farmers there contemplating either individual insurance or area insurance should either not insure or take out high coverage.

[11] The Moran's I is a frequently used correlation coefficient that measures overall spatial autocorrelation in a given dataset. Its value lies within the range [-1, 1]. A zero-value statistic indicates the related variable is perfectly randomly distributed in space. When a positive (negative) value is observed, then there is a positive (negative) spatial autocorrelation across the regions. We use the *spatgsa* command in Stata, developed by Pisati (2001), to calculate the Moran's I statistics for all investigating variables.

4.2. Systemic Risk and County Growing Conditions

To study how systemic risk is determined by county growing conditions, we now regress the three components of systemic risk on county weather and land quality variables. The following log-linear model is estimated by OLS method,

(30) $$\ln(X_c) = \delta_0 + \delta_1 G_c + \delta_2 S_c + \delta_3 D_c + \delta_4 W_c + \delta_5 L_c + \nu_c,$$

where $X_c \in \{\beta_c^2, \theta_c^2, \sigma_c^2\}$, δ_0 is the constant term and ν_c is the error term.

Regression results for equation (30) appear in Table 3. Column (1) shows that GDD has a significantly negative effect on unit yield's sensitivity to county yield while SDD has a significantly positive effect. Thus, more accumulation of excessive heat increases unit yield's sensitivity to county yield while more accumulation of beneficial heat reduces the sensitivity. This result also explains why counties in the Southern Corn Belt, where GDD and SDD are both high, are less likely to have high unit yield sensitivity to county yield than do counties in the Western Corn Belt where SDD only is high. Severe wetness also has a significant negative effect on unit yield's sensitivity to county yield.

Column (2) shows that unit's idiosyncratic yield variance is significantly decreasing in GDD and land quality. Thus, greater accumulation of beneficial degrees and better land may provide a higher benchmark yield shared by all units in the county and making idiosyncratic production practice less important in determining unit yield. The negative effect of land quality on unit's idiosyncratic yield variance also explains the low idiosyncratic yield variance in Iowa and Illinois where land quality is high.

Column (3) shows that GDD and SDD both have a significantly positive effect on county yield variance. This is consistent with the observation that counties in the southern and western parts of the Corn Belt where GDD and SDD are high generally have higher county yield variance. Severe drought also has a significantly positive effect on county yield variance while severe wetness tends to reduce county yield variance. Land quality has a marginally significantly negative effect on county yield variance.

To find the overall effect of county growing conditions on systemic risk, and consistent with (3), we first regress $\ln[R_c^2/(1-R_c^2)]$ on $2\ln(\theta_i)$, $2\ln(\beta_c)$ and $2\ln(\sigma_c)$. As shown by column (4) of Table 3, consistent with Proposition 1, systemic risk is increasing in unit yield's sensitivity to county yield and county yield variance and is decreasing in unit's idiosyncratic yield variance[12]. The overall effect of GDD on systemic risk, for example, is calculated as $(-0.068\times 1.048)+(-0.050\times -0.625)+(0.126\times 1.004)=0.086.$ Standard errors are calculated by the delta method.

Column (5) of Table 3 reports the aggregate effects of county growing conditions on systemic risk. GDD and land quality both have a significantly positive effect on systemic risk by ensuring a high benchmark yield for all units. The highly significantly positive effects of SDD and severe drought suggest that, consistent with previous findings, yields are more closely correlated in extreme weather years and systemic risk is higher in counties where extreme weather occurred more often. The negative effect of wetness constitutes a counterexample to the previous statement, which may possibly be because severe wetness brings excessive water supply for some farms but plentiful water supply in the soil for other farms.

In addition to the sign and significance of county growing conditions on systemic risk, it might also be important to check which growing condition variable plays the most important role in determining systemic risk. Following Huettner and Sunder (2012), we use the Shapley

---

[12] Equation (3) predicts that the coefficient of $2\ln(\theta_i)$, $2\ln(\beta_c)$ and $2\ln(\sigma_c)$ on $\ln[R_c^2/(1-R_c^2)]$ should be exactly -1, 1, and 1. The difference between the prediction and our estimation result comes from the fact that $\beta_c^2\mathrm{Var}(\tilde{y}_{c,t}-\mu_{c,t})/\mathrm{Var}(\tilde{y}_{i,t}-\mu_{i,t})\neq\beta_c^2\mathrm{Var}(\tilde{y}_{c,t})/\mathrm{Var}(\tilde{y}_{i,t})=\beta_c^2\sigma_c^2/\sigma_i^2$. This is because in real data $\mu_{c,t}$ and $\mu_{i,t}$ are not constant but vary by year and are correlated with $\tilde{y}_{c,t}$ and $\tilde{y}_{i,t}$. If we regress $\ln[R_c^2/(1-R_c^2)]$ on $2\ln(\theta_i)$, $2\ln(\beta_c)$ and $2\ln\big(\mathrm{Var}(\tilde{y}_{c,t}-\mu_{c,t})\big)$ we will obtain coefficients equal -1, 1, and 1.

value to measure the power of growing condition variables in explaining the explainable part 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 R2. Column (1) of Table 4 shows that, consistent with the finding that the geographic distribution of systemic risk generally lines up with the geographic distribution of county yield variance, county yield variance explains about 65% of systemic risk. Unit yield sensitivity to county yield explains about 28% while unit idiosyncratic yield variance explains the remaining 6%.

For unit yield's sensitivity to county yield, column (2) in Table 4 shows that GDD explains about 47% of the variability that can be explained by county growing conditions. SDD explains about 10% and wetness explains about 35%. Column (3) shows that land quality is the most important determinant of unit's idiosyncratic yield variance as it solely explains about 86% of explained variability unit's idiosyncratic yield variance. For county yield variance, column (4) shows that GDD explains the largest part of its growing-condition-explainable variability while SDD explains the second largest part. wetness variable also explains more than 16% of the explainable part of county yield variability. Overall, Column (5) shows that GDD is the most important growing condition variable in explaining systemic risk variability. SDD also explains a significant part of systemic risk's variability. This result is consistent with the fact that systemic risk is high in the southern and western parts of the Corn Belt, where both GDD and SDD are high. Thus, county heat conditions can be viewed as the most important factors affecting systemic risk. High heat accumulation counties are more likely to have higher systemic risk, especially for those that are less likely to experience severe wetness.

4.3. Systemic Risk Estimate and Inadequate Investigating Time Horizon

Our estimation of systemic risk is based entirely on Miranda's one factor capital market model, which implicitly assumes that the correlation between county yield and unit yield is constant within the sample period. However, empirical evidences have demonstrated that the spatial correlation of crop yields tends to be higher in extreme weather years than in a typical year (Goodwin 2001; Okhrin et al. 2013). Thus, our model might overrestimate a county's systemic risk in years that a county experiences exceptionally good weather and underestimate its systemic risk in years that the county experiences exceptionally bad weather. To see which county might suffer more from such bias, we first investigate whether unit yield's sensitivity to county yield is higher when county yield is below its expectation than when county yield is above its expectation, and then check how this pattern correlates with systemic risk.

By introducing an interaction term into equation (25), we get,

$$\tilde{y}_{i,t} - \mu_{i,t} = \alpha_c + \beta_c(\tilde{y}_{c,t} - \mu_{c,t}) + \eta_c I_{c,t}(\tilde{y}_{c,t} - \mu_{c,t}) + \theta_i \varepsilon_{i,t}, \tag{31}$$

where $I_{c,t} = 1$ whenever $\tilde{y}_{c,t} < \mu_{c,t}$ and $I_{c,t} = 0$ whenever $\tilde{y}_{c,t} \geq \mu_{c,t}$. Now, there are three mutually exclusive groups that complete the value set of $\tilde{y}_{c,t} - \mu_{c,t}$, i.e., *i*) $\tilde{y}_{c,t} = \mu_{c,t}$, *ii*) $\tilde{y}_{c,t} > \mu_{c,t}$, and *iii*) $\tilde{y}_{c,t} < \mu_{c,t}$. Parameter α_c measures the mean value of $\tilde{y}_{i,t} - \mu_{i,t}$ when $\tilde{y}_{c,t} = \mu_{c,t}$, β_c measures unit yield' sensitivity to county yield whenever $\tilde{y}_{c,t} > \mu_{c,t}$, $\beta_c + \eta_c$ measures unit yield's sensitivity to county yield whenever $\tilde{y}_{c,t} < \mu_{c,t}$, and η_c captures the difference. A positive value of η_c indicates that unit yield's sensitivity to county yield is higher when county yield is below its expectation than when it is above its expectation.

Table 5 lists the sign and significance of η_c. Among the 579 sample counties, 48.7% have a significantly positive η_c and only 24.7% have a significantly negative η_c. Moreover, about 13.6% counties have insignificantly positive η_c and about 13% have insignificantly negative η_c. Thus, for most sample counties, unit yield is more sensitive to county yield when county yield is below its expectation than when county yield is above its expectation. Since systemic risk is increasing in unit yield's sensitivity to county yield, this finding confirms the conjecture that we might underestimate a county's systemic risk if its county yields were exceptionally bad over the sample period and overestimate the systemic risk if the county yields were exceptionally good.

We then plot the geographic distributions of η_c to ascertain which area's systemic risk estimate is more likely to suffer from potential bias caused by misspecification. Panel A in Figure 4 plots the geographic distribution of η_c's value and panel B plots the geographic distribution of η_c's sign and significance. There does not exist a clear pattern in η_c's geographic distributions and the Moran's I statistics of Panel A is 0.009 and that of Panel B is -0.001, asserting that η_c can be treated as independently distributed among sample counties. The Pearson's correlation test coefficient between η_c and R_c^2 is -0.0115 and is statistically insignificant. Given the short time interval at hand, our view is that we have insufficient information available to establish whether systemic risk estimate and the bias resulting from the short time period. Thus, there is no clue about whether counties with high systemic risk or low systemic risk are more likely to suffer from the bias. To obtain more accurate estimate of typical systemic risk, a longer unit yield time series is required.

4.4. Calibrating TRASR

We now turn to calibrating TRASR and investigating the relationship between TRASR and systemic risk.

