# Farmers' Willingness and Expected Economic Benefit to Adopt BMPs: an Application of Multivariate Imputation by Chained Equation Method

**Hua Zhong**
**Ph.D. Candidate, University of Kentucky**
**hua.zhong@uky.edu**

**Wuyang Hu**
**Professor, University of Kentucky**
**wuyang.hu@uky.edu**

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

**Abstract:** Water Quality Trading (WQT) programs may offer farmers compensation to adopt Best Management Practices (BMPs). We conducted a survey of farmers in the Kentucky River watershed from 2011 to 2012. With respect to the five types of BMPs considered in the survey, about 20% of respondents did not indicate how much they will adopt. Missing responses are common for surveys on farming decisions. We compare three methods to handle the missing data: deleting the observations with missing value, mean imputation, and Multivariate Imputation by Chained Equation (MICE). Following these missing data treatments, we estimate the factors affecting how much farmers may engage in BMPs using Tobit or Poisson model. The results show that increasing the compensation for using BMPs is more likely to encourage farmers to adopt riparian buffers. In addition, land area, percentage of household income from farming, percentage of total household income reinvested back to farm, and current experience of BMPs will affect BMP adoption. The results obtained after using the MICE are more promising and reasonable than using the deletion or the mean imputation method. Implications are discussed for farmers' BMP adoptions under WQT while missing observations are present.

Key words: Best management practices, water quality trading, Multivariate Imputation by Chained Equation

Water quality trading (WQT) programs are created to assist dischargers in a watershed to meet the Total Maximum Daily Load (TMDL) provisions of the Clean Water Act of 1972 (CWA). Under TMDLs, point and nonpoint sources (PSs and NPSs) dischargers are encouraged to trade emission permits, thus water quality standards are achieved at a lower cost than traditional regulations. In WQT programs, agricultural NPSs are considered to create credits for the trading market by adopting Best Management Practices (BMPs).

However, point-nonpoint WQT programs have not been developed successfully. Only 4 programs have trading occurred to date in the 15 established point-nonpoint trading programs, especially in the trading market related to agricultural NPSs. Shortle (2013) states that most of economic research on WQT have focused on market design instead of market prediction and uncertainty, such as how much participants will trade, and what factors are likely to hinder trading. If farmers in a watershed have already adopted most BMPs on their land, their capacity to use additional BMPs may limit trading (Ribaudo and Gottlieb, 2011). Therefore, our research is interested in: whether are farmers willing to further reduce agricultural runoff and what are the factors affecting their intention? By how much will they adopt additional BMPs on their land, upon payment from a trading program, to generate trading credits?

We conducted a survey of farmers in the Kentucky River watershed from 2011 to 2012 to explore how much farmers may engage their lands in Best Management Practices (BMPs) through a water quality trading (WQT) program in Kentucky, and also investigate the factors affecting farmers' ability to implement additional BMPs. The survey asked questions about whether and how much farmers may adopt the BMPs (in addition to what they have already used) if they are offered compensation through WQT. Five BMPs are featured: riparian buffers, fencing off animals, no-till, waste storage facilities, and nutrient management. With respect to five different types of BMPs, about 21.5%, 26.9%, 24.2%, 23.2%, and 18.2% of respondents did not indicate how much they will adopt BMPs. Before analyzing, missing responses in our survey need to be addressed since the percentage of missing data is more than the 5% rule of thumb (Schafer, 1999). Therefore, the research goals of this paper is to (1) analyze how much farmers in Kentucky may engage their lands in BMPs through WQT programs and (2) address the missing issue in our survey.

Missing data problems occur in most of primary data-sets, and are common in surveys of farmers. Weber and Clay (2013) research the nonresponse issues in the USDA Agricultural Resource Management Survey (ARMS) and conclude that the time consumed and disutility from answering questions account for main reasons for nonresponse in ARMS. These effects are amplified for farms with larger sales. A naive method to handle the missing responses is to delete the observations with missing portions, known as listwise deletion method. This method assumes that the missing responses are independent with the observed and unobserved variables. Unfortunately, this assumption is rarely satisfied in empirical studies and the listwise deletion method may lead to nonresponse bias (Lin and Schaeffer (1995), Groves (2006), and Groves and Peytcheva (2008)).

