Cost-share Effectiveness in the Adoption of Cover Crops in Iowa

María Jimena González Ramírez*[1,2], Catherine L. Kling[1,2] and J. Gordon Arbuckle, Jr.[3]

[1]Center for Agricultural and Rural Development, Iowa State University
[2]Department of Economics, Iowa State University
[3]Department of Sociology, Iowa State University

May 27, 2015

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.[4]

Abstract

Motivated by the importance of the adoption of cover crops in Iowa and the availability of cost-share funding for this practice in 2013, we use matching methods combined with regression analysis to study the effectiveness of cost-share funding in the adoption of cover crops in Iowa in 2013 using the Iowa Farm and Rural Life Poll. After matching treatment and control groups, we use a Tobit model to estimate the effect of cost-share funding on cover crops acres and on the proportion of cover crops planted relative to total farm land. We find that receiving cost-share funding has a positive and statistically significant effect on cover crops acres and on the proportion of cover crops planted. Furthermore, we find that the mean marginal effect on the proportion of cover crops acres is around 15% among adopters and non-adopters and 14% among adopters. Consequently, having cost-share funding increases the proportion of cover crops acres planted by 15% among all farmers and by 14% among adopters on average. Furthermore, we find that cost-share funding increases cover crops acres on average, but the size of its marginal effect varies with different matched data sets.

JEL Codes: Q15

*This research is part of a regional collaborative project supported by the USDA-NIFA, award number 2011-68002-30190: "Cropping Systems Coordinated Agricultural Project (CAP): Climate Change, Mitigation, and Adaptation in Corn-Based Cropping Systems."

1 Introduction and Literature Review

Despite the creation of the Hypoxia Task Force in 1997 (EPA Task Force), the 2014 Gulf Hypoxia zone of oxygen-depleted bottom-water was roughly 13,000 square kilometers, an area much higher than the Hypoxia Task Force goal of 5,000 square kilometers (EPA 2014). In 2013, Iowa developed a statewide Nutrient Reduction Strategy, which is a science and technology-based framework to assess and reduce nutrients to Iowa water and the Gulf of Mexico (Iowa NRS 2013). The strategy calls for a significant adoption of cover crops, crops that are planted between harvest and the planting of cash crops, which are able to reduce both nitrogen and phosphorus losses by approximately 30%. In August of 2013, $2.8 million became available statewide to implement conservation practices based on the Iowa NRS through the Water Quality Initiative (WQI) (Iowa NRS 2014). The funds were allocated for practices that could be implemented in a short time, with the goal of providing water quality benefits in 2013 and spring of 2014 (Iowa NRS 2014). One of the practices that was promoted through this cost-share program was cover crops. According to the Iowa NRS 2013-2014 Annual Progress Report, roughly 95,000 acres of cover crops were established through this state cost-share program. This number is very small relative to the amount of total farmland in Iowa, which is around 30.5 million acres (USDA NASS 2014). Given the importance of cover crops for water quality, we are interested in assessing the effectiveness of cost-share programs in the adoption of cover crops in Iowa. We use matching methods combined with regression analysis to study the effect of cost-share programs on the adoption of cover crops in Iowa in 2013 using the Iowa Farm and Rural Life Poll.

In order to assess the effect of cost-share funding in the adoption of cover crops, we would like to know the cover crop planting decision of farmers who received cost-share funding in the absence of the funds. Nonetheless, we can never observe the counterfactual (Imbens & Wooldridge 2008). Furthermore, since the participation in cost-share programs for cover crops is not random, we also face a selection problem that can come from both observable and unobservable factors. For instance, a farmer who participates in a cover crop cost-share program may have invested more in conservation practices in the past than a farmer who does not participate. Consequently, directly comparing the adoption of cover crops between farmers enrolled in cost-share programs and farmers who are not enrolled could result in incorrect estimates that suffer from selection bias. We use matching methods to pair treated and untreated (control) farmers based on observable characteristics measured before treatment to overcome the selection problem and to have a valid counterfactual. After the data is matched, we employ regression analysis to estimate the treatment effect of cost-share funding on the adoption of cover crops, the number of cover crops acres planted, and the proportion of cover crops acres relative to total farm land.

Matching methods have been employed for program evaluations related to conservation. Liu and Lynch (2011), use matching methods to study the effect of land-use policies focused on the reduction of farmland loss. Ferraro et al. (2007), study the effectiveness of the U.S. Endangered Species Act on species recovery rates using matching methods. Adam et al. (2008), estimate the effectiveness of protected area networks on deforestation rates in Costa Rica. Conservation Programs have also been studied using difference-in-difference matching. Chabé-Ferret & Subervie (2013) study European Union Agro-environmental schemes (AES) implementation in France. AES pay farmers to adopt greener practice. The AESs they study include schemes that are meant to increase crop diversity, the planting of cover crops, the planting of buffer strips, and the conversion to organic farming. They estimate additional and windfall effects of each AES they study. Additional effects occur if the AES encourages farmers to adopt environmentally friendly practices. The windfall effect occurs if the AES program pays for practices that would have been adopted in the absence of the program. Using propensity score matching and difference-in-difference, they estimate the average treatment effect on the treated and use it to calculate these effects for each practice (Chabé-Ferret & Subervie 2013). They find that the AES increases the area planted in cover crops and that the average treatment effect on the treated is around 10 ha.

Mezzatesta et al. (2013) also estimate the average treatment effect of cost-share programs and also address additionality concerns from conservation practices. They use matching techniques to estimate the average treatment effect on the treated of cost-share funding for several conservation practices, including cover crops, in Ohio. Their outcome variable is the proportion of acres under a particular conservation practice relative to total farm acres. Furthermore, they decompose the average treatment effect on the treated according to relative contributions of adopters and non-adopters. They define adopters as enrolled farmers who would have adopted the practice in the absence of the cost-sharing fund. Non-adopters are defined as enrolled farmers who would not have adopted the practice without funding. They find that the average treatment effect on the treated of enrollment in cost-share programs is roughly 23% for cover crops.