When calibrating TRASR from equation (13) we need to know the distributions of both county yield and the unit's idiosyncratic yield. Studies investigating crop yield distributions mainly adopt two distinct methodologies, parametric and nonparametric. Parametric methods often assume that crop yield follows a specific distribution, such as the normal, gamma or beta distribution (Botts and Boles 1958; Gallagher 1987; Nelson 1990; Sherrick et al. 2004; Harri et al. 2011). Nonparametric methods, on the other hand, do not assume that crop yield follows a specific distribution and thus offer flexibility in capturing local idiosyncrasies in yield distribution that may not be captured by parametric methods (Goodwin and Ker 1998). Since our study contains many counties and units while appropriate corn yield distribution specifications might differ by location, flexibility considerations lead us to employ nonparametric kernel density estimation.

Since kernel density estimation requires a stationary data series, before estimating a distribution, we need to first retrend county yield data. The retrending model follows the detrending model and retrended yield for county *c* in year *t* is given by

(32) $$\check{y}_{c,t} = \left(\frac{\hat{v}_{c,t}}{\hat{y}_{c,t}}\right) \times \hat{y}_{c,2008} + \hat{y}_{c,2008},$$

where $\hat{y}_{c,2008}$ and $\hat{y}_{c,t}$ ($t \in \{1958, 1959, \ldots, 2007\}$) are county yield predictions from equation (26), and the $\hat{v}_{c,t}$s are prediction errors.

The kernel density function for county *c* is given by

(33) $$f(x_c) = \frac{1}{\sum_{t=1950}^{2007} w_{c,t}} \sum_{t=1950}^{2007} \frac{w_{c,t}}{h_c} K\left(\frac{x_c - \check{y}_{c,t}}{h_c}\right),$$

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

There is general consensus among researchers in the area that the kernel function choice is less importance 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 (MISE), which is the most common optimality criterion used to select bandwidth. As for bandwidth choice, we follow Goodwin and Ker (1998) to adopt Silverman's modified rule-of-thumb method to select the bandwidth parameters. The bandwidth parameter's formula is given by

$$h_c = \frac{0.9 \times \min\left(\sigma_y, \frac{IQR_y}{1.349}\right)}{n^{0.2}}, \tag{34}$$

where σ_y is the standard deviation of $\check{y}_{c,t}$ and IQR_y is the interquartile range of possible realization of $\check{y}_{c,t}$.

Since we only have fifty-year observations for each county, kernel estimation of such a short time period might be imprecise. To extend the data pool, following Goodwin and Ker (1998), we use yield information from adjacent counties. To be qualified for the calibration, the central county and its adjacent counties must have a total number of 200 yearly yield records, i.e., the central county must have at least three adjacent counties that have no missing yield records over 1958-2007. This requirement leaves us a sample comprised of 208,549 units in 540 counties. Each yield record in adjacent counties is assigned with weight $1/[(2N+1)\times 50]$ while each yield record in the central county is assigned with weight $(N+1)/[(2N+1)\times 50]$, where N is the number of adjacent counties.

After getting the density estimate for each point we want to evaluate, $\hat{f}(x)$, the cumulative density at each point, $\hat{F}(x)$, is calculated by the Trapezoid rule in the form

$$\hat{F}(x) = \int_0^x \hat{f}(\tilde{y}_c)d\tilde{y}_c \approx \sum_{k=1}^{N} \frac{\hat{f}(\tilde{y}_{c,k-1}) + \hat{f}(\tilde{y}_{c,k})}{2}\Delta\tilde{y}_{c,k}, \tag{35}$$

where $\Delta\tilde{y}_{c,k} = \tilde{y}_{c,k} - \tilde{y}_{c,k-1}$ and N is the total number of density-evaluation points below x. The two integrals, $\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c)d\tilde{y}_c$ and $\int_{\mathrm{M}_c^l}^{\mathrm{M}_c} F(\tilde{y}_c)d\tilde{y}_c$, are also calculated by the Trapezoid rule in a similar 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 nonparametric kernel estimation. Rather, within each county we pool units with similar 2008 APH values together and then perform nonparametric kernel estimation on the APH group level. By doing so, we are assuming that units with similar 2008 APH values also have similar idiosyncratic yield distribution. Since APH is determined by historical yield, which is in turn determined by land quality, it is reasonable to expect that units with similar APH values should have similar idiosyncratic yield distributions.

Empirically, within each county, the 2008 APH values are grouped in tens. That is, units with the smallest to the tenth smallest 2008 APH values are assigned to group 1, units with the eleventh to twentieth smallest 2008 APH values 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 yields 2,393 unique APH groups in total. As shown by Table 6, 37 counties have only one APH group and 54 counties have only two AHP 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 the idiosyncratic yield distributions from other units. However, as Figure 5 shows, counties with fewer APH groups also have smaller APH ranges. Moreover, since counties at the southern and western fringes of the Corn Belt, where systemic risk is high and which are counties we care most about, generally have four or more APH groups, bias caused by pooling units with different idiosyncratic yield distributions into the same APH group should be minor. Among the 2,393 APH groups, only 7.5% have less than 200 unit-year observations. Thus, the nonparametric kernel density estimation for most APH groups is 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 $\int \int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c \, dG(\varepsilon_i)$, are again evaluated by the Trapezoid rule.

Table 7 presents descriptive statistics for coverage-level conditional unit-level TRASR. We set the protection factor, ρ, to equal to 1.2 as it is the protection factor level that brings the highest expected AYP indemnity payments. Since 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. Consistent with Proposition 4, the mean and median of TRASR both increase with MPCI coverage level. Although Proposition 4 makes no predictions about the effect of AYP coverage level on TRASR, results in Table 7 show that TRASR decreases with increase in ϕ_c. This finding suggests that the negative effect of ϕ_c on unscaled AYP indemnity payments dominates the positive effect of ϕ_c on the AYP scaling factor.

We focus on TRASRs where MPCI coverage level is no less than 75% because 75% is the minimum coverage level chosen by most producers in the Corn Belt (Schnitkey and Sherrick 2014). Since some units have extremely large TRASR[13], we thus focus on medians other than

[13] Extremely large TRASR values occur whenever unit's $\mathrm{M}_i(\varepsilon_i)$s are very large or M_c is in the low density area of $F(\tilde{y}_c)$, or both apply. The first case occurs when a unit's idiosyncratic yield is negative and large in magnitude because $\mathrm{M}_i(\varepsilon_i)$ is decreasing in ε_i. The second case occurs when the AYP coverage level is low. In the first case, $\int_0^{\mathrm{M}_i(\varepsilon_i)} F(\tilde{y}_c) d\tilde{y}_c$ is large and in the second case, $\int_{\mathrm{M}_c^L}^{\mathrm{M}_c} F(\tilde{y}_c) d\tilde{y}_c$ is small. Table 5 also shows that some units have a zero-value TRASR. This is because kernel density estimation will assign zero probabilities to $\mathrm{M}_i(\varepsilon_i)$s whose values that are too far away from observed county retrended yields. Thus, zero values of TRASR occur when a unit's $\mathrm{M}_i(\varepsilon_i)$ is either too large or too small.

means as medians are less affected by extreme values than means. Results in Table 5 show that the medians of TRASR are generally higher than 60% when AYP coverage level is no greater than than 80%. The medians of TRASR are even higher than 100% for groups $\phi_i = 80\%, \phi_c = 70\%$, $\phi_i = 85\%, \phi_c = 70\%$ and $\phi_i = 85\%, \phi_c = 75\%$. Thus, generally, a necessary condition to induce most producers to choose a low coverage level AYP contract over a high coverage level MPCI contract is to set the related AYP subsidy rate higher than the related MPCI subsidy rate. The medians of TRASR are generally smaller than 50% when AYP coverage level is higher than 80%.

We then check whether the current AYP subsidy rate discourages producers from choosing AYP over MPCI. Here we define a new concept, threshold area subsidy rate, or TASR, which is the product of TRASR and MPCI subsidy rate. By equation (8), TASR thus constitutes the lower bound on the AYP subsidy rate below which the expected AYP indemnities would be strictly less than the expected MPCI indemnities and risk averse growers would never choose AYP over MPCI. Since MPCI offers contracts at different unit levels and subsidy rates differ by unit levels, we choose the enterprise unit level subsidy rate to derive our TASR because enterprise unit level MPCI contract is the most similar to AYP and in recent years it has covered the largest share of insured corn acres (Coble 2017). Subsidy rate information for AYP and different unit-level MPCI contracts are listed in Table 8. Enterprise unit level subsidy rates are higher than AYP subsidy rates at all available coverage levels except for the level 85%, but the coverage-level conditional subsidy rate gap between the two insurance contracts decrease with increase in coverage level.

Panel A of Table 9 reports sample means of the coverage-level conditional percent of units whose TASR is higher than the current AYP subsidy rate. To save space, we only report results corresponding to a MPCI coverage level no less than 75%. The complete results are listed in Table 12 in Appendix B.

Consistent with results in Table 7, the percent of units whose TASR is higher than the current AYP subsidy rate is increasing in MPCI coverage level and is generally decreasing in AYP coverage level. When AYP coverage level is less than 80%, generally more than half of producers will find their TASR higher than the current AYP subsidy rate. This result suggests that the low coverage level AYP contracts which are currently available do not provide the minimum required subsidy rate to compete with the subsidized MPCI contracts as subsidy transfer instruments. Raising the related subsidy rate is a necessary condition if more low coverage level AYP demand is to be induced. When AYP coverage level is no less than 80%, generally only less than 30% of producers will find their TASR higher than the current AYP subsidy rate. Thus, for most producers, the currently available high coverage level AYP contracts have met the minimum subsidy rate requirement to compete with MPCI contracts whenever producers are risk-neutral. This finding suggests that the current AYP subsidy rate might not be a major deterrence for producers to choose high coverage level AYP contracts over MPCI contracts and raising AYP subsidy rate might not be able to induce more high coverage level AYP demand. However, since risk aversion will increase the minimum required AYP subsidy rate, the percent of risk aversion producers whose TASR is higher than the current AYP subsidy rate will surely exceed the percent listed in Table 7. Raising subsidy rate might still be a viable option to induce more demand even for high coverage level AYP contracts.

Besides comparing with the current AYP subsidy rate, we also compare TASR to the 100% level to see whether some producers will not choose AYP over MPCI even when AYP is offered for free. Panel B of Table 9 presents the results. As expected, because the MPCI subsidy rate is smaller than 100%, the percent of units with TASR below 100% is much smaller than the percent with TASR below the current AYP subsidy rate. For all coverage level combinations, at most 41% of producers will find their TASR higher than 100%, so only a small proportion of producers will find that freely offered AYP contracts are not value for money when MPCI contracts are available. Thus, raising AYP subsidy rate to a level no greater than 100% might be a feasible option to induce greater AYP demand. High coverage level AYP contracts, at 80% or higher, might benefit more from this option because less than 10% producers will find their TASR higher than 100%.