Our data include two variables facing the issue of missing responses: one is the missing response in the yes/no choice question of whether farmers would like to accept our offer to implement BMPs through WQT programs, and the other is the missing response in the follow-up questions on how much farmers will adopt BMPs if they decide to accept our offer. In this research, we use Multiple Imputation (MI) to address the issue of missing data. MI, introduced by Rubin (1978), is a statistical method that impute *m* plausible missing values for each missing unit to create *m* completed datasets; each completed dataset is analyzed using a statistical method separately; then the *m* results, point estimation and covariance matrices, are averaged into final estimates using Rubin's formula (1978) (Raghunathan, Lepkowski, Hoewyk and Solenberger 2001; King, Honaker, Joseph and Scheve 2001).

To be specific, we applied a multivariate MI method referred to as the Multivariate Imputation by Chained Equation (MICE) introduced by Raghunathan et al. (2001). This current study imputes the missing data in six scenarios. The first scenario is to delete observations with missing responses in the analysis. The second scenario is to replace missing values of "how-much" by the observed mean. In the third scenario, we impute the missing response in the follow up question if they accept the offer given in the survey. The fourth to sixth scenarios are the multi-stage procedure that firstly imputes the missing response in the yes/no question; then imputes the missing value in the follow-up question if the answer to the first question is yes. The specific procedure is introduced in the empirical strategy section. Given the imputed value, we estimate the factors affecting how much farmers may engage in BMPs using Tobit or Poisson regression, and combine the *m* results to a final estimate.

The next section describes our empirical survey and missing issues in our research. Following the survey and missing data problem section, we introduce the theory of the mechanism of missing data. Then, we discuss empirical strategies to address the missing data in our survey and the imputation procedure. The last two sections display the result of the data analysis, and conclude with the policy implication of our research.

## Survey and missing problem

### Survey

The survey data were collected from randomly chosen farmers across 35 counties in the Kentucky River watershed from 2011 to 2012. The response rate was 23%, and there were 357 valid observations out of 459 responses. The survey questions included farmers' participation in current government-funded environmental or conservation programs, their potential adoption of additional BMPs through a WQT program, farm characteristics, as well as respondents' demographic characteristics.

The key BMP adoption questions asked farmers: "Regardless of whether you are currently participating in any government cost share programs, if you knew that by using water quality management practices on your land, a nearby waste/sewage water treatment plant or factory will cover X% of your cost of implementing these practices, would you be interested in using additional water quality management practices (BMPs) in the form of the following activities: riparian buffers, fencing off animals, no-till, waste storage facility and nutrient management?" In the actual survey, X% is replaced with one of the following levels with equal probability: 75%, 80%, 85%, 90%, 95%, 100%, 105%, 110%, 115% and 120%. When answering the survey, each respondent will see only one questionnaire with one of the possible levels of compensation. A respondent could answer "yes", "no", or "not possible for me" with respect to each practice. The "not possible for me" option captures the possibility that farmers have already maximized their potential to adopt BMPs, or whether BMPs are applicable on their land.

If respondents would like to consider adopting a BMP given the compensation in the survey, the following question was asked "if yes, by how much in addition to what you have adopted already would you like to adopt this practice?" The respondents could answer exact values for how much they would like to adopt the practice. The measurement unit for the practices of Riparian buffers and Animal fences is “feet”; the measurement unit for the practices of No till and Nutrient management is “acre”; and the measurement unit for the practices of Waste storage facility is number of facilities, “unit”.

Furthermore, the survey is designed with four types of information explaining the meaning of WQT programs. One of the four levels of the information is randomly assigned with equal probability to the survey. This design is to examine whether the different levels of information will influence an individual’s response. The first type of information is the baseline with basic explanation of WQT programs. The information does not contain any further description or interpretation of WQT programs. The second type of information includes the information in the first type but also includes additional message on WQT programs focusing on their cost saving implications. The third type contains the baseline information and also information emphasizes the environmental benefit from WQT programs. The fourth type provides the baseline as well as explanation of WQT programs focusing on both cost saving and environmental benefit information.

Table 1 presents all variables and summary statistics for the entire sample. Table 2 explains discrete levels in explanatory variables.