We focus on two outcome variables studied before: (1) cover crops acres planted (Chabé-Ferret & Subervie 2013), and (2) proportion of cover crops planted relative to total farm land (Mezzatesta et al. 2013). However, instead of calculating the ATT, we estimate the treatment effect using Tobit regressions on the matched data. Whether we focus on acres or the proportion, our outcome variables are censored at zero due to the lack of adoption of this conservation practice. Hence, using Tobit regressions seems appropriate to calculate the marginal effect of receiving cost-share. We find that receiving cost-share has a positive effect on both outcome variables and its coefficient under each regression is statistically significant at the 99% level for both outcome variables. We find that the mean marginal effect on the proportion of cover crops acres is roughly 15% among both adopters and non-adopters and around 14% among adopters using three different matching methods with good balance. However, the mean marginal effect on cover crops acres differs among matched data sets, ranging between 106 and 123 acres among adopters and non-adopters, and ranging from 102 to 116 acres among adopters only. For the adoption of cover crops, we use a Logistic regression and find that the cost-share treatment indicator is not statistically significant. In the end, we conclude that cost-share funding increases cover crops acres and the proportion of cover crops acres planted relative to total farm land.

2 Methodology

For the estimation of treatment effects, we would like to know the way the treatment participant would behave in the absence of the treatment as first formalized by Rubin (1974). The treatment effect for individual i is the comparison of i's outcome with treatment, denoted by $Y_{1,i}$, and i's outcome without treatment, denoted by $Y_{0,i}$. The fundamental problem when estimating treatment effects is that we only observe one of these potential outcomes for each individual (Holland, 1986). Basically, when estimating causal effects, we face a missing data problem, where we want to predict the unobserved potential outcomes (Rubin, 1976). In order to estimate treatment effects, $E(Y_1 - Y_0|X)$, we would like to compare treated and control individuals that are very similar. Following Rosenbaum and Rubin (1983) and Heckman et al. (1998), two assumptions are made to estimate treatment effects: (1) strong ignorability assumption, in which the treatment assignment, denoted by T, is independent of potential outcomes (Y_0, Y_1) given the covariates X (i.e. $T \perp (Y_0, Y_1)|X$); and (2) overlap assumption, in which there is a positive probability, denoted by $P(T = 1)$ of receiving each treatment for all values of X (i.e. $0 < P(T = 1|X) < 1$ for all X). A weaker version of (1), in which $E(Y_0|X, T) = E(Y_0|X)$ and $E(Y_1|X, T) = E(Y_1|X)$, suffices for estimating the average treatment effect on the treated, defined as $ATT = E(Y_1 - Y_0|X, T = 1)$. For our research question, we focus on three outcome variables: Y^1 which is the amount of cover crops planted in 2013; Y^2 which is the proportion of cover crops planted relative to total farm land; and Y^3 which is an adoption indicator variable that takes the value of one if a farmer planted cover crops in 2013 and zero otherwise. The treatment indicator, T, is defined as follows:

$$T = \begin{cases} 1 & \text{if farmer is enrolled in cost-share program for cover crops} \\ 0 & \text{if farmer is not enrolled in cost-share program for cover crops} \end{cases}$$

In order to estimate treatment effects, the literature suggests a two-step process. To start, researchers use pretreatment information to select comparable treated and control units to analyze the treatment effect without using the outcome variable. Secondly, using the matched sample, researchers estimate treatment effects (Stuart & Rubin 2008). For the first step, matching techniques are employed to balance the distribution of covariates in the treated and control groups (Stuart 2010). In essence, by controlling for pretreatment differences between treatment and control, researchers are able to reduce bias by using a valid counterfactual. For the second step, some researchers calculate the ATT. Another possibility is to use the matched data to perform regression analysis by regressing the outcome variable on the treatment status and other relevant covariates. Matching methods and regression adjustment models can complement each other (Rubin & Thomas 2000, Glazerman, Levy & Myers 2003, Abadie & Imbens 2006). Intuitively, by selecting matched samples, the bias due to covariance differences is reduced and regression analysis for remaining small covariance differences increases the efficiency of treatment estimates (Stuart & Rubin 2008) and makes results less sensitive to model specifications (Ho et. al. 2007).

For the first step, propensity score matching (PSM) is typically employed in non-experimental studies to attain balance and overcome the selection problem (Rosenbaum & Rubin 1983). First, a propensity score is calculated, which is each individual's probability of being included in the treatment, and it is calculated using observed covariates (Wooldridge 2010). Smith and Todd (2005) recommend the inclusion of covariates that influence both treatment status and outcome when estimating the propensity score. As emphasized by Ho et. al. (2007), the selection of covariates to be included in regressions can be based on previous research (i.e. Chabé-Ferret & Subervie 2013 & Mezzatesta et al. 2013) and scientific understanding. Furthermore, using covariates measured prior to treatment assignment is fundamental to avoid including variables that may have been affected by the treatment (Stuart & Rubin 2008).

Choosing appropriate covariates, a variety of matching algorithms are employed to obtain valid counterfactuals including Nearest Neighbor, Optimal, and Genetic matching. Under nearest neighbor matching, best controls are found by minimizing a distance measure for each treated unit one at a time (Ho et al. 2011). Optimal matching provides matched samples with the smallest average absolute distance across all matched pairs (Ho et al. 2011). Genetic matching is a multivariate matching method that maximizes the balance of covariates across treatment and control(Diamond and Sekhon 2012). In essence, the method minimizes the discrepancy between distribution of potential cofounders in the treated and control groups, which allows for a maximized covariate balance (Sekhon 2011).