4.5. TRASR, Systemic Risk, and Area Insurance Demand

We now turn to study the correlation between TRASR and systemic risk and its implication for area insurance demand. Our conceptual framework suggests that TRASR should be negatively correlated with systemic risk. To test whether this conjecture holds in data, we then conduct a series of correlation tests between county systemic risk and the two county-level TRASR variables: county-level percent of units whose TASR is higher than the current AYP subsidy rate and county-level percent of units whose TASR is higher than 100%.

Table 10 reports the Pearson's correlation coefficients between systemic risk and the two county-level TRASR variables. Panel A shows that generally, systemic risk is significantly positively correlated with the percent of units whose TASR is higher than the current AYP subsidy rate at low AYP coverage levels and is significantly negatively correlated with it at high AYP coverage levels. Thus, for low coverage level AYP contracts, producers in high systemic risk counties are more likely to find their TASR to exceed the current AYP subsidy rate than producers in low systemic risk counties, while for high coverage level AYP contracts, producers in high systemic risk counties are less likely to find their TASR higher than the AYP subsidy rate than producers in low systemic risk counties. This finding indicates that the current AYP subsidy rate is more likely to be a major deterrence for producers in high systemic risk counties to choose low coverage level AYP contracts while it is more likely to be a major deterrence for producers in low systemic risk counties to choose high coverage level AYP contracts.

Panel B of Table 10 show that systemic risk is significantly negatively correlated with the percent of units whose TASR is higher than 100% at most coverage level combinations, suggesting that producers in high systemic risk counties are less likely to have TASR exceeding 100% than are producers in low systemic risk counties. Thus, high systemic risk counties generally have larger room to raise AYP subsidy rate to induce more AYP demand.

Since the correlation test coefficients only capture the overall relationships between systemic risk and TRASR variables among all sample counties and we care more about whether high systemic risk counties are truly associated with lower TASR, we then plot the geographic distributions for county-level TRASR variables.

Figures 6 through 8 plot the geographic distributions for county-level percent of units whose TASR exceeds the current AYP subsidy rate. These figures again show that the percent of units whose TASR is higher than the current AYP subsidy rate is increasing in MPCI coverage level and is decreasing in AYP coverage level. For all coverage level combinations, counties at the Corn Belt's northeastern fringe, where systemic risk is low, generally have the highest percent of units whose TASR surpasses the current AYP subsidy rate. Counties in Central Corn Belt generally have the lowest percent of units with TASR higher than the current AYP subsidy rate. Western and Southern Corn Belt counties, where systemic risk is high, also have relative low percent of units whose TASR is higher than the current AYP subsidy rate, especially when AYP coverage level is no less than 80%. The geographic distribution patterns remain for the percent of units whose TASR is higher than 100% and Western and Southern Corn Belt counties also have low percent of units whose TASR is higher than 100%, as plotted by figures 9 through 11. These results assert that producers in Western and Southern Corn Belt counties not only enjoy better risk protection from AYP, but also require less subsidies to possibly choose AYP over MPCI than counties in other areas. Subsidy rates also have larger room to increase to induce more AYP demand in these counties.

Although our evidence suggests that counties in the Corn Belt's southern and western fringes have the best potential to grow area insurance demand, this conjecture does not match the reality of area insurance demand. Working with data from RMA Summary of Business reports we draw the geographic distribution for share of acres insured by area insurance in Figure 12. Clearly, although the overall demand for area insurance is low except in year 2006, counties with larger share of acres insured by area insurance are mainly located in the eastern and northeastern parts of the Corn Belt, where systemic risk is low while TRASR is high. This fact suggests that area insurance demand might not be mainly driven by producers' risk attitude and the contract designs, but by some factors unobserved in our data, such as the marketing strategy of insurance companies (Skees et al. 1997).

CHAPTER 5. CONCLUSION AND DISCUSSION

The Federal Crop Insurance Program has long been afflicted by high operation costs where asymmetric information and systemic risk play major roles. Area-based insurance programs have been widely proposed as viable options to deal with these problems. However, program take-up rates remain low and knowledge of determinants of the low demand is lacking in literature.

In this paper, we investigated two factors that determine the demand of area insurance programs. First, we directly measured systemic risk in corn production across the Midwest. The magnitude of systemic risk determines the effectiveness of risk protection for area insurance programs. The greater the systemic risk, the more workable is area-based insurance (Skees et al. 1997). We find that, in general, systemic risk explains slightly more than one-third of total unit yield variability. It is highest in Southern and Western Corn Belt counties but low in the Central and Northern Corn Belt. Further investigations show that the geographic distribution of systemic risk is most likely to be driven by the geographic distribution of county yield variance where systemic risk significantly increases with county heat conditions and the appearance of severe drought.

In addition to systemic risk, we also investigate the role of premium subsidy in producer's choice between area and individual insurance contracts. By calibrating TRASR, the relative subsidy rate at which AYP expected indemnity payments equals those from MPCI indemnity payments, we find that most producers would require the AYP subsidy rate to be higher than the enterprise unit level MPCI subsidy rate to possibly choose AYP over MPCI. The percent of producers holding this requirement is increasing in MPCI coverage level and is decreasing in AYP coverage level. We also find that the current AYP subsidy rate might be the major deterrence for producers to not choose AYP at low coverage levels, but not the major deterrence for them to not choose AYP at high coverage levels. Since only a small fraction of producers would require the subsidy rate of high coverage level AYP contracts to be higher than 100% to possibly choose these AYP contracts over MPCI, raising the subsidy rate of high coverage level AYP contracts might be a feasible option to induce more demand for AYP contracts.

We also find a negative correlation between systemic risk and TRASR and unit's idiosyncratic yield variance seems to be the best predictor of TRASR. Producers in Southern and Western Corn Belt counties, where systemic risk is high, generally have low TRASR. These counties are ideal areas to implement area insurance contracts as producers there would enjoy better risk protection and require lower subsidies to compensate for their risk exposure. However, this conjecture is at variance with the fact that counties with relatively high area insurance take-up rate are mainly located in the eastern and northeastern portions of the Corn Belt. Thus, some unobserved factors, such as insurance companies' marketing strategy, might play more important role than producers' risk attitude and area insurance's contract design in determining producers' demand of area insurance.

Our study largely extends current literature studying demand of area insurance contract by providing some basic facts about systemic risk and pecuniary motivations for potential area insurance buyers. The correlation between these two factors are also explored. Our study also reveals that only AYP contracts at high coverage levels has the potential to induce producers to choose it over MPCI contracts by raising its subsidy rate to a still reasonable level. But we also find that the current AYP subsidy rate might not be the major deterrence for producers to not choose high coverage level AYP contracts.

Our study does not explicitly include the risk attitude of producers. This helps us free from choosing among various utility function forms and risk aversion parameters, but it also leads us to underestimate the lower bound of the relative subsidy rate that is required by risk averse producers to choose AYP over MPCI. Our study also does not explore the case for revenue crop insurances, which has a much larger market share than yield insurance nowadays. Exploring the effects of these aspects asks for more work, and our study provides an analysis framework for future studies to work on.

APPENDICES

APPENDIX A

Major Tables and Figures

Table 1. Descriptive statistics for yield variables and county growing condition variables

Variable	N	Mean	St. Dev	Min	Max
Unit yield (bu./ac)	2,133,320	149	39.4	0	374
County yield (bu./ac)	5,790	137	28.9	27	204
G_c	579	1,290	152	360	1,612
S_c	579	14.8	14.1	0.49	92.4
D_c	579	4.72	1.89	1	13
W_c	579	5.53	2.04	0	13
L_c	579	47.3	22.5	2.20	93.5

Notes: Mean unit yield is the average of all unit-year yields over the 213,332 units and over the period 1998-2007. Mean county yield is the average of all county-year yields over the 579 counties and over 1998-2007.

Table 2**. Descriptive statistics for county systemic risk and its components**

Variable	N	Mean	St. Dev	Min	Max
Systemic risk, R_c^2	579	0.37	0.15	0.02	0.78
County-level unit yield sensitivity, β_c	579	1.05	0.20	0.19	2.05
Unit idiosyncratic yield standard deviation, θ_i	579	23.4	4.66	9.32	43.0
County yield standard deviation, σ_c	579	19.4	5.90	7.51	43.5

Table 3**. Regression results for equation (30)**

	(1)	(2)	(3)	(4)	(5)
	$2\ln(\beta_c)$	$2\ln(\theta_i)$	$2\ln(\sigma_c)$	$2\ln(\tau_c)$	$2\ln(\tau_c)$
G_c/100	-0.068***	-0.050***	0.126***		0.086**
	(0.017)	(0.015)	(0.022)		(0.026)
S_c/10	0.040**	0.002	0.081***		0.121***
	(0.019)	(0.015)	(0.023)		(0.032)
D_c	0.008	0.006	0.044***		0.048***
	(0.010)	(0.008)	(0.012)		(0.016)
W_c	-0.032***	-0.008	-0.039***		-0.068***
	(0.011)	(0.009)	(0.012)		(0.015)
L_c (10 %)	-0.005	-0.083***	-0.020*		0.027**
	(0.008)	(0.006)	(0.010)		(0.014)
$2\ln(\beta_c)$				1.048***	
				(0.071)	
$2\ln(\theta_i)$				-0.625***	
				(0.043)	
$2\ln(\sigma_c)$				1.004***	
				(0.028)	
constant	1.047***	7.324***	4.203***	-2.636***	
	(0.193)	(0.206)	(0.299)	(0.356)	
Obs.	579	579	579	579	
R-squared	0.057	0.255	0.308	0.831	

Notes: Robust standard errors are in paretheses; *, **, and *** denote p<0.1, p<0.05, p<0.01, respectively.

