Estimating Distributional Impacts of Federal Crop Insurance Program

Jialing Yu
Department of Applied Economics
Oregon State University

Selected Paper prepared for presentation for the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28.

Introduction

Agricultural production faces risks from various sources, such as, weather conditions, pests, natural disasters, management errors, diseases, and price fluctuations. Federal crop insurance program first authorized in 1938 has become a major risk management tool for the government. In 2013, crop insurance covered 295 million acres with federal expenditure of $ 6.0 billion (Shields 2013). Furthermore, the 2014 Farm Bill expands the federal crop insurance outlay by $ 5.7 billion for the next ten years; it mostly stems from two newly introduce crop insurance products: Stacked Income Protection Plan (STAX) for cotton, and Supplemental Coverage Option (SCO) for other crops (Chite 2014). The new insurance products provide area-based coverage in combination with traditional crop insurance policies, and offers more subsidy than the traditional ones at higher coverage levels. As crop insurance plays a more vital role in farm programs, the important empirical questions to ask are what policy impacts on the production net return, and how they are distributed across population. This study focuses on the application in the Corn Belt region.

Crop insurance is a voluntary program, which makes program participation a critical issue for the government to deliver the program and also for researchers to correctly assess impacts. The government has increased premium subsidies several times in the history to encourage participation. Glauber (2012) concludes that with adequate subsidy, farmers will buy the insurance. Yet not all producers participate in the program. It is essential to understand participation process in that it selects the population who will be affected, and therefore determines the distribution and magnitudes of policy impacts. Correspondingly, farmers making participation decisions based on potential components of outcomes gives rise to the selection problem in the policy evaluation. In other words, participants select themselves into the insurance program because they might gain more when participate, and therefore observable outcomes for participants with insurance are systematically different from that for non-participants which are unobservable. The impact assessment requires constructing counterfactuals for both participants and non-participants. Furthermore, the procedure is complicated with the fact that the observed production outcomes are not the decision-relevant expected outcomes (ex-ante parameters). That is, the decision making process is based on the expected outcomes (ex-ante) with information set available at the beginning of the season, the ex-post outcomes are observed after all uncertainties are realized. How to recover ex-ante parameters from ex-post parameters is also a challenge in impact assessment.

There are extensive econometric studies dedicated to solving the selection problem by in the field of policy evaluation and technology adoption (Suri 2011), such as, instrumental variable approach and difference in difference. In the crop insurance literature, most studies focus on impact analysis or determination of influencing factors for participation; they solve the selection problem by using instrumental variables. The IVs chosen are not related to outcome but related to participation, such as, intra-state variation in premiums, lagged yield, loss ratio and insurance decision (Walters 2012, Schoengold, Ding and Headlee 2014). Another related strand of studies treat participation decision as simultaneous decision with other production decisions, such as input uses, land allocations and etc. they participation is modeled as part of s simultaneous decision-making system, Wu (1999), Serra and Goodwin (2003) and Shaik, Coble and Knight (2005). These structural econometric studies require extensive and detailed data from surveys that are not always easy to access. Alternatively, this study deals with the selection problem from an alternative approach developed by Antle (2011), Tradeoff Analysis Model for Multi-Dimensional Impact Assessment (TOA-MD) model. As oppose to structural econometric models which rely heavily on economic theory and are data-demanding, TOA-MD is a parsimonious model based on statistical assumptions. It is parsimonious in the sense that it is mainly based on a bivariate normality assumption, and only requires summary statistics to implement. It also emphasizes on the heterogeneity of population, which could link site-specific characteristics. The importance of heterogeneity is well acknowledged in the crop insurance literature, its existence leads to series of problems caused by information asymmetry, such as adverse selection (Goodwin 1993, Glauber 2004). Heterogeneity is also important for program enrollment, which has been emphasized in econometric policy evaluation and technology adoption literature (Suri 2011).

TOA-MD model, this study applies, is a population-based approach that links policy changes to the participation and simulates distributional policy impacts counting for self-selection problem. The model is based on statistical assumption that farmers make participation decision maximizing their net returns, and the net returns of two alternatives (participate or not participate) follow the bivariate normal distribution. Based on this simple assumption, we use observed summary statistics of self-selected groups to recover parameters of the population, construct counterfactuals, and therefore obtain the impact of crop insurance, such as, average treatment effect, treatment effect on the treated group, and treated effect on the untreated group. Monte-Carlo simulation is also used to capture the difference between ex-ante and ex-post parameters.

The paper is organized as follows: Section2 discusses the theoretical foundations underlying the empirical framework and how empirical models are integrated in the context of crop insurance. Section 3 discusses data structure and summarizes crop insurance program in the study area, the Corn Belt. Section 4 lays out the detailed empirical procedures. Section 5 presents estimation results. The last section makes conclusions.

Theory and empirical framework

The theoretical foundations underlying the TOA-MD model is introduced in Antle (2011). This section focuses on its application in the context of crop insurance. The impact assessment is essentially comparing two states with and without policy intervention. The estimation strategy is to recover net return counterfactuals observed scenarios to deal with the selection issue, based on the assumption of bivariate normal distribution. For ex-ante assessment to simulate the impact with policy scenarios. Finally, TOA-MD model simulate participation rate and impact on net returns.