**Missing data problem**

Missing responses to the BMP adoption questions are analyzed with three cases. The first case is when respondents answered “no” or “not possible for me” to the yes/no questions, the responses to the quantity to be adopted will be missing as well. Logically, if the respondents would not like to consider the BMPs, they are not able to implement BMPs on their land, or they do not have the additional ability to implement BMPs on their land because they have already adopted BMPs as much as they can. In these cases, the plausible values for the missing data on amount to be adopted are “zero”. Therefore, missing values in this case can be replaced by “zero”.

The second case of missing data is when respondent answered “yes” to the yes/no questions, but they did not respond how much they would like to use BMPs on their land. Because respondents have already stated they would like to adopt BMPs given the compensation in the survey, the plausible values for the missing data in the follow-up question should be some positive values, continuous for the practices of Riparian buffers, Animal fences, No till and Nutrient management, and counts for the practice of Waste storage facilities. The third case is that respondents did not answer whether they would like to consider adopting a BMP, the answer to the amount to be adopted would therefore be also missing. If respondents did not respond to any of the five BMPs, their responses are not considered in our analysis. If respondents answered the yes/no question to at least one of the practices, they are treated as in this third case as missing. In the third case, the plausible values for missing data in yes/no questions are categorical variables representing “yes”, “no” and “not possible for me”; if respondents actually answered “yes” to the yes/no questions or are imputed to be “yes” responses, the plausible values for the quantitative questions are the same as in the second case. Figure 1 illustrates the three missing cases in our survey.

Tables 3-7 show the statistical summary of the missing data in the survey with respect to each practice. For example, in Table 3, the first column shows the frequency of responses to the decision question with respect to Riparian buffers, and the numbers of respondents who did not respond. The second column displays the frequency of responses for the follow-up question for when respondents answered “yes” to consider adopting the practice. The third column shows the mean of the observed values of “how much would you like to adopt Riparian buffers”. Following Table 3, Tables 4-7 also display the number of observations who responded but did not respond to the yes/no decision questions; the number of observations who answered but did not answer the quantitative questions after the “yes” response in the decision questions; and the observed mean of the “how much” variable, with respect to the practices of Animal fences, No till, Waste storage facility and Nutrient management. Across Tables 3-7, missing data issues exist in all of the BMPs examined in our survey.