Table 4. Shapley values for each growing condition variables on systemic risk, %

	(1)	(2)	(3)	(4)	(5)
	$2\ln(\tau_c)$	$2\ln(\beta_c)$	$2\ln(\theta_i)$	$2\ln(\sigma_c)$	$2\ln(\tau_c)$
$2\ln(\beta_c)$	28.35				
$2\ln(\theta_i)$	6.27				
$2\ln(\sigma_c)$	65.38				
G_c		46.92	9.50	42.97	41.99
S_c		10.29	2.28	36.00	26.59
D_c		5.53	1.95	6.42	5.89
W_c		35.53	0.30	9.60	16.37
L_c		1.74	85.98	5.01	9.16
Total	100	100	100	100	100

Table 5. Sign and significance of η_c among 579 Corn Belt counties

Sign and P-value of η_c	N	Percent
$\eta_c > 0$ & P-value≤0.1	282	48.7
$\eta_c > 0$ & P-value>0.1	79	13.6
$\eta_c \leq 0$ & P-value≤0.1	143	24.7
$\eta_c \leq 0$ & P-value>0.1	75	13.0

Table 6. Number of counties with different number of APH groups

Number of APH groups	Number of counties	Percent of counties with associated APH groups
1	37	6.9
2	54	10.0
3	60	11.1
4	87	16.0
5	137	25.4
6	123	22.8
7	40	7.4
8	1	0.2
9	1	0.2

Table 7. Descriptive statistics for unit-level TRASR, conditional on MPCI coverage level and AYP coverage level

	N	Mean	St.Dev	Min	Median	Max
$\phi_i = 50\%, \phi_c = 70\%$	208,549	28.3	70.5	0.00	12.4	3,421
$\phi_i = 50\%, \phi_c = 75\%$	208,549	19.0	37.5	0.00	9.2	1,084
$\phi_i = 50\%, \phi_c = 80\%$	208,549	13.6	23.3	0.00	6.9	482
$\phi_i = 50\%, \phi_c = 85\%$	208,549	10.0	15.8	0.00	5.1	280
$\phi_i = 50\%, \phi_c = 90\%$	208,549	7.5	11.3	0.00	3.8	179
$\phi_i = 55\%, \phi_c = 70\%$	208,549	38.1	84.5	0.00	19.4	4,022
$\phi_i = 55\%, \phi_c = 75\%$	208,549	25.7	44.7	0.00	14.2	1,275
$\phi_i = 55\%, \phi_c = 80\%$	208,549	18.3	27.6	0.00	10.6	566
$\phi_i = 55\%, \phi_c = 85\%$	208,549	13.5	18.6	0.00	8.0	328
$\phi_i = 55\%, \phi_c = 90\%$	208,549	10.1	13.3	0.00	5.9	199
$\phi_i = 60\%, \phi_c = 70\%$	208,549	51.4	101.0	0.01	29.0	4,711
$\phi_i = 60\%, \phi_c = 75\%$	208,549	34.5	53.1	0.00	21.3	1,493
$\phi_i = 60\%, \phi_c = 80\%$	208,549	24.6	32.6	0.00	15.9	659
$\phi_i = 60\%, \phi_c = 85\%$	208,549	18.1	21.9	0.00	11.9	382
$\phi_i = 60\%, \phi_c = 90\%$	208,549	13.5	15.6	0.00	8.9	232
$\phi_i = 65\%, \phi_c = 70\%$	208,549	69.2	120.8	0.05	42.1	5,500
$\phi_i = 65\%, \phi_c = 75\%$	208,549	46.4	62.9	0.04	30.9	1,743
$\phi_i = 65\%, \phi_c = 80\%$	208,549	33.0	38.3	0.03	23.0	761
$\phi_i = 65\%, \phi_c = 85\%$	208,549	24.2	25.5	0.02	17.3	442
$\phi_i = 65\%, \phi_c = 90\%$	208,549	18.0	18.1	0.02	12.9	268
$\phi_i = 70\%, \phi_c = 70\%$	208,549	93.0	144.4	0.26	60.1	6,397
$\phi_i = 70\%, \phi_c = 75\%$	208,549	62.2	74.4	0.19	43.9	2,028
$\phi_i = 70\%, \phi_c = 80\%$	208,549	44.1	44.9	0.15	32.7	873
$\phi_i = 70\%, \phi_c = 85\%$	208,549	32.2	29.6	0.11	24.5	507
$\phi_i = 70\%, \phi_c = 90\%$	208,549	23.9	20.8	0.09	18.3	307
$\phi_i = 75\%, \phi_c = 70\%$	208,549	125.1	173.1	1.07	85.1	7,408
$\phi_i = 75\%, \phi_c = 75\%$	208,549	83.3	88.0	0.80	61.7	2,348
$\phi_i = 75\%, \phi_c = 80\%$	208,549	58.9	52.4	0.60	45.7	996
$\phi_i = 75\%, \phi_c = 85\%$	208,549	42.8	34.1	0.47	34.2	578
$\phi_i = 75\%, \phi_c = 90\%$	208,549	31.8	23.8	0.37	25.6	351
$\phi_i = 80\%, \phi_c = 70\%$	208,549	168.2	208.5	3.62	119.6	8,538
$\phi_i = 80\%, \phi_c = 75\%$	208,549	111.5	104.2	2.69	86.3	2,706
$\phi_i = 80\%, \phi_c = 80\%$	208,549	78.6	61.0	2.03	63.6	1,129
$\phi_i = 80\%, \phi_c = 85\%$	208,549	57.0	39.1	1.60	47.4	655
$\phi_i = 80\%, \phi_c = 90\%$	208,549	42.1	26.9	1.25	35.5	397
$\phi_i = 85\%, \phi_c = 70\%$	208,549	226.1	252.8	9.68	165.5	9,791
$\phi_i = 85\%, \phi_c = 75\%$	208,549	149.4	123.9	7.19	120.1	3,103
$\phi_i = 85\%, \phi_c = 80\%$	208,549	104.9	71.1	5.44	88.1	1,273
$\phi_i = 85\%, \phi_c = 85\%$	208,549	75.8	44.6	4.28	65.3	739
$\phi_i = 85\%, \phi_c = 85\%$	208,549	55.8	30.1	3.34	48.8	448

Table 8. Premium subsidy rates for individual and area insurance contracts, conditional on coverage level

Insurance Plan	Coverage Level (%)									
	CAT	50	55	60	65	70	75	80	85	90
Basic and Optional Units	100	67	64	64	59	59	55	48	38	n/a
Enterprise Units	n/a	80	80	80	80	80	77	68	53	n/a
Area Yield Plans	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

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

Table 9. Coverage-level conditional percent of units whose TASR is higher than the current AYP subsidy rate and percent of units whose TASR is higher than 100%

MPCI coverage level	AYP coverage level				
	70%	75%	80%	85%	90%
	Panel A: Percent of units whose TASR>current AYP subsidy rate				
75%	57.2	35.1	21.4	10.9	6.7
80%	71.8	49.5	30.8	14.2	8.2
85%	76.6	56.4	34.9	14.1	7.2
	Panel B: Percent of units whose TASR>100%				
75%	25.5	11.8	5.8	2.6	1.0
80%	36.3	16.7	7.1	3.1	1.1
85%	40.9	18.4	6.9	2.8	0.8

Table 10. Coveragel-level conditional Pearson's correlation test coefficients between county systemic risk and county-level TRASR variables

MPCI coverage level	AYP coverage level				
	70%	75%	80%	85%	90%
	Panel A: Pearson's correlation test coefficients between systemic risk and percent of units whose TASR>current AYP subsidy rate				
75%	0.1397*	0.0483	-0.0367	-0.1118*	-0.1112*
80%	0.1417*	0.0785*	-0.0041	-0.1106*	-0.1302*
85%	0.0541	-0.0031	-0.0568	-0.1716*	-0.1908*
	Panel B: Pearson's correlation test coefficients between systemic risk and percent of units whose TASR>100%				
75%	-0.0973*	-0.1741*	-0.2111*	-0.2110*	-0.1684*
80%	-0.0885*	-0.1624*	-0.2184*	-0.2230*	-0.1767*
85%	-0.1343*	-0.1812*	-0.2244*	-0.2162*	-0.1804*

Note: * denotes p<0.1.

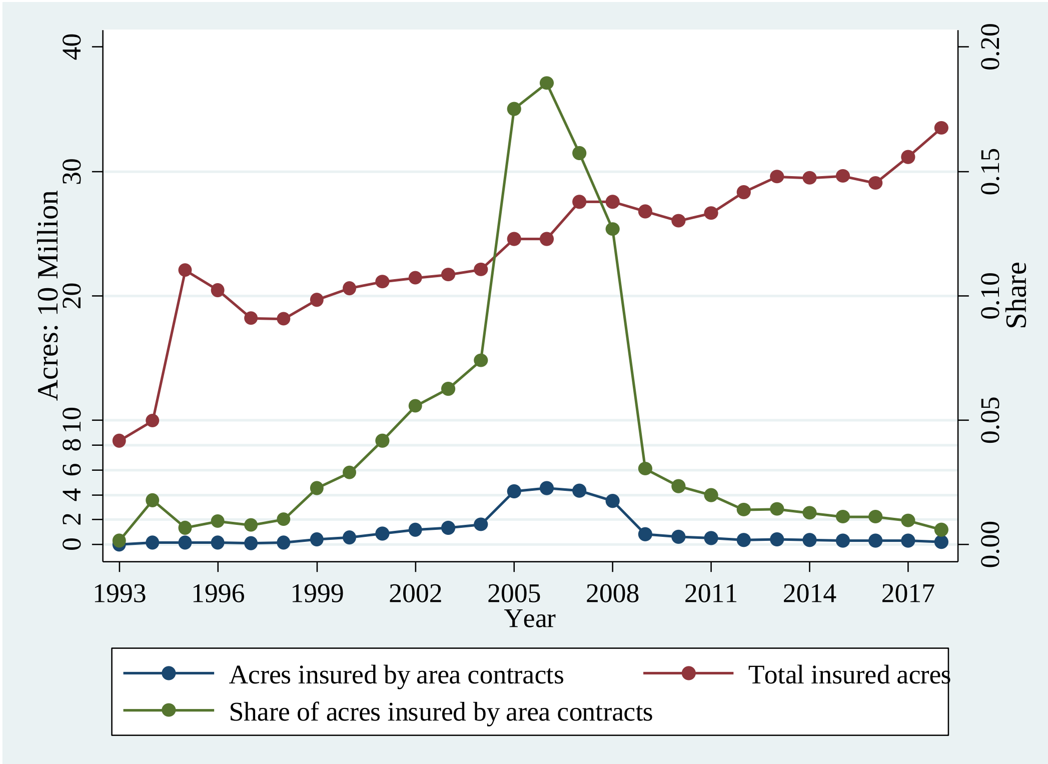

Figure 1. Share of acres insured by area insurance contracts for all crops, 1993-2018

Notes: Source: Summary of Business, 1993-2018, Risk Management Agency (RMA).
This figure plots acres insured by area insurance contracts and acres insured by any kind of crop insurance for all crops over 1993-2018 (left y-axis), and plots the share of all insured acres that are insured by area insurance contracts for all crops over 1993-2018 (right y-axis). Area insurance before year 2014 are Group Risk Plan (GRP), Group Risk Income Protection (GRIP), Group Risk Income Protection with Harvest Revenue Option (GRIP-HRO). Area insurance starting in 2014 are under Area Yield Protection (AYP), Area Revenue Protection (ARP), and Area Revenue Protection with Harvest Price Exclusion (ARP-HPE).