There are various kinds of crop insurance programs that have been implemented. Most of them takes the form of conventional insurance[1], that is, farmers pay premiums for certain level of coverage and get indemnity payment when certain criteria is triggered. Programs differ by protection type (yield or revenue), coverage levels, and indemnity criteria (farm-based or county-based). For the convenience of presentation, we use the example of area-based revenue insurance. The following conceptual framework can be applied to analysis of other types of insurance too.

Before going into details, it is helpful to lay out a few concepts to avoid confusions later. First, we approach the crop insurance mechanism from both farm net return distribution and then population net return distribution. There are ex-ante distribution and ex-post distribution. The ex-ante distribution reflects farmers' expectations. And the ex-post one is the realized, or observed distribution. Second, the farm net return includes returns from production and government programs. And the net return of crop insurance may be from two types, one is strictly from the contract, that is, the indemnity payment minus the farmers' paid premium; we call it contract return; the other is from changes of production behavior, it

[1] There are some pilot index insurance programs available in some counties, these index programs provide coverage mostly for forage, pastureland and apiculture. It only accounts for a small proportion compared with conventional crop insurance, therefore, this study only focuses on the conventional crop insurance.

is indirect return, because it is through changes in yield or cost. Third, a state refers to the case with or without insurance; the population distribution of farmers' each state is called system in TOA-MD model. In the realized world, only one state is observable, and the other is called its counterfactual. Different types and levels of policy interventions, such as different insurance programs, coverage levels or subsidy levels are called policy scenarios.

Individual level and population level distributions

The net return of individual farm is a distribution over uncertain factors such as weather conditions, natural disasters, and price fluctuations. The individual distribution is idiosyncratic in that it is affected by farm characteristics, local weather conditions, soil qualities, management practices, and etc. The net return is the revenue minus the production cost, for farmer *i* in county *j*, $NR_{ij} = R_{ij} - C_{ij}$. Assume it follows Normal distribution: $N(\mu_{ij}, \sigma_{ij}^2)$. Population distribution characterizes parameters of individual distributions over heterogeneous populations. For instance, the means of individual net return distributions are often parameters of interest[2]. They are distributed over heterogeneous individual and spatial characteristics. Again, assume it follows the normal distribution, the distribution of μ_{ij}, the mean of individual net return distribution, is $N(\mu, \sigma^2)$. In sum, the individual level distribution is over uncertainty in production; the population distribution characterizes parameters of individual distributions over population with different characteristics.

Crop insurance mechanism

When crop insurance comes into play, it truncates distributions at both individual level and population level. The following explain how it affects net returns at individual and population levels.

Individual net return with insurance is the production net return plus contract return from insurance: $NR_{ij}^I = R_{ij}^I - C_{ij}^I + In_{ij}^I - (1-s) \cdot Pr_{ij}$. The revenue and cost with insurance (R_{ij}^I *and* C_{ij}^I) can be different from the ones without insurance (R_{ij}^N *and* C_{ij}^N). The state with insurance is denoted by superscript *I*, and the state without insurance is denoted by superscript *N*. This setting allows the potential for moral hazard behaviors. Pr_j is the insurance premium which is based on the county-level loss history and *s* is the premium subsidy rate. The farmers paid premium can be simplified to $b_{ij} = (1-s) \cdot Pr_{ij}$. The indemnity payment $In_{ij}^I = \begin{cases} 0 & if\ R_{ij}^I \geq a_{ij} \\ a_{ij} - R_{ij}^I & if\ R_{ij}^I < a_{ij} \end{cases}$, where a_{ij} is the cutoff level. It triggers the indemnity which is selected by the farmer in the contract. The insurance coverage guarantees the net return not falling below the cutoff a_{ij}, that is, the individual net return distribution is truncated at the cutoff level. Figure 1 shows that the individual distributions with and without insurance, and the shaded area is net return distribution with insurance. The distribution without insurance might also be on the left side of the one with insurance.

---

[2] If without the risk neutrality assumption, the higher moments of the individual level distribution would be of interest too.

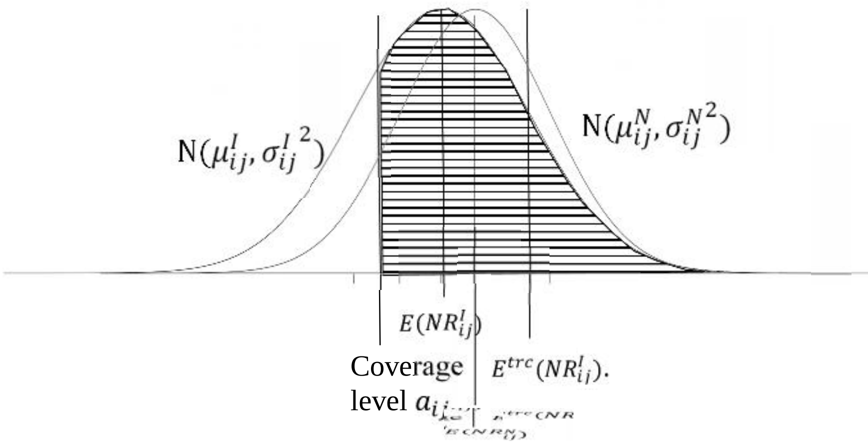

Figure 1. Individual distributions of net returns

Notice that there are a variety of insurance programs, the differences are all absorbed into the term a_{ij}.