After data is matched, we use a Logistic regression for the adoption outcome variable, Y^3, and we use the same set of covariates employed in the matching process, X, to account for remaining covariance differences. For Y^1 and Y^2, we employ Tobit regressions since both are censored at zero for significant fraction of the observations. For censored dependent variables, using conventional regression methods such as OLS yield biased results (Greene 2008). Under the the standard Tobit model (Tobin 1985), the dependent variable is left censored at zero.

$$y_i = \beta_0 + \beta_T T_i + x_i^{'}\beta + \varepsilon_i \tag{1}$$

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } y_i^* > 0 \end{cases} \tag{2}$$

where i indicates the observation, y_i^* is the latent variable, $x_i^{'}$ is a vector of explanatory variables, T_i is the treatment, β is a vector of unknown parameters, and ε_i is the error term. Both Y^1 and Y^2 are censored at zero, since several farmers have zero cover crops acres making the proportion outcome variable also censored at zero. Likewise, no farmer planted cover cover crops in the entire farm, making $Y^2 < 1$ for all farmers.When estimating the treatment effect through these Tobit models, we apply the same set of covariates utilized in the matching process, X, to control for remaining covariance differences. To conclude the analysis, we focus on two marginal effects. First, we calculate the marginal effect on the expected value for our outcome variables for uncensored observations (i.e. both adopters and non-adopters of cover crops):

$$\frac{\partial E[Y^j|Y^j > 0]}{\partial T} = \beta_T \left[1 - \lambda(\alpha)\left(\frac{X_i\beta}{\sigma} + \lambda(\alpha)\right)\right] \tag{3}$$

for $j = 1, 2$ where $\lambda(\alpha) = \frac{\phi\frac{X_i\beta}{\sigma}}{\Phi\frac{X_i\beta}{\sigma}}$ is the Inverse Mills Ratio, σ is the Tobit scale, $\Phi()$ is the standard normal cdf and $\phi()$ is the standard normal pdf. This marginal effect indicates the treatment effect on uncensored observations (i.e. on observations where $Y^j > 0$). Secondly, we calculate the marginal effect on the expected value for both censored and uncensored observations (i.e. only cover crops adopters):

$$\frac{\partial E[Y^j]}{\partial T} = \Phi\left(\frac{X_i\beta}{\sigma}\right)\beta_T \tag{4}$$

We calculate these marginal effects for both outcome variables: Y^1 and Y^2.

3 Data

We use data from the Iowa Farm and Rural Life Poll (IFRLP) to estimate the effect of cost-share programs in the adoption of cover crops in Iowa. The IFRLP is an annual longitudinal survey of Iowa farmers that started in 1982, which has a sample of roughly 2000 Iowa operators. It is the longest-running survey of its kind in the United States (Arbuckle, Jr. et al. 2013). The survey focuses on a few particular subjects each year. For our analysis, we use data from the 2010 and 2011 polls, which provide pretreatment covariates[1]. In addition, the 2014 poll is used since it contains information for our three outcome variables. Furthermore, the 2014 poll is used to determine which farmers received cost-share funding to plant cover crops, which determines our treatment variable. Lastly, the 2014 poll also contains some pretreatment variables that do not change after harvest, when cover crops decisions are made and when cost-share is received. Using this data, we study the effect of cost-share funding on our three outcome variables.

Focusing on the 2014 poll solely, roughly 14 percent of surveyed farmers stated that they planted cover crops in 2013. The mean among cover crops adopters was 98 acres (IFRLP 2014). After merging the polls from 2010, 2011 and 2014, we have 588 observations. Using this merged data set, we construct several explanatory variables. Through this process, we drop observation with missing variables or inconsistent responses, resulting in 530 observations. Table 1 summarizes adoption among these observations. With this subset of the polls, we observe that roughly 17% of respondents adopted cover crops in 2013. There are almost twice as many adopters without cost-share as they are with this funding. Focusing on our outcome variables, Table 2 contains some summary statistics among those who adopted cover crops. We observe that the mean cover crops acres planted is around 109 acres in 2013. The range goes from 1 to 1700 acres. The median is 45 cover crops acres, which is much lower than the mean. For the proportion, the mean is around 20%, and the median is roughly 12%. The range of the proportion is between 0.2% and 80%.

In order to match treatment observations to valid counterfactuals, we use a list of covariates that affects both treatment and outcome variables. Following the literature, we use similar covariates as Mezzatesta et al. (2013) as well as additional variables available in the IFRLP. We make sure our covariates occurred prior to receiving cost-share funding and prior to planting cover crops in 2013. Table 3 describes each covariate used in the matching process and subsequent Logistic and Tobit models, displaying a combination of demographic and farm characteristics as well as some conservation information that might affect the enrollment into the

[1]Each poll contains questions about the previous year

cost-share program and the subsequent cover crops planting decision. Pretreatment outcome variables are ideal explanatory variables to include in both Propensity Score and regression models. However, we do not have information about previous cover crops acres planted. As a proxy, we use an indicator variable that captures the farmers who adopted cover crops in the last five years prior to 2010. Farm characteristics such as soil erosion problems, the presence of water running through or along the farm, farm size, proportion of farm acreage rented, the management of livestock, gross farm sales, and the proportion of farm acreage devoted to grain crops are included to help predict program enrollment as well as outcome variables. We also include location information to capture some of the differences among geographic locations based on weather, land characteristics, and other factors that are different among agricultural districts. Demographic information such as age, experience, farm income, education level, and the number of days worked off farm are also included. Lastly, we include the farmers' attitudes towards reducing nutrient or sediment runoff into waterways and previous conservation costs and drainage expenditures. For the latter, besides capturing drainage preference, we hope to use this variable as an indicator of the slope of the land. Land with little slope is more likely to require drainage systems. We note that slope and other agronomic characteristics are missing from our data set, which might affect our estimation results. Using available data, we try our best to capture explanatory variables that affect both treatment and outcome variables.

Table 4 summarizes explanatory variables prior to any matching process, showing some statistical significant difference in means between treated and control groups prior to matching. For instance, the sample mean of the dummy variable indicating water running on or along the farm is 0.90 for farmers receiving cost-share funding and 0.72 for farmers not receiving cost-share funding, a difference that is significantly different at a 99% level. The difference on the natural log of farm land is significantly different among treatment and control groups at the 95%. Lastly, differences in age, age squared, and the indicator variable for prior use of cover crops are significantly different at a 90% level. This table illustrates the importance of matching before any treatment analysis, since the treatment and control groups have some explanatory variables that are significantly different.

4 Matching

For the first step of our analysis, we use several matching algorithms including Nearest Neighbor, Optimal matching, and Genetic matching. We use different caliper levels, discarding options, distance measures (i.e. Logit, Probit, and Mahalanobis), and different number of control units to match to treated observations. As emphasized by Stuart (2010), we choose the best matched data set without using the outcome variable. For this section, we report results from the best matching method for our data. Robustness checks based on other matching methods are reported under the Robustness Checks Section.[2] We choose the best method based on the lowest standardized mean differences among all covariates (including the Propensity Score) and the overlap assumption. We obtain the best matching result using Nearest Neighbor matching with a

[2]Results from other matching methods are available upon request.