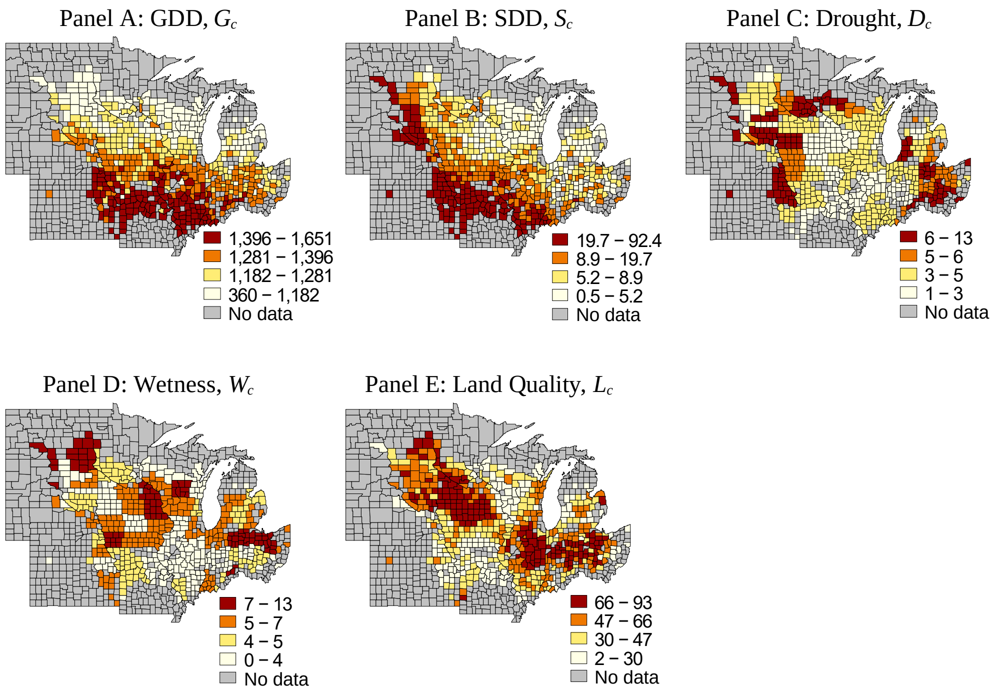

Figure 2. Geographic distributions of county growing conditions

Notes: This figure plots geographic distributions of county GDD, county SDD, frequency of severe drought, frequency of severe wetness, and the proportion of county land that is in land capability classes I or II. Numbers in legends are quartile ranges.

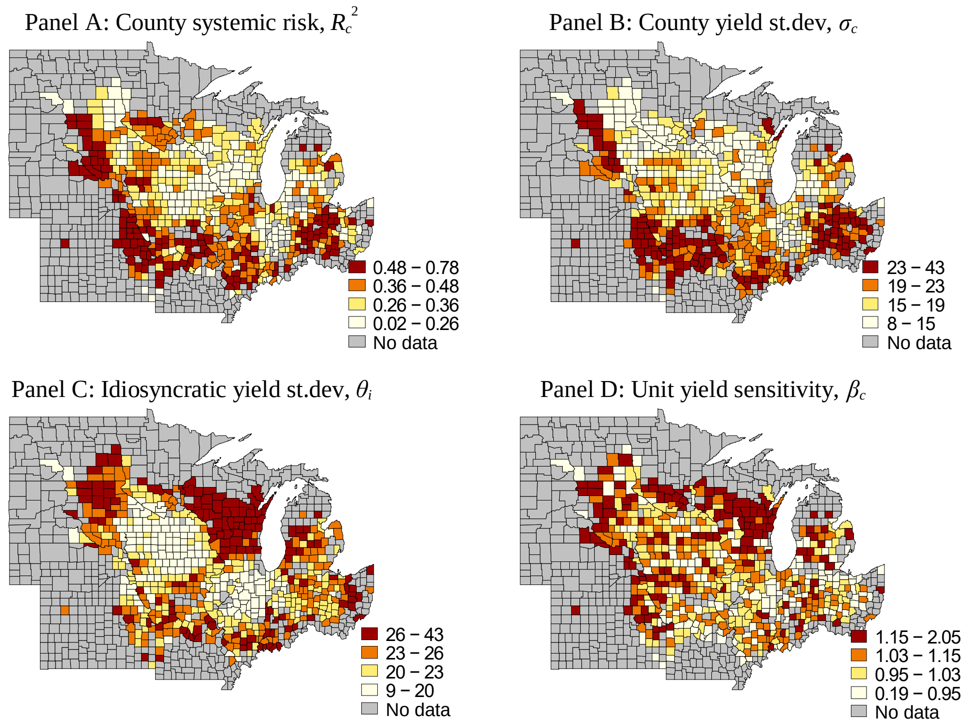

Figure 3. Geographic distributions of county systemic risk and its components

Note: This figure plots geographic distributions of county systemic risk, county yield standard deviation, unit's idiosyncratic yield standard deviation, and unit yield's sensitivity to county yield. Numbers in legends are quartile ranges.

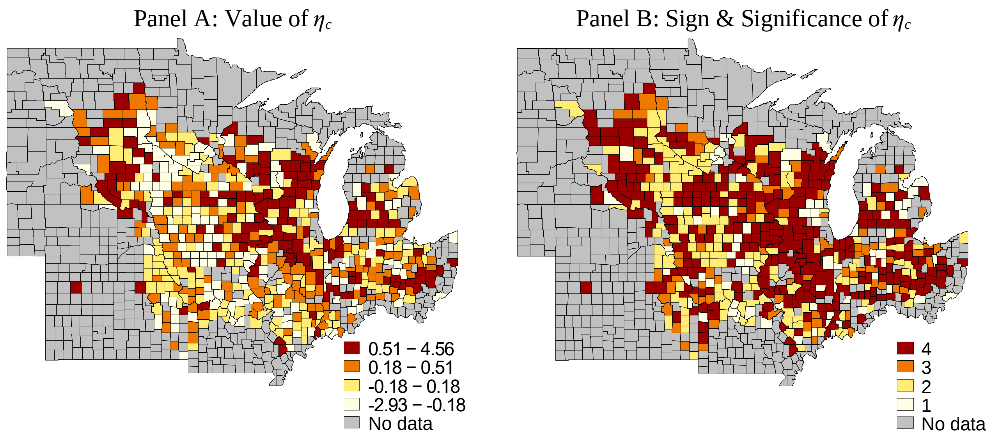

Figure 4. Geographic distributions of η_c's value, and sign/significance category.

Note: This figure plots geographic distributions of η_c. Panel A plots the geographic distribution of η_c' values and numbers in the legend are quartile ranges. Panel B plots the geographic distribution of η_c's sign and significance. Counties labeled 4 are significantly positive η_c, counties labeled 3 are insignificantly positive η_c, counties labeled 2 are significantly negative η_c and counties labeled 1 are insignificantly negative η_c.

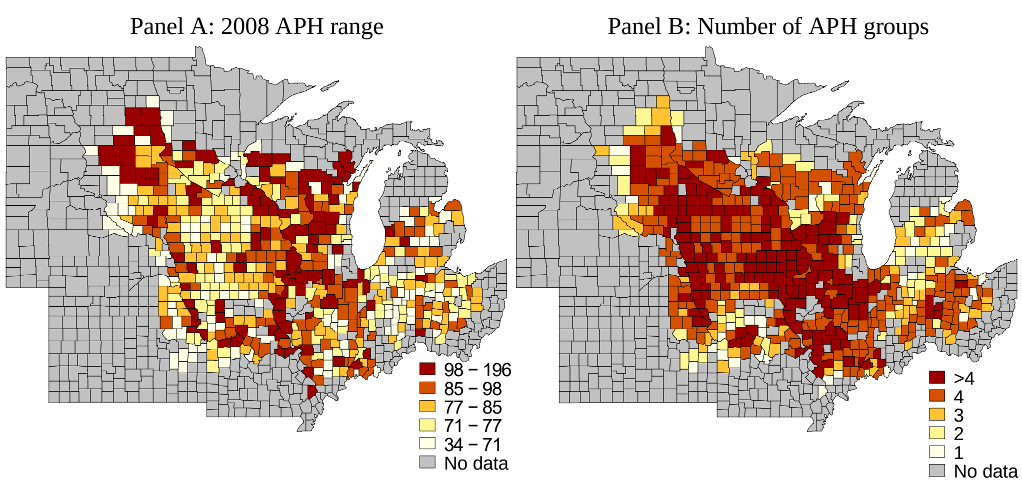

Figure 5. Geographic distributions of 2008 APH range and number of APH groups

Notes: The 2008 APH range equals the maximum 2008 county APH value minus the minimum 2008 county APH value. Numbers in the legend of Panel A are quintile ranges.