The opportunity cost of participating insurance program for a risk-neutral farmer is $\omega_{ij} = NR_{ij}^{N} - NR_{ij}^{I}$. A risk-neutral and profit maximizing farmer will participate if expected opportunity cost is negative: $\mu_{\omega_{ij}} = E(\omega_{ij}) = E\left(R_{ij}^{N} - R_{ij}^{I}\right) - E\left(C_{ij}^{N} - C_{ij}^{I}\right) - E(In_{ij}^{I}) + b_j$. If participation doesn't change farmer's production behavior, then the first two terms vanish, and participation decision only depends on the expected indemnity payment and subsidized premium payment.

TOA-MD model links the population distribution of outcomes to the participation process. At the population level, the parameter of interest is the mean of individual distribution (μ_{ij}^{N}, μ_{ij}^{I} and $\mu_{\omega_{ij}}$). To be consistent with the terminology of TOA-MD model, following Antle (2011), call the population distribution of μ_{ij}^{N} system1, and the population distribution of μ_{ij}^{I} system2. They can be described as a bivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$. Distribution of $\mu_{\omega_{ij}}$ determines the participation rate, which can be written by the parameters of the bivariate normal distribution. Figure 2 is adapted from Antle (2011), illustrating population distributions (system1 and system2), as well as their relationships with opportunity cost and how the participation truncates population distributions. Each ellipsoid shows the bivariate distribution of net return (means of individual distributions) with opportunity cost. The population participation rate is the proportion of population that have negative opportunity cost, as shown in Figure 2, participants are dotted part of system2 ellipsoid, denoted as B, and the rest will stay out of insurance programs (the checked part of system1 ellipsoid denoted as A).

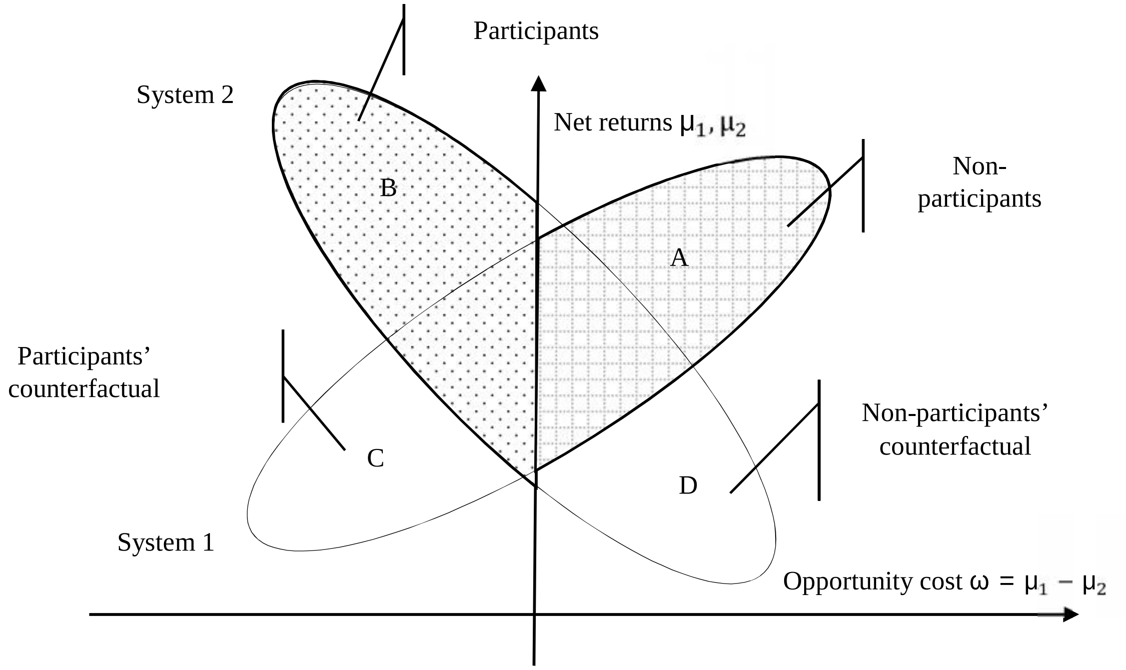

Figure2. The conceptual model of program participation and impact assessment

Recover counterfactuals and treatment effects

In the impact assessment, one goal is to characterize distributions of treatment effects by policy interventions. And average treatment effects (ATE), average treatment effect on the treated (ATT), treatment effect on the untreated (ATU) are often parameters of interest. The population average of negative opportunity cost is the ATE; the average of negative opportunity cost across participants is ATT; and the average of negative opportunity cost across non-participants is ATU. They can be obtained by comparing net returns in two states. However, only one state is observable (farmers are either in crop insurance program or not). To obtain the distribution of opportunity cost, recovering net returns of counterfactual states for both participants and non-participants is necessary. The traditional approach is to estimate policy-invariant econometric models and then simulate counterfactuals under different policy scenarios. In this study, following the parsimonious spirit of TOA-MD model and the bivariate normal distribution assumption, it is convenient to recover counterfactual distribution parameters from the observed truncated ones.