Probit Propensity Score, four nearest neighbors, a 0.25 caliper, allowing for replacement and the discarding of controls.

We first report the Probit Propensity Score results in Table 5. The Probit estimation shows that having cover crops for the last five years prior to 2010 has a positive effect in receiving treatment and its coefficient is statistically significant at a 99% level. Farm size and having water going through or along the farm also have a positive effect and their coefficients are statistically significance at a 95% level. Age affects treatment selection negatively, meaning that younger farmers are more likely to enroll in a cost-share program for cover crops. Differently, farm experience increases the likelihood to enroll in a cost-share program.

After matching using the specified method, we have 28 treated units matched to 90 control units for a total of 118 observations. Table 6 summarizes the standardized mean differences between treatment and control for all covariates and propensity score. These differences are the mean difference divided by the standard deviation in the original treated group (Ho et al. 2007). The highest absolute standardized mean difference is 0.1231. Only six covariates have absolute standardized mean differences above 0.10. The propensity score absolute difference in means is very small. According to Rubin (2001), absolute standardized mean difference of propensity score should be closer to zero. Following Rosenbaum and Rubin (1985), we want absolute standardized mean differences below 0.20. As depicted by Table 6, we obtain a good match based on this criteria.

To assess the common support of the matched sample, we use Figure 1, which depicts the overlap of the propensity scores between matched treated units and matched control units. Figure 1 displays jittered estimated propensity scores of treated and control units, which illustrates the overall distribution of propensity scores in treated and control groups. The size of each dot is proportional the weight given to the unit in the matching process (Ho et al. 2007). From this figure, we can observe that matched treated and matched control units have overlapping propensity scores. Lastly, we also illustrate the histogram of the estimated propensity scores before and after matching. Figure 2 shows the estimated propensity scores from the raw treated and control groups on the left. On the right, it shows the weighted histograms of the estimated propensity scores in the matched treated and control groups (Ho et al. 2007). We can see that matched treated and control groups look more similar after matching.

5 Estimation Results

After finding the matching method that attained the best balance among treatment and control groups, we proceed with regression analysis. For the adoption outcome variable, Y^3, we use a Logistic regression using X and T as independent variables and the weights from the matching results. We find that the coefficient on the treatment indicator is not statistically significant. We run logistic regressions with other matched data sets and find that the treatment coefficient is not statistically significant.[3] We also try using a Probit model and find the same results. Hence, we conclude that the effect of receiving cost-share funding on the adoption of cover crops is not statistically significant.

Since the other two outcome variables are censored at zero due to the lack of adoption of cover crops, we utilize a Tobit model to estimate the treatment effect of cost-share funding on both cover crops acres (Y^1) and the proportion of cover crops acres relative to total farm land (Y^2), as explained in the methodology session. For each Tobit regression, we use the weights from the matching procedure, and we employ the same set of covariates used in the matching process (X) in addition to the treatment indicator (T). Table 7 summarizes the results from the Tobit regression on cover crops acres. We observe that the treatment indicator (i.e. $T =$ cost.share.I) affects cover crops acres positively, and its coefficient is statistically significant at a 99% level. Other covariates are statistically significant, correcting for residual covariate imbalance between the groups (Ho et. al. 2007).

In order to assess the treatment effect, we compute two marginal effects as explained in the methodology section. Using Equations (3) and (4), each marginal effect is computed for each observation in the matched sample. After computing marginal effects for each observation, we calculate some summarize statistics reported in Table 8. The first row of Table 8 shows the marginal effect of cost-share funding for both adopters and non-adopters of cover crops. We observe that while the mean marginal effect of receiving cost-share funding is around 123 acres for both adopter and non-adopters of cover crops, we also observe that the mean is much higher than the median. The range of marginal effects goes from roughly zero to roughly 425 acres, which shows a large spread in marginal effects among matched observations. Focusing on adopters only, the mean marginal effect of receiving cost-share funding is around 116 acres. Again, this mean is higher than the median marginal effect, but less different than above.

For the proportion of cover crops planted relative to total farm land, we also employ a Tobit regression with the same set of covariates as in Table 7. Table 9 summarizes the results from this Tobit regression. Again, the coefficient on the treatment indicator is positive and it is statistically significant at the 99%. A few other covariates are also statistically significant, controlling for residual covariate imbalance between treated and control groups (Ho. et. al. 2007). Again, we use Equations (3) and (4) to compute each marginal effect for each observation in the matched sample. After this calculation, we compute some summarize statistics that are reported in Table 10. For the whole sample, including both adopters and non-adopters, we observe that the mean marginal effect of receiving cost-share funding on the proportion of cover crops relative to total farm land is roughly 16%, with a median of 9% and a range of roughly 45 percentage points. The mean marginal effect for only adopters is around 15%, with a median of 10% and a range of roughly 43 percentage points. We conclude that having cost-share funding increases cover crops acres and the proportion of farm land devoted to cover crops.

[3] Regression results are available upon request

6 Robustness Checks

In order to evaluate the robustness of our results, we report results from two other matching methods that attained a good balance from the first step of our research analysis. Nearest Neighbor matching with a Probit Propensity Score, five nearest neighbors, no caliper, allowing for replacement and the discarding of controls also provides a decent balance. The lowest absolute standardized mean difference is 0.1337.[4] Furthermore, Optimal matching with a Logit Propensity Score, four nearest neighbors, no caliper, and without replacement or discarding also offers a good match. The lowest standardized mean difference under this matching model is 0.1474.[5]. Tobit regressions are run for each matched data set with each outcome variable as the dependent variable. For the four regressions, the coefficient on the treatment indicator is positive and it is statistically significant at the 99% level, which agrees with our previous results.