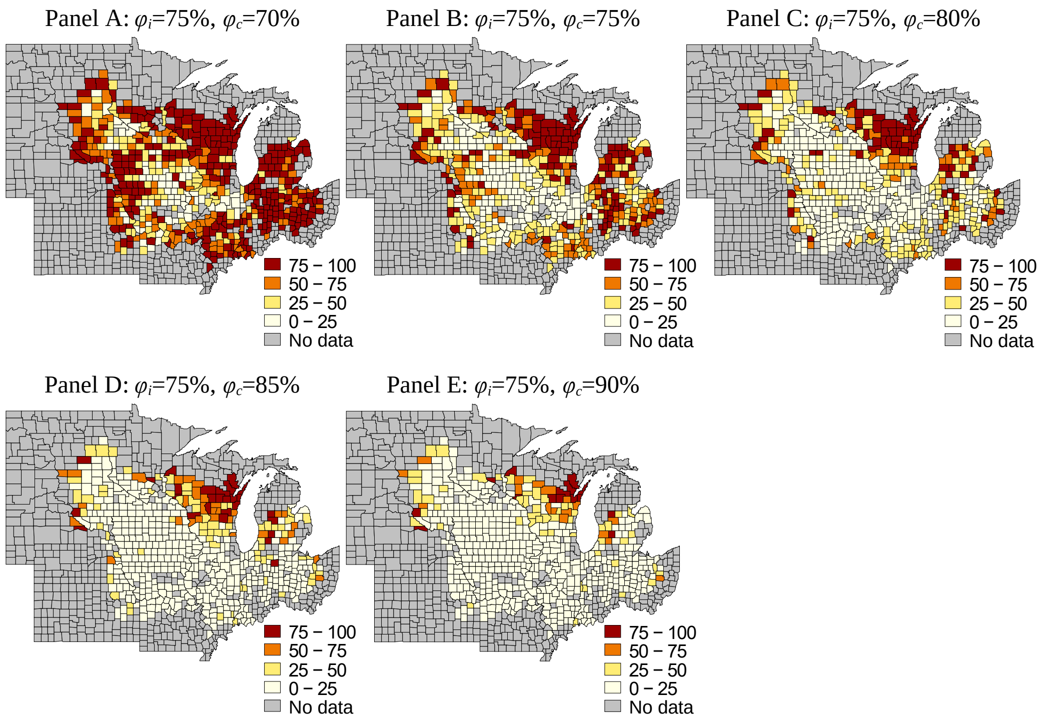

Figure 6. Geographic distributions of county-level percent of units whose TASR is higher than the current AYP subsidy rate, conditional on MPCI coverage level = 75%

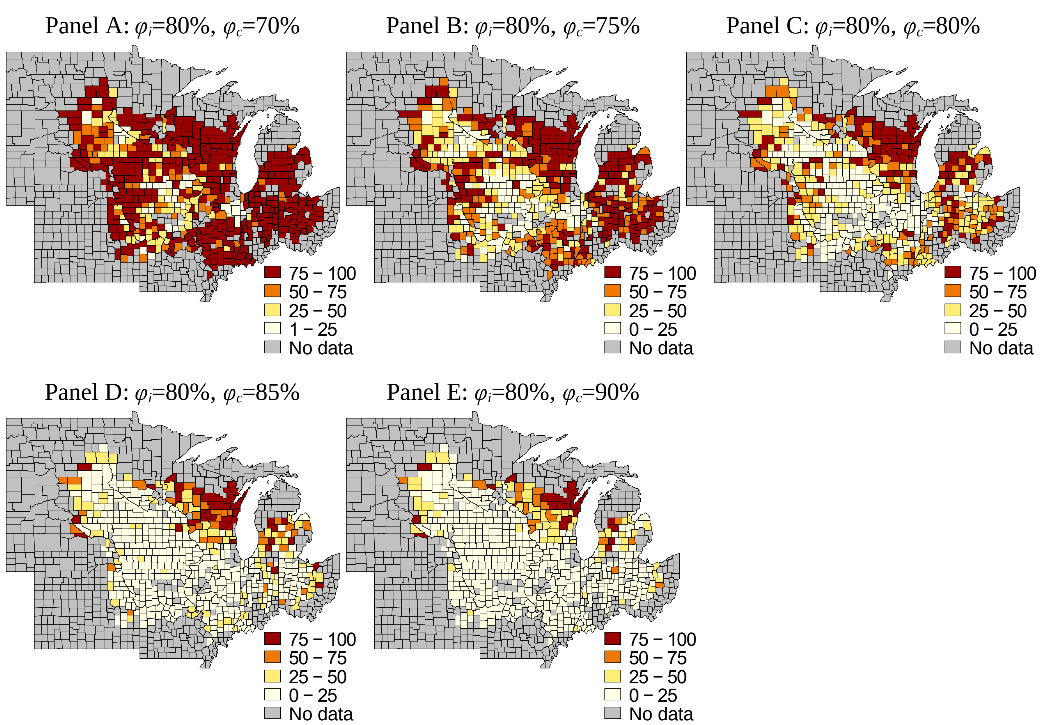

Figure 7. Geographic distributions of county-level percent of units whose TASR exceeds the current AYP subsidy rate, conditional on MPCI coverage level = 80%

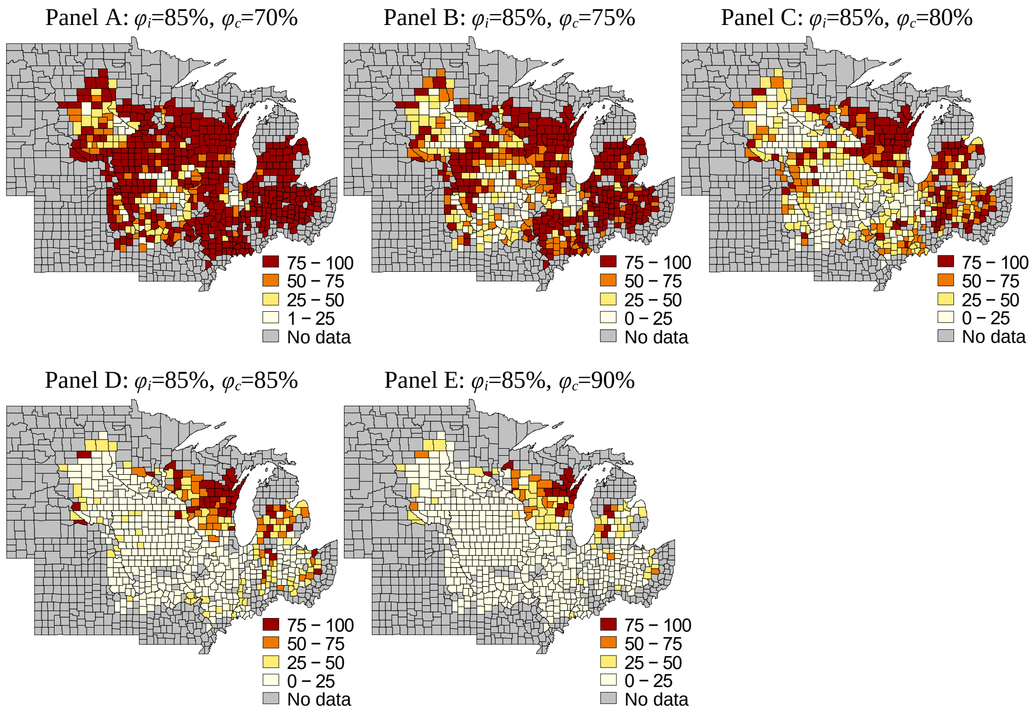

Figure 8. Geographic distributions of county-level percent of units whose TASR exceeds the current AYP subsidy rate, conditional on MPCI coverage level = 85%

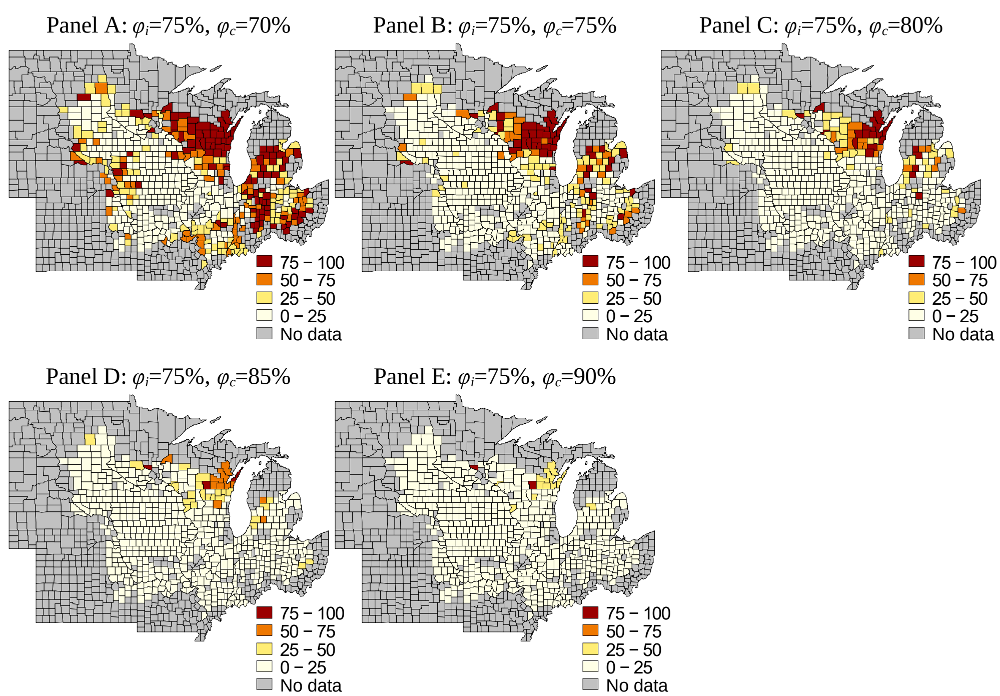

Figure 9. Geographic distributions of county-level percent of units whose TASR exceeds 100%, conditional on MPCI coverage level = 75%

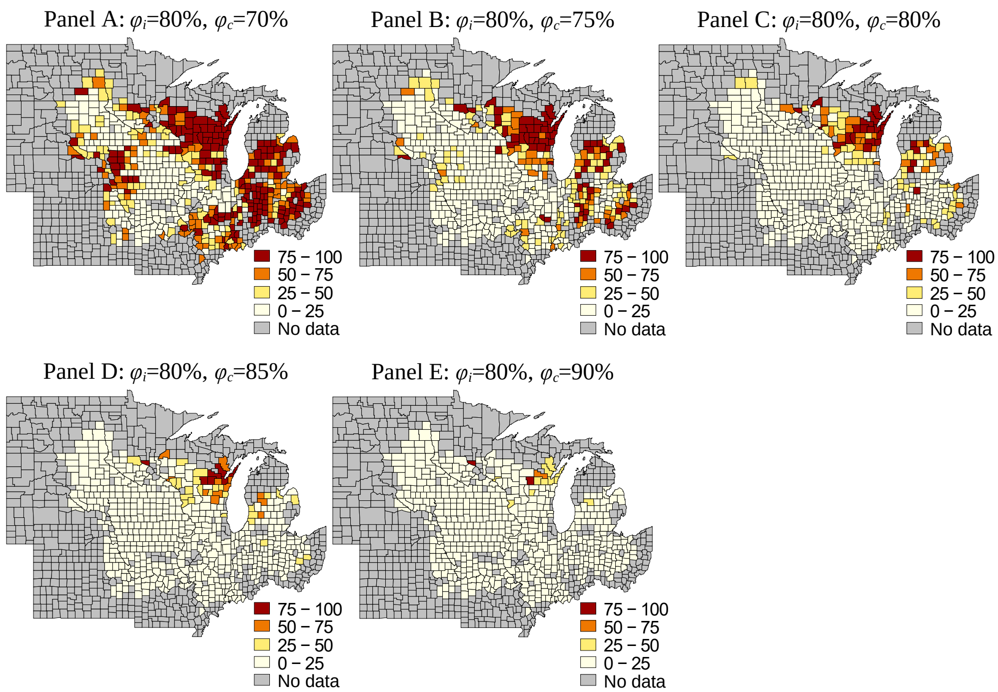

Figure 10. Geographic distributions of county-level percent of units whose TASR exceeds 100%, conditional on MPCI coverage level = 80%