The technique to recover the counterfactuals is called ex-post solver here. Means and variances of observed distributions, i.e., region A and B (μ_{1o}, σ_{1o}^2 *and* μ_{2o}, σ_{2o}^2) are parameters of truncated distributions. They can be written as equations by parameters of bivariate normal distribution, that is the bivariate distribution of two systems with and without insurance (equation (2)-(5)), and the realized participation rate is cumulative distribution function of normal distribution valued at 0 (equation (1)). There are five known parameters ($r, \mu_{1o}, \sigma_{1o}^2, \mu_{2o}$ *and* σ_{2o}^2) and five unknown parameters (correlation coefficient of the bivariate distribution ρ, means and variances of bivariate normal distributions μ_1, σ_1^2, μ_2 *and* σ_2^2) in the system of equations. The parameters of bivariate normal distribution can be solved from equation (1)-(5).

$$r = \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}}\right) \qquad (1)$$

$$\mu_{ho} = \mu_h - (-1)^h \sigma_{ho}\theta_h\lambda_h \qquad (2), (3)$$

$$\sigma_{ho}^2 = \sigma_h^2\{1 + (-1)^h(\mu_1 - \mu_2)\theta_h^2 - \theta_h^2\lambda_h^2\} \qquad (4), (5)$$

Where h=1 and 2 for system1 without insurance and system2 with insurance, and λ_h is the inverse Mills ratio and θ_h is the correlation between ω and μ_h

Another related issue in parameterization is to obtain ex-ante parameters, because farmers make decision based on expected net returns with information set only known at the beginning of the growing season. Since the assessment is based on participation decisions, the true parameters should be ex-ante. Monte Carlo simulation is used to simulate ex-ante parameters from ex-post ones. The logic is as follows: the observed population distribution of outcome is one set of realization from ex-ante distribution; it can be thought of as a set of samples drawn from all possible outcome in the ex-ante distribution. By repeatedly drawing samples, we could simulate the distribution over all possible outcomes, i.e. the ex-ante distribution. In this sense, the observed means and variances are sample means and variances, they also following normal distributions. The standard error for the sample means and variances are: $SE(\mu_{ho}) = \sqrt{\sigma_{ho}^2/n_h}$ and $SE(\sigma_{ho}^2) = \sigma_{ho}^2\sqrt{2/n_h - 1}$. The procedure goes, first randomly drawing a set of parameters from distributions of sample means and variances, and second, form a system of equations (1)-(5), and solve for parameters of bivariate normal distribution ($\mu_1, \sigma_1^2,\ \mu_2, \sigma_2^2$ and ρ); and then the average of solved parameters are the ex-ante parameters of bivariate normal distribution, the counterfactual means and variances for participants and non-participants can also be calculated. Other ex-ante parameters for ATE, ATT and ATU can all be calculated.

In practice, the system of equations are solved numerically, not every combination of random sampling can lead to a solution due to the high non-linearity and complexity of the equation system. The solvable solutions are the ones that make more sense in reality. We use other criteria based on experience to confine the range of the solutions which simulates the ex-ante distribution of systems. In sum, ex-post solver serves for two purposes. The first is to construct parameters of bivariate normal distribution from parameters of truncated distributions, in order to solve the counterfactual issue. The second is to convert ex-post parameters into ex-ante parameters by Monte-Carlo simulation.

Data and crop insurance in the Corn Belt

Data

There are mainly three data sources for empirical estimation: 2007 Agricultural Census, and crop insurance administrative data from Risk Management Agency's website: (a) Statistics (mean, variance, coefficient of variation and correlation) of the following variables are extracted from 2007 Agricultural Census for TOA-MD model: (1) revenue and acreage of five crop activities, namely, corn, soybean, wheat, hay and the other crops as a numeraire; (2) farm characteristics: total land acreage, total expenditure, total sales, and livestock sales; (3) policy variables: acreage under crop insurance, crop and livestock insurance payment; direct payment and other subsidy payments. (b) The summary of business data from Risk Management Agency (RMA)'s website are used to simulate net returns from crop insurance.

Crop insurance in the Corn Belt

This section describes summary statistics of farm characteristics to provide background information about the Corn Belt region. The Corn Belt is one of the major production regions in the U.S., including Illinois, Iowa, Indiana, Missouri and Ohio. Corn and soybean are the major crops growing in the region, accounting for over 50% of national corn and soybean production; hay and small grains such as oats and winter wheat and dairying are other important agricultural activities. We mainly look at corn, soybean and wheat in the region. In order to better fit the bivariate normal distribution assumption. The population is first grouped by farm types. And each farm type group is further stratified into two stratums by median farm size. Stratums are treated as independent groups with each other. The following table lists some summary statistics of each stratum.

Table 1. Farm type stratification

Stratums	Number of farms	Farm sizes	Insurance Participation rate
1.Corn-soy farms (L)	35101	849	77%
2.Corn-soy farms (S)	36624	142	51%
3.Non-irrigated other farms (L)	24196	204	6%
4.Non-irrigated other farms (S)	26504	72	3%
5.Irrigated Corn-soy farms (L)	833	1694	77%
6.Irrigated Corn-soy farms (S)	1038	434	61%
7.Beef cattle farms (L)	21950	808	55%
8.Beef cattle farms (S)	23230	177	20%
9.Dairy cattle farms (L)	1216	306	36%
10.Dairy cattle farms (S)	1102	90	6%
11.Hog farms (L)	2536	1088	86%
12.Hog farms (S)	2383	273	69%

Note: L for large farms and S for small farms, they are grouped by the median farm size in the farm type group.