Table 11 summarizes the marginal effects under each matched data set. While both matching models provide very similar results for each marginal effect in Table 11, we observe some discrepancies in marginal effects for cover crops acres planted with previous results. Specifically, we observe differences in mean marginal effects, in which Table 8 reports the mean of $\frac{\partial E[Y^1]}{\partial T}$ equal to 122.53 and the mean of $\frac{\partial E[Y^1|Y^1>0]}{\partial T}$ equal to 115.95. These means are higher than their analogous mean in Table 11. However, median marginal effects differ less between both tables. For the proportion outcome variable, we see more consistency across results. The mean of $\frac{\partial E[Y^2]}{\partial T}$ is between 14% and 16% and the mean of $\frac{\partial E[Y^2|Y^2>0]}{\partial T}$ is between 14% and 15%. The median marginal effects among the three matched data sets are also very close. Based on these marginal effects, we can conclude that having cost-share funding increases the proportion of cover crops acres planted, on average, by roughly 15% among adopters and non-adopters of cover crops. Among adopters only, having cost-share funding increases the proportion of cover crops acres planted, on average, by 14%. The marginal effects for cover crops acres are large and positive, but they vary in size between matching models.

7 Conclusions

Cover crops are a conservation practice that has been promoted by the Iowa Nutrient Reduction Strategy. In August of 2013, cost-share funds became available to establish cover crops, among other conservation practices, with the goal of providing water quality benefits in 2013 and spring 2014 (Iowa NRS 2014). Motivated by the importance of the adoption of cover crops in Iowa and the availability of cost-share funding for this practice in 2013, we use matching methods combined with regression analysis to study the effect of cost-share funding on the adoption of cover crops in Iowa in 2013 using the Iowa Farm and Rural Life Poll. Following a two-step process, we first match treated and control units based on pretreatment information using a variety of matching methods including Nearest Neighbor, Optimal, and Genetic matching. We choose

[4] Complete matching results are available upon request
[5] Complete matching results are available upon request

the best matched data set based on standardized mean difference and the overlap of propensity scores between treated and control groups. For the second step, we perform regression analysis on the chosen matched data by regressing the outcome variable on the treatment indicator and other relevant covariates. We study three outcome variables: the adoption of cover crops, cover crops acres, and the proportion of cover crops relative to total farm land. For the adoption of cover crops, we use a Logistic regression and find that the coefficient of the treatment indicator is not statistically significant across several models using various matched data sets. For the other two outcome variables, we use a Tobit model, since a substantial portion of the outcome variables are censored at zero. We find that receiving cost-share funding has a positive effect on both cover crops acres and on the proportion of cover crops. Moreover, the treatment coefficients are statistically significant at the 99% level under each Tobit model, implying that receiving cost-share acres has a positive effect on cover crop acres and on the proportion of cover crops. Moreover, the mean marginal effect on the proportion of cover crops acres is around 15% among adopters and non-adopters and 14% among adopters. Consequently, having cost-share funding increases the proportion of cover crops acres planted by 15% among all farmers and by 14% among adopters on average. For cover crops acres, we find that the mean marginal effect differs among matched data sets, ranging from 106 to 123 acres among adopters and non-adopters and ranging from 102 to 106 acres among adopters only. We conclude that cost-share funding increases cover crops acres on average, but the size of its marginal effect varies with different matched data sets. In the end, cost-share funding is effective in increasing cover crops acres and the proportion of cover crops planted.

8 Tables and Figures

Figure 1: Distribution of Propensity Scores

Table 1: Cover Crops Adoption

Number of Non-adopters	Number of Adopters with Cost-Share	Number of Adopters without Cost-Share	Number of Observations
442	29	59	530

Table 2: Summary Statistics for Outcome Variables among Adopters

Outcome Variable	Min	1st Quarter	Median	Mean	3rd Quarter	Max
Y^1: cover crops acres	1	19.25	45	109.20	110.50	1700
Y^2: proportion	0.002	0.0439	0.1237	0.196	0.2817	0.80

This table focuses on the behavior of adopters only (i.e $Y^1 > 0$ and $Y^2 > 0$).