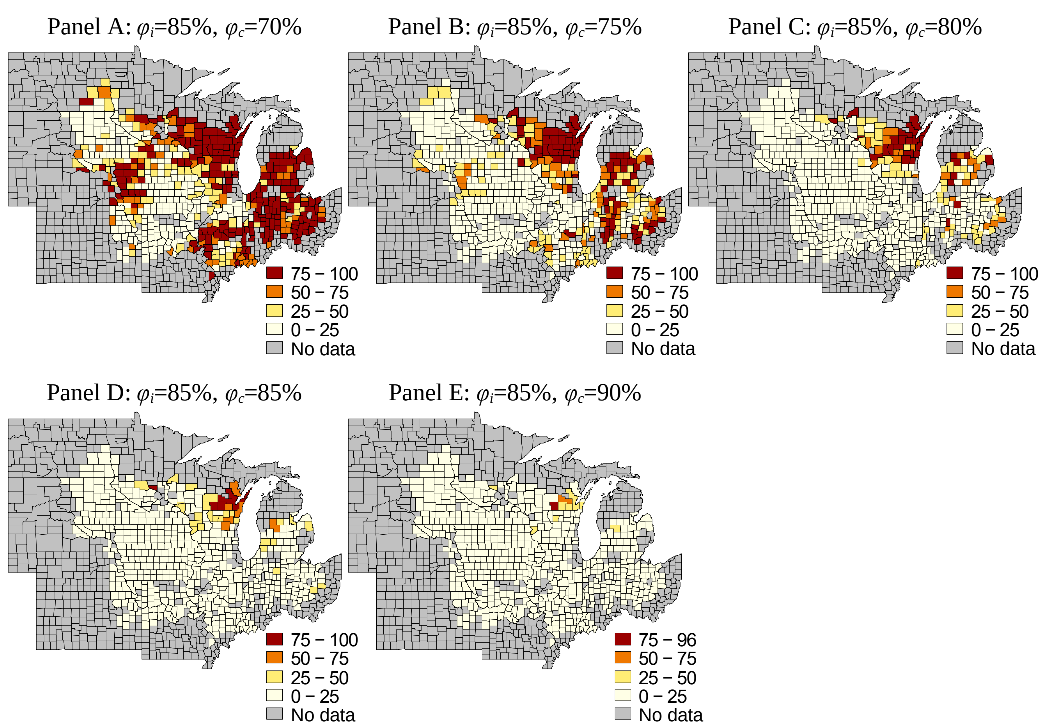

Figure 11. Geographic distributions of county-level percent of units whose TASR exceeds 100%, conditional on MPCI coverage level = 85%

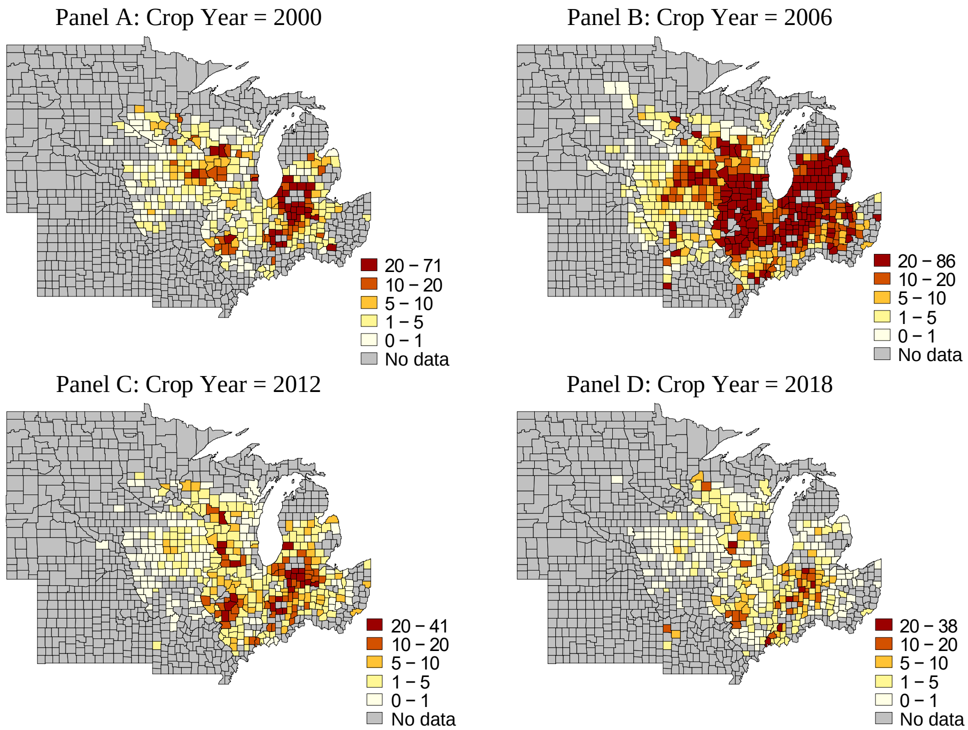

Figure 12. Geographic distribution of share of corn acres insured by area insurance contracts

Notes: Data Source: Summary of Business, 2000, 2006, 2012, and 2018, RMA. This figure plots geographic distributions of share of corn acres insured by area insurance contracts for year 2000, 2006, 2012, and 2018. For each county, the share is calculated as the rate of all corn acres insured by area insurance contracts over all corn acres insured by any kind of insurance contract.

APPENDIX B

Supplementary Tables

Table 11. Observation loss after each data screening and merging step

	Number of counties left	Number of units left
Initial observations	1,031	14,700,150
Drop units with any non-actual historical yield types	1,003	5,455,480
Drop units with historical year before 1998	980	2,669,560
Drop irrigated units and counties with irrigation rate higher than 20%	847	2,267,540
Drop counties with less than 50-year data records between 1958 and 2007	712	2,234,220
Drop counties with less than 30 units	618	2,220,260
Merge with weather and land quality data	579	2,133,320

Table 12. Coverage-level conditional percent of units whose TASR is higher than the current AYP subsidy rate and percent of units whose TASR is higher than 100%

MPCI coverage level	AYP coverage level				
	70%	75%	80%	85%	90%
	Panel A: Percent of units whose TASR>current AYP subsidy rate				
50%	7.1	4.3	2.7	1.3	0.7
55%	10.0	6.0	3.9	2.0	1.1
60%	14.6	8.6	5.7	3.1	1.8
65%	23.2	13.1	8.5	4.7	2.9
70%	38.0	21.5	13.4	7.3	4.6
75%	57.2	35.1	21.4	10.9	6.7
80%	71.8	49.5	30.8	14.2	8.2
85%	76.6	56.4	34.9	14.1	7.2
	Panel B: Percent of units whose TASR>100%				
50%	3.4	1.7	0.8	0.3	0.0
55%	4.7	2.5	1.2	0.5	0.1
60%	6.6	3.6	1.7	0.7	0.2
65%	9.8	5.3	2.6	1.1	0.4
70%	15.8	7.9	4.0	1.8	0.7
75%	25.5	11.8	5.8	2.6	1.0
80%	36.3	16.7	7.1	3.1	1.1
85%	40.9	18.4	6.9	2.8	0.8

Table 13. Coveragel-level conditional Pearson's correlation test coefficients between county systemic risk and county-level TRASR variables

MPCI coverage level	AYP coverage level				
	70%	75%	80%	85%	90%
	Panel A: Pearson's correlation test coefficients between systemic risk and percent of units whose TASR>current AYP subsidy rate				
50%	-0.1360*	-0.1543*	-0.1504*	-0.1391*	-0.1080*
55%	-0.1238*	-0.1515*	-0.1514*	-0.1467*	-0.1201*
60%	-0.1024*	-0.1291*	-0.1407*	-0.1495*	-0.1256*
65%	-0.0273	-0.1026*	-0.1194*	-0.1391*	-0.1360*
70%	0.0665	-0.0382	-0.0897*	-0.1292*	-0.1167*
75%	0.1397*	0.0483	-0.0367	-0.1118*	-0.1112*
80%	0.1417*	0.0785*	-0.0041	-0.1106*	-0.1302*
85%	0.0541	-0.0031	-0.0568	-0.1716*	-0.1908*
	Panel B: Pearson's correlation test coefficients between systemic risk and percent of units whose TASR>100%				
50%	-0.1866*	-0.1841*	-0.1549*	-0.0939*	-0.0380
55%	-0.1846*	-0.1915*	-0.1715*	-0.1209*	-0.0770*
60%	-0.1803*	-0.1962*	-0.1871*	-0.1527*	-0.0935*
65%	-0.1701*	-0.1971*	-0.2060*	-0.1720*	-0.1137*
70%	-0.1387*	-0.1855*	-0.2039*	-0.1963*	-0.1453*
75%	-0.0973*	-0.1741*	-0.2111*	-0.2110*	-0.1684*
80%	-0.0885*	-0.1624*	-0.2184*	-0.2230*	-0.1767*
85%	-0.1343*	-0.1812*	-0.2244*	-0.2162*	-0.1804*

Note: * denotes p<0.1

REFERENCES

REFERENCES

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

Barnett, B.J., Black, J.R., Hu, Y. and Skees, J.R., 2005. Is area yield insurance competitive with farm yield insurance?. *Journal of Agricultural and Resource Economics*, pp.285-301.

Bourgeon, J.M. and Chambers, R.G., 2003. Optimal area-yield crop insurance reconsidered. *American Journal of Agricultural Economics*, 85(3), pp.590-604.

Bulut, H., Collins, K.J. and Zacharias, T.P., 2012. Optimal coverage level choice with individual and area insurance plans. *American Journal of Agricultural Economics*, 94(4), pp.1013-1023.

Botts, R.R. and Boles, J.N., 1958. Use of normal-curve theory in crop insurance ratemaking. *Journal of Farm Economics*, 40(3), pp.733-740.