Insurance contract returns

The following table shows the average premium, subsidies, indemnity payment and net returns from crop insurance compiled from RMA crop insurance report of business.

Table 3. 10-year average of insurance contract returns (1998-2007, Unit: $/acre)

	Premium	Premium subsidy	Liability	Indemnity	Net return
All crops	15.6	8.2	218.8	7.0	-0.4
Corn (41)	19.6	10.2	266.4	8.0	-1.4
Soybean (81)	11.4	6.1	172.4	5.9	0.6
Wheat (11)	8.8	5.4	93.6	5.7	2.4

Parameterization

The goal of this study is to estimate distributions of treatment effects on net returns by comparing net return distributions with and without insurance treatments. This section describes parameterization of net return distributions for TOA-MD model. The estimation strategy is first to obtain statistics of production net returns and insurance net returns for participants and non-participants, and then recover ex-ante parameters for population distributions, which are the true parameters for TOA-MD model.

The stratums are treated as independent with each other, and the parameterization procedures are the same. Therefore, lose subscripts for stratums for notation convenience. Based on the notation used above, i indicates for individual; j is the county where individual i locates; h indicates for the two systems $h = 1,2$; k indicates for outcome variables.

<u>Net returns per acre</u>

Statistics of farm-level net return distribution can be obtained from the agricultural census. In the risk-neutral case, the mean of individual net return distribution is the parameter of interests, which can represented by realized net returns in cross-sectional data. Production net return per acre is revenue per acre minus the cost per acre: $ANR_{ij} = \frac{R_{ij}-C_{ij}}{A_{ij}}$. The revenue includes government payment and insurance payment. Group means and variances are calculates.

<u>Truncated population distributions of net returns</u>

Realized individual net returns from production and indemnity payment are used to construct parameters for truncated population distributions of net returns. The mean and variance of net returns for non-participants (d=0) is the truncated distribution of system1 (Region A). These parameters can be directly collected from the census data.

$$\mu_{1o} = E[ANR_{ij}|d = 0]$$

$$\sigma^2{}_{1o} = E\left[\left(ANR_{ij}\right)^2\middle|d = 0\right] - (\mu_{1o})^2$$

The mean and variance of net returns for participants (d=1) is the truncated distribution of system2 (Region B), it includes net returns from production and insurance. The indemnity payment is included in ANR_{ij}, Premium rate setting primarily uses historical loss experiences for a crop in a county (Coble et al. 2010) and subsidy rates are fixed countrywide. Therefore, county-level average of farmers' paid premium In_j^I is used, and they are obtained from business of summary data set at Risk Management Agency (RMA) website.

$$\mu_{2o} = E[ANR_{ij} - In_j^I | d = 1]$$

$$\sigma^2{}_{2o} = E\left[\left(ANR_{ij} - In_j^I - \mu_{2o}\right)^2 \middle| d = 1\right]$$

Parameters for correlated truncated distributions $\mu_{1o}, \sigma^2{}_{1o}$ and $\mu_{2o}, \sigma^2{}_{2o}$, as well as the participation rate, are used to recover parameters of bivariate normal distributions $(\mu_1, \sigma^2{}_1, \mu_2, \sigma^2{}_2 \; and \; \rho)$. Ex-ante parameters are simulated by Monte Carlo simulation.

Empirical Results

The results focused on major corn-soybean farms. The following presents estimation results of net return distributions and impact assessment.

We recover the bivariate normal distribution of net returns with and without net returns using MODEL procedure in SAS program. It essentially computes a simultaneous solution of a system of non-linear equations. Due to high non-linearity and complexity of the equation system, sometimes the numerical method cannot reach a solution. We relax one restriction in the system, the one about the correlation coefficient of the two systems. It is assumed to be known and follow a uniform distribution from -1 and 1. By relaxing this restriction, the efficiency of program increases significantly. And among the solutions, we can further set restrictions to narrow down the range of solutions.

Ex-ante and ex-post distributions

In the following, we mainly discuss results from corn-soybean farm groups (stratum1 and 2), which is a major farm type in the Corn Belt region. The following table shows both ex-ante and ex-post parameters of the bivariate normal distribution for corn-soybean farms. Column 3, 5, 7, and 9 show parameters for the truncated distributions for participants and non-participants. Colum 2, 4, 6, and 8 show parameters of the population distribution recovered from the truncated distributions. Ex-ante parameters come from Monte Carlo simulation that varying observed parameters, and ex-post parameters are recovered from exact observed parameters.

Comparing truncated with population distributions, the truncated distribution means are greater than the population means no matter for system 1 or system2. It makes sense because farms opt themselves into the system that will generate greater returns. For stratum 1 (larger farms), the standard deviations are similar between truncated and population distribution for both system1 and system2.

Between ex-post and ex-ante distributions, in stratum1, corresponding parameters for truncated variables are close. Probably due to the fact that the corn-soybean farm is one of the major farm types, and relatively homogeneous in the Corn Belt region. Also the Corn Belt is a relatively low-risk production region. Both participants and non-participants have more experiences and also form better expectations of production outcomes in their own systems. For population distributions, expectations for system2 is more accurate than the system1. It is because the adoption rate is high (77%) in the region, farms are more familiar with the system with insurance.