Figure 2: Histogram of Propensity Scores

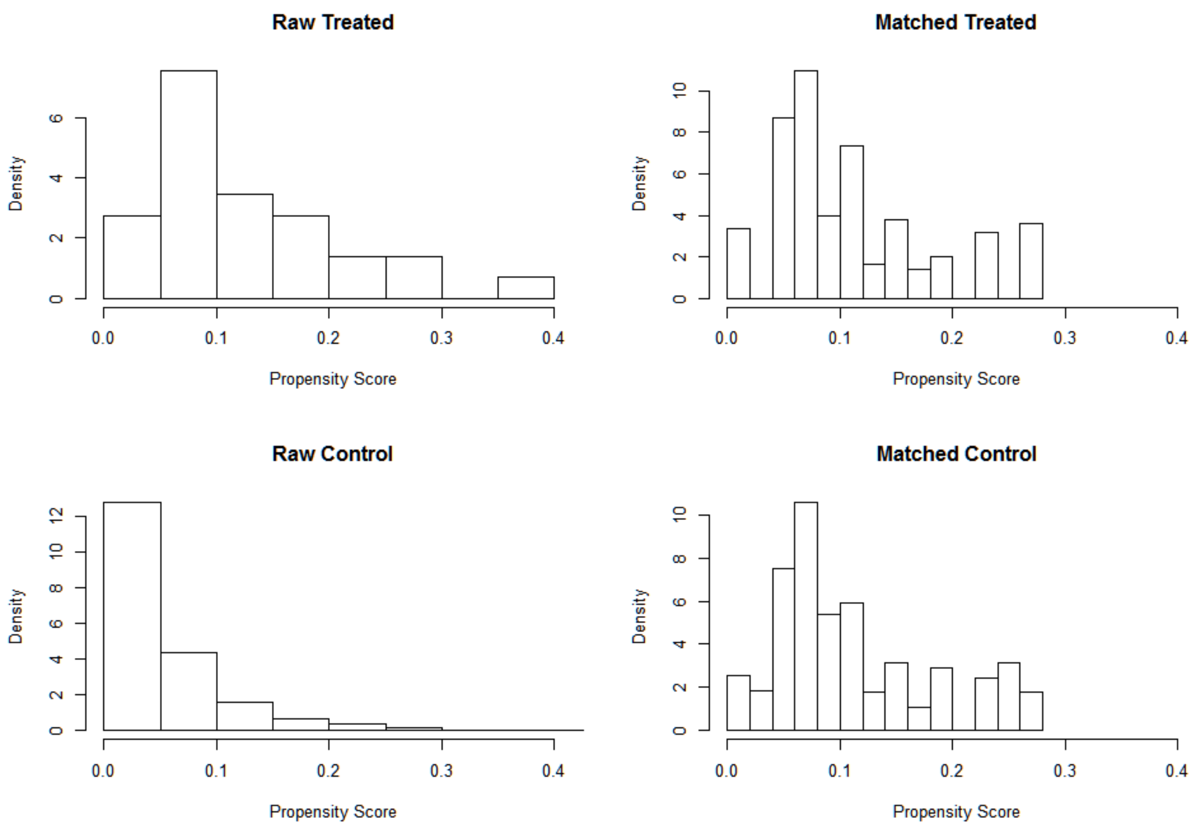

Table 3: Explanatory Variables Description

Covariate	Definition
cover.crops.2010.I	=1 if farmer planted cover crops in the last five years prior to 2010
water.on.or.along.farm	=1 if farmer indicated that creeks, streams, or rivers run through or along the farm
soil.erosion	=1 if farmer indicated to have had significant soil erosion on any of his or her land in the last five years in 2011
attitude.reduction	=1 if farmer believes that Iowa farmers should do more to reduce nutrient and or sediment runoff into waterways
conservation.costs.I	=1 if farmer had incurred in any costs associated with conservation practices (excluding tile or similar drainage systems) over the past 10 years in 2010
drainage.expenditure.I	=1 if farmer had any expenditure associated with agricultural drainage systems over the past 10 years in 2010
log.ag.land	natural log of total farm acreage operated in 2013
rented	proportion of farm acreage rented in 2013
labor.off.farm	number of days worked off the farm in 2009
gross.farm.sales.I	=1 if farmer had gross farm sales above $250,000 in 2009
farm.income.I	=1 if percent of total net household income from the farm was above 51% in 2009
age	age of farmer
age.sq	age squared
college	=1 if the highest level of education completed was at least a Bachelor's degree in 2011
Central	=1 if farm is located in Central Agricultural District
East.Central	=1 if farm is located in East Central Agricultural District
West.Central	=1 if farm is located in West Central Agricultural District
North.Central	=1 if farm is located in North Central Agricultural District
North.East	=1 if farm is located in North East Agricultural District
North.West	=1 if farm is located in North West Agricultural District
South.Central	=1 if farm is located in South Central Agricultural District
South.West	=1 if farm is located in South West Agricultural District
livestock.I	=1 if farmer managed livestock in 2013
exp	number of years farming in the USA
exp.sq	experience squared
grains	proportion of farm acreage devoted to grain crops in 2013

Table 4: Summary Statistics of Explanatory Variables

	Treatment	Control	Difference in Means
cover.crops.2010.I	0.24	0.09	0.15 .
water.on.or.along.farm	0.90	0.72	0.18 **
soil.erosion	0.31	0.28	0.03
attitude.reduction	0.83	0.83	0.00
conservation.costs.I	0.59	0.51	0.08
drainage.expenditure.I	0.59	0.54	0.05
log.ag.land	6.14	5.63	0.51 *
rented	0.37	0.32	0.05
labor.off.farm	90.48	78.94	11.54
gross.farm.sales.I	0.38	0.28	0.10
farm.income.I	0.52	0.51	0.01
age	62.66	66.18	-3.52 .
age.sq	4025.28	4463.62	-438.34 .
college	0.38	0.34	0.04
Central	0.17	0.14	0.03
East.Central	0.07	0.14	-0.07
West.Central	0.10	0.12	-0.02
North.Central	0.10	0.13	-0.03
North.East	0.17	0.12	0.05
North.West	0.14	0.14	0.00
South.Central	0.14	0.06	0.08
South.West	0.07	0.07	0.00
livestock.I	0.28	0.24	0.04
exp	39.79	41.40	-1.61
exp.sq	1671.17	1852.33	-181.16
grains	0.83	0.78	0.05

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Statistical significance is based on Welch Two Sample t-tests.

Table 5: Propensity Score Results - Probit Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.5179	3.5331	0.15	0.8835
cover.crops.2010.I	0.7391	0.2732	2.71	0.0068 **
water.on.or.along.farm	0.5947	0.3012	1.97	0.0483 *
soil.erosion	-0.0268	0.2306	-0.12	0.9073
attitude.reduction	-0.1646	0.2676	-0.62	0.5385
conservation.costs.I	-0.1535	0.2381	-0.64	0.5190
drainage.expenditure.I	-0.1136	0.2343	-0.48	0.6278
log.ag.land	0.3062	0.1530	2.00	0.0453 *
rented	-0.1975	0.3536	-0.56	0.5765
labor.off.farm	0.0006	0.0011	0.53	0.5946
gross.farm.sales.I	-0.1487	0.2857	-0.52	0.6028
farm.income.I	-0.1656	0.2605	-0.64	0.5249
age	-0.2086	0.1217	-1.71	0.0864 .
age.sq	0.0014	0.0010	1.51	0.1323
college	-0.0072	0.2252	-0.03	0.9745
Central	0.6780	0.5641	1.20	0.2294
East.Central	0.3855	0.6004	0.64	0.5209
West.Central	0.4840	0.5941	0.81	0.4153
North.Central	0.3949	0.5900	0.67	0.5033
North.East	0.7596	0.5711	1.33	0.1835
North.West	0.6671	0.5871	1.14	0.2559
South.Central	1.1256	0.6072	1.85	0.0638 .
South.West	0.5587	0.6335	0.88	0.3778
livestock.I	-0.0161	0.2489	-0.06	0.9484
exp	0.1300	0.0719	1.81	0.0706 .
exp.sq	-0.0015	0.0009	-1.71	0.0869 .
grains	0.1054	0.4948	0.21	0.8313