Carriquiry, M.A., Babcock, B.A. and Hart, C.E., 2008. Using a farmer's beta for improved estimation of expected yields. *Journal of Agricultural and Resource Economics*, pp.52-68.

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

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

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

Coble, K.H. and Barnett, B.J., 2008. Implications of integrated commodity programs and crop insurance. *Journal of Agricultural and Applied Economics*, 40(2), pp.431-442.

Coble, K.H. and Dismukes, R., 2008. Distributional and risk reduction effects of commodity revenue program design. *Review of Agricultural Economics*, 30(3), pp.543-553.

Coble, K.H., Heifner, R.G. and Zuniga, M., 2000. Implications of crop yield and revenue insurance for producer hedging. *Journal of Agricultural and Resource Economics*, pp.432-452.

Coble, K.H., Knight, T.O., Pope, R.D. and Williams, J.R., 1996. Modeling farm-level crop insurance demand with panel data. *American Journal of Agricultural Economics*, 78(2), pp.439-447.

Congressional Budget Office, 2017. Options to reduce the budgetary costs of the federal crop insurance program. Reports 53375. December. Congressional Budget Office. https://www.cbo.gov/publication/53375

Cooper, J., Zulauf, C., Langemeier, M. and Schnitkey, G., 2012. Implications of within county yield heterogeneity for modeling crop insurance premiums. *Agricultural Finance Review*, 72(1), pp.134-155.

Deng, X., Barnett, B.J. and Vedenov, D.V., 2007. Is there a viable market for area-based crop insurance?. *American Journal of Agricultural Economics*, 89(2), pp.508-519.

Dismukes, R., Arriola, C. and Coble, K.H., 2010. *ACRE Program Payments and Risk Reduction: An Analysis Based on Simulations of Crop Revenue Variability* (No. 262248). United States Department of Agriculture, Economic Research Service.

Du, X., Feng, H. and Hennessy, D.A., 2016. Rationality of choices in subsidized crop insurance markets. *American Journal of Agricultural Economics*, 99(3), pp.732-756.

Du, X., Hennessy, D.A., Feng, H. and Arora, G., 2017. Land resilience and tail dependence among crop yield distributions. *American Journal of Agricultural Economics*, *100*(3), pp.809-828.

Du, X., Yu, C.L., Hennessy, D.A. and Miao, R., 2015. Geography of crop yield skewness. *Agricultural Economics*, 46(4), pp.463-473.

Duncan, J. and Myers, R.J., 2000. Crop insurance under catastrophic risk. *American Journal of Agricultural Economics*, 82(4), pp.842-855.

Edwards, W.M., 2011. *Proven yields and insurance units for crop insurance* (No. 34425). Iowa State University, Department of Economics.

Edwards, E., 2012. Trend-Adjusted Actual Production History (APH). *Ag Decision Maker, Iowa State University Extension and Outreach*. https://www.extension.iastate.edu/agdm/crops/html/a1-56.html

Elabed, G., Bellemare, M.F., Carter, M.R. and Guirkinger, C., 2013. Managing basis risk with multiscale index insurance. *Agricultural Economics*, 44(4-5), pp.419-431.

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

Goodwin, B.K. and Hungerford, A., 2014. Copula-based models of systemic risk in US agriculture: implications for crop insurance and reinsurance contracts. *American Journal of Agricultural Economics*, 97(3), pp.879-896.

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

Goodwin, B.K. and Smith, V.H., 1995. *The Economics of Crop Insurance and Disaster Aid*. American Enterprise Institute. September 1. Washington, D.C..

Goodwin, B.K. and Smith, V.H., 2013. What harm is done by subsidizing crop insurance?. *American Journal of Agricultural Economics*, 95(2), pp.489-497.

Goodwin, B.K., Vandeveer, M.L. and Deal, J.L., 2004. An empirical analysis of acreage effects of participation in the federal crop insurance program. *American Journal of Agricultural Economics*, 86(4), pp.1058-1077.

Halcrow, H.G., 1949. Actuarial structures for crop insurance. *Journal of Farm Economics*, 31(3), pp.418-443.

Hill, R.V., Robles, M. and Ceballos, F., 2016. Demand for a simple weather insurance product in India: theory and evidence. *American Journal of Agricultural Economics*, 98(4), pp.1250-1270.

Harri, A., Coble, K.H., Ker, A.P. and Goodwin, B.J., 2011. Relaxing heteroscedasticity assumptions in area-yield crop insurance rating. *American Journal of Agricultural Economics*, 93(3), pp.707-717.

Huettner, F. and Sunder, M., 2012. Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values. *Electronic Journal of Statistics*, 6, pp.1239-1250.

Jensen, N.D., Mude, A.G. and Barrett, C.B., 2018. How basis risk and spatiotemporal adverse selection influence demand for index insurance: Evidence from northern Kenya. *Food Policy*, 74, pp.172-198.

Karl, T.R., 1986. The sensitivity of the Palmer Drought Severity Index and Palmer's Z-index to their calibration coefficients including potential evapotranspiration. *Journal of Climate and Applied Meteorology*, 25(1), pp.77-86.

Ker, A.P. and Goodwin, B.K., 2000. Nonparametric estimation of crop insurance rates revisited. *American Journal of Agricultural Economics*, 82(2), pp.463-478.

Van Kerm, P., 2003. Adaptive kernel density estimation. *The Stata Journal*, 3(2), pp.148-156.

Lusk, J.L., 2016. Distributional effects of crop insurance subsidies. *Applied Economic Perspectives and Policy*, 39(1), pp.1-15.

Neild, R.E. and Newman, J.E., 1987. *Growing season characteristics and requirements in the Corn Belt*. Cooperative Extension Service, Iowa State University.

Nelson, C.H., 1990. The influence of distributional assumptions on the calculation of crop insurance premia. *North Central Journal of Agricultural Economics*, *12*(1), pp.71-78.

Mahul, O., 1999. Optimum area yield crop insurance. *American Journal of Agricultural Economics*, 81(1), pp.75-82.

Miranda, M.J., 1991. Area-yield crop insurance reconsidered. *American Journal of Agricultural Economics*, 73(2), pp.233-242.

National Oceanic and Atmospheric Administration. 2014. Statewide Regional Drought. Available at: ftp://ftp.ncdc.noaa.gov/pub/data/cirs/climdiv/drought-readme.txt

Newton, H.J., 1988. *Timeslab: a time series analysis laboratory*. Wadsworth Publ. Co..

Miranda, M.J. and Glauber, J.W., 1997. Systemic risk, reinsurance, and the failure of crop insurance markets. *American Journal of Agricultural Economics*, 79(1), pp.206-215.

Okhrin, O., Odening, M. and Xu, W., 2013. Systemic weather risk and crop insurance: the case of China. *Journal of Risk and Insurance*, 80(2), pp.351-372.

O'Donoghue, E., 2014. The effects of premium subsidies on demand for crop insurance. *USDA-ERS Economic Research Report*, July 1 (169). Washington, D.C.

Quiggin, J.C., Karagiannis, G. and Stanton, J., 1993. Crop insurance and crop production: an empirical study of moral hazard and adverse selection. *Australian Journal of Agricultural Economics*, 37(429-2016-29192), p.95.

Parzen, E., 1962. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3), pp.1065-1076.

Schlenker, W., Hanemann, W.M. and Fisher, A.C., 2006. The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions. *Review of Economics and Statistics*, 88(1), pp.113-125.

Schlenker, W. and Roberts, M.J., 2009. Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37), pp.15594-15598.

Schnitkey, G. and B. Sherrick., 2014. Coverage Levels on Crop Insurance and the SCO Alternative. *Farmdoc Daily* (4):78, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, April 29. https://farmdocdaily.illinois.edu/2014/04/coverage-levels-on-crop-insurance-and-sco.html

Shaik, S., Coble, K.H., Knight, T.O., Baquet, A.E. and Patrick, G.F., 2008. Crop revenue and yield insurance demand: a subjective probability approach. *Journal of Agricultural and Applied Economics*, 40(1379-2016-112787), p.757.

Sherrick, B. and G. Schnitkey., 2016.Crop Insurance Decisions for 2016. *Farmdoc Daily* (6):40, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, March 1. https://farmdocdaily.illinois.edu/wp-content/uploads/2016/04/fdd010316.pdf

Sherrick, B.J., Zanini, F.C., Schnitkey, G.D. and Irwin, S.H., 2004. Crop insurance valuation under alternative yield distributions. *American Journal of Agricultural Economics*, 86(2), pp.406-419.

Skees, J.R., Black, J.R. and Barnett, B.J., 1997. Designing and rating an area yield crop insurance contract. *American Journal of Agricultural Economics*, 79(2), pp.430-438.

Shen, Z. and Odening, M., 2013. Coping with systemic risk in index-based crop insurance. *Agricultural Economics*, 44(1), pp.1-13.

Smith, V.H., Chouinard, H.H. and Baquet, A.E., 1994. Almost ideal area yield crop insurance contracts. *Agricultural and Resource Economics Review*, 23(1), pp.75-83.

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

Tack, J.B. and Holt, M.T., 2016. The influence of weather extremes on the spatial correlation of corn yields. *Climatic Change*, 134(1-2), pp.299-309.

Tobin, J., 1958. Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2), pp.65-86.

Tapia, R.A. and Thompson, J.R., 1978. *Nonparametric probability density estimation*. Baltimore, MD: Johns Hopkins Univ., 1978.

Vercammen, J.A., 2000. Constrained efficient contracts for area yield crop insurance. *American Journal of Agricultural Economics*, 82(4), pp.856-864.

Xu, Z., Hennessy, D.A., Sardana, K. and Moschini, G., 2013. The realized yield effect of genetically engineered crops: US maize and soybean. *Crop Science*, 53(3), pp.735-745.

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

Yu, J., Smith, A. and Sumner, D.A., 2017. Effects of crop insurance premium subsidies on crop acreage. *American Journal of Agricultural Economics*, 100(1), pp.91-114.

Zulauf, C.R., Demircan, V., Scnhitkey, G., Arthur, B., Ibendahl, G. and Herbel, K., 2013. *Examining Contemporaneous Farm and County Losses Using Farm Level Data* (No. 157812). Agricultural and Applied Economics Association.

Zulauf, C., G. Schnitkey, N. Paulson, and J. Coppess., 2017. Comparing NASS and RMA County Yields for Corn. *Farmdoc Daily* (7):202, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, November 2. https://farmdocdaily.illinois.edu/2017/11/comparing-nass-and-rma-county-yields-for-corn.html