In stratums2, we do observe bigger difference between ex-post and ex-ante parameters, especially for the means, that is, bigger difference between expected net returns and realized net returns. First, small farms might be less stable than large farms. For instance, if there are some disturbances, the impact might be even out eventually due to variations within the farm. It is also possible that extreme disturbances skew outcome distribution into one direction even more for large farms, depending on the situation. Second, greater biases between expected values and realized values come from the fact that there are bigger variations and uncertainties among small farms. Notice that the population means are over expected for both systems, and the expected and realized means are fairly close. It means farms tend to over-estimate the counterfactual net returns, it also shows that small farms are less predictable.

Table 4: Bivariate normal distributions of net returns ($/acre)

Parameters	S1.Large corn-soybean farms				S2.Small corn-soybean farms			
	Ex-post		***Ex-ante***		***Ex-post***		***Ex-ante***	
	Population	Truncated	Population	Truncated	Population	Truncated	Population	Truncated
System 1	154.6	167.0	147.8	164.7	107.8	102.6	108.4	103.3
(No insurance)	(132.8)	(132.4)	(133.0)	(132.4)	(156.0)	(155.9)	(156.0)	(155.9)
System 2	162.4	166.2	162.1	166.7	110.9	136.4	110.5	135.7
(Insurance)	(132.8)	(132.7)	(132.9)	(132.7)	(159.3)	(156.9)	(159.2)	(157.0)

Note: Numbers in parentheses are standard deviations.

Treatment effects

Table 5. Treatment effects of crop insurance and correlations between systems

	Large corn-soy farms		Small corn-soy farms	
	ex-post	ex-ante	ex-post	ex-ante
Participation rate (percentage of farms)	77.7	77.7	51.4	51.4
ATE ($/acre)	7.8	14.4	3.1	2.1
ATU ($/acre)	-16.9	-17.7	-18.5	-19.4
ATT ($/acre)	15.3	24.1	23.6	22.5
corr (μ_1, μ_2)	1.0	1.0	1.0	1.0
corr (ω, μ_1)	0.1	0.1	0.0	0.0
corr(ω, μ_2)	-0.1	-0.1	-0.2	-0.2

From the net return bivariate normal distribution, we could calculate average treatment effect (ATE), average treatment effect on the treated group (ATT) and average treatment effect on the untreated group (ATU). ATE is positive for both stratums, and it is higher for larger farms resulting in higher participation in stratum1. As expected, ATU is negative and TT is positive, because they think they will earn more in the system they choose, and lose in the other. Comparing ATU and ATT between larger farms and small farms, it seems that participants with smaller sizes benefit more from insurance than larger participants on the per-acre basis; at the same time smaller non-participants will lose more if they participated. It might come from the fact that characteristics of small farms have more variations, and therefore impacts by insurance are more likely to spread out. And also if farmers think they could benefit by changing production behavior, smaller farms are more flexible in this aspect. Notice that the ex-ante treatment effects are greater than the ex-post parameters in magnitudes, showing farmers tend to overestimate the benefit or loss from insurance.

Notice the treatment effect is much greater than the region's average gains from insurance contract (indemnity payment minus farmers' paid premium). It suggests the impact on net return might come from other sources too, such as, changing input uses, converting marginal land to production and etc., which are considered as moral hazard behavior. The treatment effect is a composite term that includes impacts on net return from insurance contract and associated production adjustment, though it cannot be decomposed to different sources. Assume it also follows a normal distribution in this case. Farmers who have positive treatment effects are participants, and some who don't are non-participants. For nonparticipants, maybe it is highly costly for them to convert marginal land into production (constraints of scale, labors); because of locating in a high risk area, the premium is high.

Though it is hard to decompose the treatment effects into sources by mechanism under the framework of this paper, we could make comparisons with evidences from other studies. First, consider the effect by converting marginal land into production. The land under the Conservation Reserve Program (CRP) can be convert into production after their contracts expire. The CRP rent payment is from $80 to $130 per acre on average which can be considered as proxy for production net returns. The average increase in net return per acre if CRP land is released to production are $2 to $6 on average. There is higher increase for smaller stratums, since the percentage of CRP acreage is higher for smaller farms. Second, the treatment effect may come from land use changes among crops, as studies find that crop insurance participation provokes acreage responses (Goodwin, Vandeveer and Deal 2004), such as shift from hay and pasture to corn (Wu 1999). Third, in the ex-ante context, it is reasonable to assume the insurance is actuarially fair since it is legislated to be actuarially fair. An expected benefit would be premium subsidies, which ranges from $4 to $18 per acre at county level. Fourth, as for input uses, studies found different evidence: Smith and Goodwin (1996) find that on average, insured farm spend $4.23 less on fertilizer than uninsured farms; Horowitz and Lichtenberg's (1993) results show that insured farmers tend to use more chemicals. Thought these impacts by sources discussed above cannot be added up directly to be compared with the treatment estimated here. The fact that these numbers fall into the range of the treatment effects indicates the estimation of treatment effects are reasonable and comparable with other studies.