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 6: Matching Results

	Absolute Std. Mean Difference
Propensity Score	0.0050
cover.crops.2010.I	0.0273
water.on.or.along.farm	0.0864
soil.erosion	0.1138
attitude.reduction	0.0232
conservation.costs.I	0.0356
drainage.expenditure.I	0.0238
log.ag.land	0.0575
rented	0.0907
labor.off.farm	0.0019
gross.farm.sales.I	0.0904
farm.income.I	0.1112
age	0.1231
age.sq	0.1174
college	0.0663
Central	0.0465
East.Central	0.0692
West.Central	0.0576
North.Central	0.0000
North.East	0.0774
North.West	0.0000
South.Central	0.0254
South.West	0.0231
livestock.I	0.0327
exp	0.1171
exp.sq	0.1227
grains	0.0400

Table 7: Tobit Model for Cover Crops Acres Outcome Variable

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2463.94	1398.19	1.762	0.07803 .
cost.share.I	425.38	77.81	5.467	4.58e-08 ***
cover.crops.2010.I	276.02	97.58	2.829	0.00468 **
water.on.or.along.farm	-80.67	121.00	-0.667	0.50496
attitude.reduction	104.09	108.23	0.962	0.33619
soil.erosion	-5.35	92.20	-0.058	0.95372
conservation.costs.I	87.68	87.47	1.002	0.31616
drainage.expenditure.I	27.72	91.25	0.304	0.76135
log.ag.land	1.38	58.02	0.024	0.98096
rented	231.92	146.62	1.582	0.11368
labor.off.farm	0.27	0.44	0.624	0.53262
gross.farm.sales.I	252.9	109.24	2.315	0.02061 *
farm.income.I	20.46	107.36	0.191	0.84884
age	-82.59	49.80	-1.659	0.09722 .
age.sq	0.74	0.39	1.896	0.05798 .
college	-159.01	93.95	-1.693	0.09054 .
Central	249.33	212.60	1.173	0.24090
East.Central	49.80	252.70	0.197	0.84378
West.Central	74.27	224.57	0.331	0.74087
North.Central	-145.56	228.54	-0.637	0.52417
North.East	52.23	221.75	0.236	0.81379
North.West	-2.998	250.84	-0.012	0.99046
South.Central	-30.28	235.25	-0.129	0.89757
South.West	-72.24	239.58	-0.302	0.76302
livestock.I	80.19	89.52	0.896	0.37041
exp	-48.41	34.50	-1.403	0.16060
exp.sq	0.75	0.46	1.650	0.09884 .
grains	-332.00	162.19	-2.047	0.04066 *
Log(scale)	5.48	0.11	50.791	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8: Marginal Effects of Cost-Share Funding on Cover Crops Acres

Marginal Effect	Min	Median	Mean	Max
$\frac{\partial E[Y^1]}{\partial T}$	0.00	55.03	122.53	425.38
$\frac{\partial E[Y^1 \| Y^1>0]}{\partial T}$	16.45	78.52	115.95	425.34

Y^1 : cover crops acres

Table 9: Tobit Model for Proportion Outcome Variable

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.7148	1.0524	0.679	0.496981
cost.share.I	0.4529	0.0593	7.640	2.17e-14 ***
cover.crops.2010.I	0.2439	0.0729	3.344	0.000825 ***
water.on.or.along.farm	-0.1148	0.0898	-1.279	0.201004
attitude.reduction	-0.0072	0.0784	-0.092	0.926721
soil.erosion	-0.0583	0.0721	-0.809	0.418694
conservation.costs.I	0.1289	0.0676	1.907	0.056523 .
drainage.expenditure.I	0.0245	0.0669	0.365	0.714904
log.ag.land	-0.1264	0.0452	-2.798	0.005141 **
rented	0.1509	0.1078	1.400	0.161438
labor.off.farm	0.0002	0.0003	0.886	0.375650
gross.farm.sales.I	0.2258	0.0822	2.748	0.005989 **
farm.income.I	0.0148	0.0764	0.194	0.846469
age	-0.0168	0.0375	-0.448	0.654247
age.sq	0.0002	0.0003	0.641	0.521262
college	-0.1724	0.0714	-2.412	0.015852 *
Central	0.1518	0.1554	0.977	0.328535
East.Central	0.0007	0.1809	0.004	0.996771
West.Central	-0.0594	0.1675	-0.355	0.722821
North.Central	-0.2604	0.1707	-1.525	0.127191
North.East	-0.0693	0.1647	-0.421	0.674049
North.West	0.0274	0.1866	0.147	0.883095
South.Central	-0.0305	0.1688	-0.181	0.856458
South.West	-0.0172	0.1788	-0.096	0.923448
livestock.I	0.0756	0.0688	1.098	0.272204
exp	-0.0042	0.0262	-0.159	0.873552
exp.sq	0.0001	0.0003	0.423	0.672088
grains	-0.1108	0.1165	-0.951	0.341415
Log(scale)	-1.6907	0.1091	-15.490	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 10: Marginal Effects of Cost-Share Funding on Proportion of Cover Crops Acres Relative to Total Farm Land

Marginal Effect	Min	Median	Mean	Max
$\frac{\partial E[Y^2]}{\partial T}$	0.00	0.09	.16	.45
$\frac{\partial E[Y^2 \mid Y^2>0]}{\partial T}$	0.02	0.10	.15	.45

Y^2 : proportion of cover crops planted relative to total farm land

Table 11: Marginal Effects of Cost-Share using other Matching Methods

Matching Model	Marginal Effect	Min	Median	Mean	Max
Nearest	$\frac{\partial E[Y^1]}{\partial T}$	0.01	53.57	107.43	399.36
Optimal	$\frac{\partial E[Y^1]}{\partial T}$	0.33	55.06	106.23	348.69
Nearest	$\frac{\partial E[Y^1 \mid Y^1>0]}{\partial T}$	17.96	74.69	103.92	399.36
Optimal	$\frac{\partial E[Y^1 \mid Y^1>0]}{\partial T}$	25.77	74.04	101.84	367.21
Nearest	$\frac{\partial E[Y^2]}{\partial T}$	0.00	0.07	.14	.47
Optimal	$\frac{\partial E[Y^2]}{\partial T}$	0.00	0.08	.15	.47
Nearest	$\frac{\partial E[Y^2 \mid Y^2>0]}{\partial T}$	0.03	0.09	.14	.47
Optimal	$\frac{\partial E[Y^2 \mid Y^2>0]}{\partial T}$	0.03	0.10	.14	.45

Y^1 : cover crops acres; Y^2 : proportion of cover crops planted relative to total farm land

References

[1] Abadie, A. & Imbens, G. W. (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, *74*, 235-267.