In this sense, the treatment effect incorporates impacts from insurance contract and a package of behavioral adjustments, and it follows a distribution, and ATT is positive and ATU is negative. However, in reality, the lower bound of treatment on the untreated group should be farmers' paid premiums. Because farmers could avoid any behavioral adjustment under crop insurance that might cause losses. That is actual ATU might be smaller than estimated in Table 5. Overall the treatment effect can be interpreted as a compound effect from a suite of sources, not just insurance contract. However, from another perspective, the treatment effects are calculated backwards from participation rate. The gap between treatment effect and contract return may indicate that farmers insure to reduce risks. It is worthwhile to explore the risk-averse setting.

For an impact assessment, one question is how policy interventions alter the relative ranking and equality of the economy. The correlation relationships of bivariate normal distribution could provide useful information in this aspect. The correlation of system 1 and opportunity are positive and correlation of system 2 and opportunity cost are negative, though they are very small in magnitudes. It suggests farms with lower net return in system1 have lower opportunity cost, farmers with lower opportunity cost have higher return in system 2. The sorting process is then farmers with lower income in system1 tend to participate and earn more in system2, which makes the economy more equal. The magnitudes of correlations between net returns and opportunity cost are small in these stratums, which makes the outcomes of the two systems are very similar (corr (μ_1, μ_1) in both systems are close to 1).

Other farm stratums

We stratify the population by their main agricultural activity and the sizes. Each stratum is treated as an independent distribution, and there are big variations across stratums. Generally, small farm groups have more variations, (high standard deviation). Therefore, smaller average impact for the population on average, but bigger impacts on the treated group and the untreated group, respectively. Larger farm groups usually gain more from the insurance.

Table 6. Parameters of net return distributions by stratums

Parameters	Corn-soy farms (L)	Corn-soy farms (S)	Irrigated Corn-soy farms (L)	Irrigated Corn-soy farms (S)	Beef cattle farms (L)	Beef cattle farms (S)	Dairy cattle farms (L)	Dairy cattle farms (S)	Hog farms (L)	Hog farms (S)
μ_1	147.8	108.4	-22.3	16.5	-14.4	-41.5	195.3	305.3	4.6	8.7
σ_1	133.0	156.0	142.9	200.4	171.9	236.8	252.3	302.6	109.6	344.6
μ_2	162.1	110.5	102.6	83.8	47.2	-231.6	111.3	206.8	136.2	182.6
σ_2	132.9	159.2	150.3	177.8	193.1	234.8	264.6	309.3	224.4	301.3
Participation rate	0.8	0.5	0.8	0.6	0.5	0.2	0.4	0.1	0.9	0.7
Corr (ω, μ_1)	0.1	0.0	0.9	0.7	0.7	0.9	0.4	-0.2	0.9	0.6
Corr(ω, μ_2)	-0.1	-0.2	-0.8	-0.5	-0.7	-0.9	-0.5	-0.3	-1.0	-0.4
Corr (μ_1, μ_2)	1.0	1.0	-0.6	0.3	0.0	-0.6	0.6	1.0	-0.7	0.5
ATE	14.4	2.1	124.8	67.3	61.5	-190.0	-84.0	-98.5	131.6	173.9
σ_{ATE}	22.8	26.3	253.9	224.4	255.8	423.4	229.1	31.4	309.2	309.5
ATU	-17.7	-19.4	-266.8	-153.1	-156.8	-381.8	-216.9	-98.9	-673.4	-166.2
ATT	24.1	22.5	240.6	207.2	243.1	559.7	155.4	-93.5	262.6	326.7

As for other farm types, the non-irrigated other farms have extremely low participation rate, 6% and 3% respectively. The program cannot get solutions because of the low participation rates. There is not enough information to recover system 2, and numerically it might run into corner solution issues as well. It also suggests that the farms' expectations are consistent that they might lose by purchasing the insurance. For the irrigated corn-soy farms treatment effects on average are higher than non-irrigated farms, but also more dispersed for participants and non-participants. This result seems to be counterintuitive. Because irrigation practice reduces the risks and constraints by weather conditions; it is a substitute to crop insurance. However, we see higher participation rates in irrigated groups and also higher treatment effects. One possible explanation might be that the Corn Belt has advantageous agriculture production conditions, and the constraints relieved by irrigation won't play a very big role in production. At least it is the case in year 2007, because the average corn yield are very close for irrigated groups and non-irrigation groups; and the soybean yields are higher in non-irrigated groups. The irrigated farms are twice or three times the sizes of non-irrigated farms, the irrigation practice is applied maybe the farm are located in relatively drier regions or they are preventative practice, the role of risk reducing it plays is very limited in normal years. Other livestock stratums have greater variations within stratum and also the distributions might be skewed by livestock activities, and the normality assumption might don't fit very well. The estimation results are less reliable.