[2] Andam, K. S., Ferraro, P.J., Pfaff, A., Sanchez-Azofeifa, G. A., & Robalino, J.A., (2008). Measuring the Effectiveness of Protected Area Networks in Reducing Deforestation. *Proceedings of the National Academy of Sciences*, *105*(42), 16089-16094.

[3] Arbuckle Jr, J.G., Lasley, P., Korshing, P., & Kast, C. (2010). *Iowa Farm and Rural Life Poll 2010 Summary Report*. Ames: Iowa State University Extension.

[4] Arbuckle, J.G., Lasley, P., & Ferrell, J. (2011). *The Iowa Farm and Rural Life Poll 2011 Summary Report*. Ames: Iowa State University Extension.

[5] Arbuckle Jr., J.G., (2013). Farmer support for extending Conservation Compliance beyond soil erosion: Evidence from Iowa. *Journal of Soil and Water Conservation*, *68*(2), 99-109.

[6] Arbuckle Jr, J.G. & Lasley, P. (2013). *Iowa Farm and Rural Life Poll 2013 Summary Report*. Ames: Iowa State University Extension.

[7] Arbuckle Jr., J.G., Morton, L.W., & Hobbs, J. (2013). Farmers beliefs and concerns about climate change and attitudes toward adaptation and mitigation: Evidence from Iowa. *Climate Change*, *118*, 551-563.

[8] Arbuckle Jr., J.G., Morton, L.W., & Hobbs, J. (2013). Understanding Farmer Perspectives on Climate Change Adaptation and Mitigation: The Roles of Trust in Sources of Climate Information, Climate Change Beliefs, and Perceived Risk. *Environment and Behavior*, 1-30.

[9] Chabé-Ferret, S., & Subervie, J. (2013). How much green for the buck? Estimating additional and windfall effects of French agro-environmental schemes by DID-matching. *Journal of Environmental Economics and Management*, *62*, 12-27.

[10] Cocharn, W. G., & Rubin, D. B. (2013). Controlling Bias in Observational Studies: A Review. *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, *35*(4), 417-466.

[11] Diamond, A., & Sekhon, J.S. (2012). Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *Review of Economics and Statistics (forthcoming)*, *2012*, 1-41.

[12] EPA (2014). *Northern Gulf of Mexico Hypoxic Zone*. Obtained from Mississippi River Gulf of Mexico Watershed Nutrient Task Force: http://water.epa.gov/type/watersheds/named/msbasin/zone.cfm

[13] Ferraro, P. J., McIntosh, C., & Ospina, M. (2007). The Effectiveness of the US Endangered Species Act: An Econometric Analysis using Matching Methods. *Journal of Environmental Economics and Management*, *54*(3), 245-261.

[14] Glazerman, S., Levy, D. M. & Myers, D. (2003). Nonexperimental versus Experimental Estimates of Earnings Impacts. *Annals of the American Academy of Political Science*, *589*, 63-93.

[15] Greene, W. H. (2008). *Econometric Analysis* 6th edition. Prentice Hall.

[16] Heckman, J.J. (1998). Characterizing Selection Bias Using Experimental Data. *Econometrica*, *66*, 1017-1098.

[17] Hansen, B.B. (2004). Full Matching in an Observational Study of Coaching for the SAT. *Journal of the American Statistical Association*, *99*(467), 609-618.

[18] Ho, D. E., Imai, K., King, G., Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, *42*(8), 1-28.

[19] Ho, D. E., Imai, K., King, G., Stuart, E. A. (2007). Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*, *15*, 199-236.

[20] Holland, P.W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, *81*, 945-960.

[21] Imbens, G.M., & Wooldridge, J.M. (2008). Recent Developments in the Econometrics of Program Evaluations. *NBER Working Papers Series*, Working Paper 14251.

[22] Iowa NRS (2013) Iowa Department of Agriculture and Land Stewardship; Iowa Department of Natural Resources. Iowa State University College of Agriculture and Life Sciences. *Iowa Nutrient Reduction Strategy.* Ames, Iowa.

[23] Iowa NRS (2014). *Iowa Nutrient Reduction Strategy 2013-2014 Annual Progress Report.* Ames: Iowa State University.

[24] Liu, X., & Lynch, L. (2011). Do Agricultural Land Preservation Programs Reduce Farmland Loss? Evidence from Propensity Score Matching Estimator. *Land Economics*, *87*(2), 183-201.

[25] Mezzatesta, M., Newburn, D.A., & Woodward, R.T. (2013). Additionality and the Adoption of Farm Conservation Practices. *Land Economics*, *89*(4), 722-742.

[26] Rosenbaum, P. & Rubin, D.B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporates the Propensity Score. *The American Statistician*, *39*(1), 33-38.

[27] Rosenbaum, P. & Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, *40*, 688-701.

[28] Rubin, D.B. (2001). Using Propensity Scores to Help Design Observational Studies: An Application to the Tobacco Litigation. *Health Services & Outcomes Research Methodology*, *2*, 169-188.

[29] Rubin, D.B. (1976). Inference and Missing Data (with discussion). *Biometrika*, *63*, 581-592.

[30] Rubin, D.B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology*, *66*, 688-701.

[31] Sekhon, J.S. (2011, May). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*, *42*(7), 1-52.

[32] Smith,J.A., & Todd, P.E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators?. *Journal of Econometrics*, *125*, 305-353.

[33] Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, *25*, 1-21.

[34] Stuart, E. A., & Rubin, D.B. (2008). 11 Best Practices in Quasi-Experimental Designs: Matching Methods for Causal Inference. In J. Osborne, *Best Practices in Quantitative Methods*, Chapter 11. SAGE Publications, Inc.

[35] Tobin, J. (1958). Estimation of relationships for limited dependent variables *Econometrica*, *26*, 24-36.

[36] USDA NASS (2014). *2014 State of Agriculture Overview.*

[37] Wooldridge, J.M. (2010) . *Econometric Analysis of Cross Section and Panel Data.* Cambridge: The MIT Press.