TOA-MD impact assessment

Table 7. Impact assessment in Corn Belt region

	Major corn-soybean farms (L)	Major corn-soybean farms (S)	Population
Participation rate (% of farms)	82	71	81
Mean net return per acre ($)	185	134	177
Mean farm net return($)	144,829	17,378	79,411
Total net gains ($)	1,309,100,973	130,674,105	1,439,775,078

TOA-MD model assess the total impact in the region for different stratums. The net return distributions estimated from ex-post solver and farm characteristics collected from Agricultural census data are input for TOA-MD model. And impacts estimated for major corn-soybean farms are shown in Table 7. Major corn-soybean farms account for 65% of the insured acreage in the region. The model tends to over-predict the participation rate. It makes sense that we assume acreages farm are homogeneous within a farm, therefore the acreage coverage of insurance is assumed to be 100% within a farm. It is reasonable to some extent for major corn-soybean farms, since the acreage coverage is 85% to 90% in these two stratums. But this number varies a lot in other stratums. Further The total net gains in the region is $1.4 billion, which accounts for 11% of total farm sales. And 91% of benefits go to large farms, even though the participants in small farm group gains more. The high participation rate and the size of insured acreage gross the most of benefit for large farms. This result provides quantitative support to the argument against crop insurance that the program is in favor of large farms. This is inherent problem with the program, in that larger farms are more stable and homogeneous, even if the benefit per unit is small, the gross benefit for the farm is sizable, and therefore they are more likely to participate in the program. Overall, they get most of the benefit from the program. It is consistent with results from other studies that the level of heterogeneity of farm characteristics affect program participation (Goodwin 1993, Glauber 2004).

Conclusions

The study focuses on farmers' participation decision to crop insurance program and then the impact of participation. The approach that solves the selection problem in impact assessment is parsimonious in the sense of data collection and simulation. The statistics of different farm groups are collected from Agricultural Census for the model. Based on the bivariate normal distribution assumption, counterfactual distributions of net returns with and without insurance can be recovered from observed statistics, and then Monte Carlo simulation is used to obtain ex-ante parameters of net return distributions from ex-post parameters. The recovered counterfactual distributions and results from the land use model are input for TOA-MD model.

Recovering the distribution of net returns for both systems with and without crop insurance shows their correlation relationships and treatment effects. The correlation relationships is the key to selection process. In general, the self-selection process improves the equality. The treatment effect is quite different from average insurance contract returns; it suggests farms participate the insurance program not only for premium subsidies or the indemnity payment. They might be able to benefit more from adjusting production behaviors. The bivariate distributions of net returns vary by farm types and sizes. Generally, smaller farms have higher variations in net returns, and opportunity cost to participate, and their participation rate is lower. The impact assessment by TOA-MD model supports the argument that the crop insurance program benefits large farms more. This is inherent problem with the program, in that larger farms are more stable and homogeneous, even if the benefit per unit is small, the gross benefit for the farm is sizable, and therefore they are more likely to participate in the program. Overall, they get most of the benefit from the program.

Future expansions can be made to other regions to test the validity of the model. In the Monte Carlo simulation, the variation of the parameters can be obtained from the panel data. The current setting of the model is for risk neutral farms, it is worthwhile to explore the setting of risk-aversion.

References

Antle, John M. “Parsimonious Multi-Dimensional Impact Assessment.” *American Journal of Agricultural Economics* 93, no. 5 (2011): 1292–1311.

Chite R.M. “The 2014 Farm Bill (P.L. 113-79): Summary and Side-by-Side.” *Congressional Research Service*. 2014.

Glauber, Joseph W. “The Growth of the Federal Crop Insurance Program, 1990–2011.” *American Journal of Agricultural Economics* 95, no. 2 (2013): 482–88.

Goodwin, Barry K., Monte L. Vandeveer, and John L. Deal. “An Empirical Analysis of Acreage Effects of Participation in the Federal Crop Insurance Program.” *American Journal of Agricultural Economics* 86, no. 4 (2004): 1058–77.

Horowitz, John K., and Erik Lichtenberg. “Insurance, Moral Hazard, and Chemical Use in Agriculture.” *American Journal of Agricultural Economics* 75, no. 4 (1993): 926–35.

Schoengold, K, Y. Ding and R. Headlee. “The Impact of AD HOC Disaster Crop Insurance Programs on the Use of Risk-reducing Conservation Tillage Practices.” *American Journal of Agricultural Economics* (2014): aau073v1-aau073.

Serra, T., B.K. Goodwin and A.M. Featherstone “Modeling Changes in the US Demand for Crop Insurance during the 1990s.” *Agricultural Finance Review* 63 no.2 (2003), 109-125

Shaik, Saleem, Keith H. Coble, and Thomas O. Knight. “Revenue Crop Insurance Demand.” In *Selected Paper Presented at AAEA Annual Meetings, Providence, Rhode Island*, 7:24–27, 2005. http://www.researchgate.net/publication/23506059_Revenue_Crop_Insurance_Demand/file/9fcfd50ace70926f3d.pdf.

Shields, Dennis A. “Federal Crop Insurance: Background.” *In Congressional Research Service. CRS Report for Congress,* December. Vol. 6, 2012. http://nationalaglawcenter.org/wp-content/uploads/assets/crs/R40532.pdf.

Smith, Vincent H., and Barry K. Goodwin. “Crop Insurance, Moral Hazard, and Agricultural Chemical Use.” *American Journal of Agricultural Economics* 78, no. 2 (1996): 428–38.

Suri T. “Selection and Comparative Advantage in Technology Adoption.” *Econometrica* 79 no.1 (2011), 159-209

Walters, Cory G., C. Richard Shumway, Hayley H. Chouinard, and Philip R. Wandschneider. “Crop Insurance, Land Allocation, and the Environment.” *Journal of Agricultural & Resource Economics* 37, no. 2 (2012).

Wu, JunJie. “Crop Insurance, Acreage Decisions, and Nonpoint-Source Pollution.” *American Journal of Agricultural Economics* 81, no. 2 (1999): 305–20.