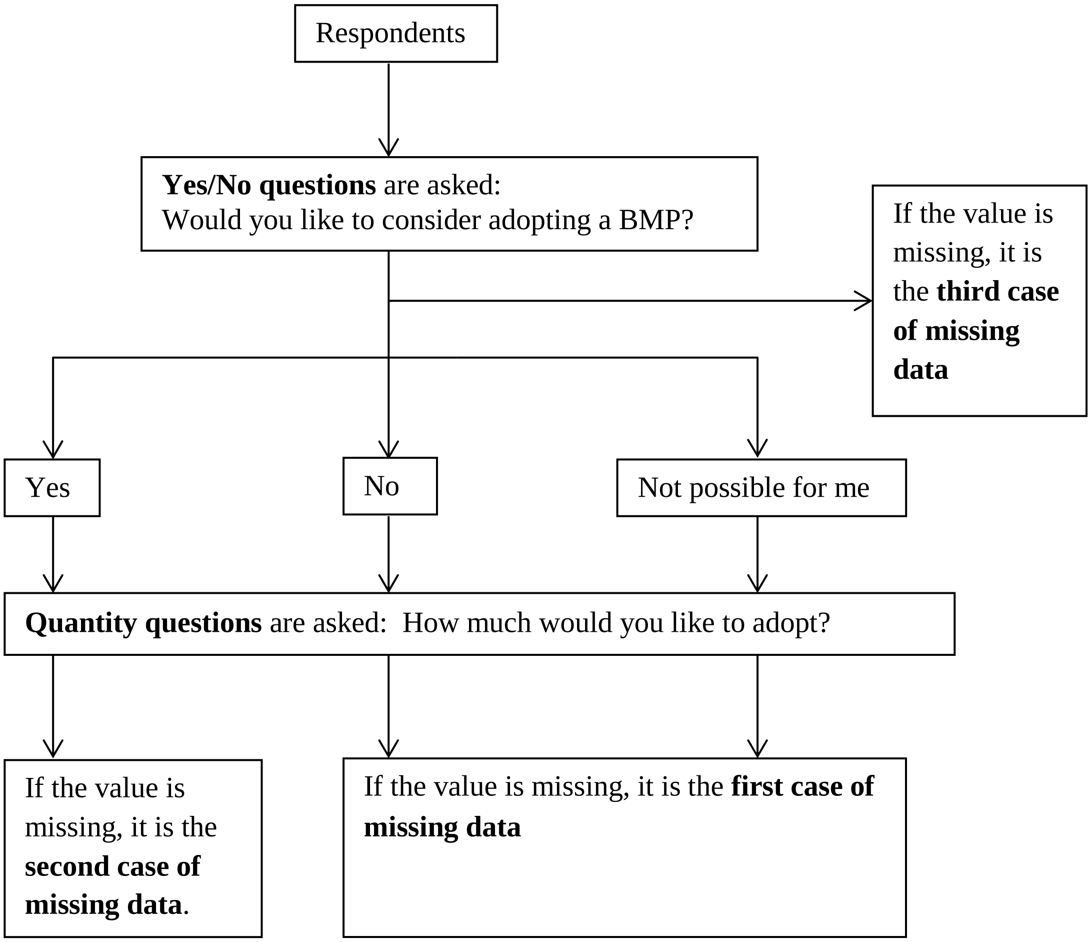


Figure 1 The explanation of missing data in the survey

# Theory of Missing Mechanism and Multiple Imputation

### Missing mechanism

This section introduces three types of missing data mechanisms. Let Y denote a variable with missing data, X denote variables observed completely, R be an indicator variable that equals one if Y is missing and zero if Y is observed. The first type of missing data is missing completely at random (MCAR) and is defined as

$$\Pr(R = 1|X, Y) = \Pr(R = 1)$$

The MCAR implies that missing data do not depend on any observed or unobserved variables. If the MCAR mechanism applies, the listwise deletion method that deletes the observations with missing data is the most common and efficient strategy to address the missing issue. However, the MCAR rarely holds true in the empirical analysis because it indicates that missing responses arise completely by chance (Kenward and Carpenter 2007).

The second type of missing data is missing at random (MAR) and is represented as

$$\Pr(R = 1|X, Y) = \Pr(R = 1|\, X)$$

$$\text{or}\ \ \Pr(R = 1|X, Y) = \Pr(R = 1|\, Y_{observed})$$

The MAR assumes that the probability of missing is related to the observed data but not to the unobserved data. The MAR is the most commonly assumed mechanism in empirical research, and is the fundamental assumption for most of imputation methods. If MAR holds, a variety of methods can be applied to handle the missing data, such as Hot deck method, MI, Full Information Maximum Likelihood (FIML). These methods will be discussed in later sections.

The third type of missing data is missing not at random (MNAR) or nonignorable (NI). It implies that the probability of missing is related to the unobserved value in the missing variable. NI is impossible to be verified unless we obtain the unobserved value or the external information beyond the survey. The current available strategies to deal with NI missing are complicated, and the results are sensitive to the choices of methods (Allison 2012). Although various studies have been introduced and developed to examine the problem of NI, no standard method has been found, and only the Heckman-type modelling may alleviate the NI missing data issue (Grittner et.al 2011).

**The MAR assumption**

We assume that the missing data in our survey satisfies the MAR mechanism. First, the MCAR is an inefficient assumption in empirical research, because the MCAR mechanism rarely exists in empirical surveys. Even if the MCAR assumption is satisfied, imputation method based on the MAR mechanism will not bias the analysis (Little and Rubin 1989). In addition, we conducted a primary test to examine whether nonresponses are related to the observed variables. That is to create a 0-1 indicator variable indicating the observations having missing data; then treat the indicator variable as dependent variable and treat all observed variables as independent variables. Then the method uses a logistic model to estimate whether there is any correlation between the nonresponse indicator and observed explanatory variables. The results show that nonresponses are correlated with several observed variables in the survey, so the MCAR assumption fails.

Second, as we concluded above, the MNAR assumption cannot be justified unless we obtain the unobserved value. Even if we can tell the pattern of missing following MNAR, we cannot test the performances of those methods for the MNAR since the missing data are not observable. Also, the results may be significantly different depending on the correction methods. A simple and plausible method to handle the MNAR is to still use the imputation method under the MAR assumption, but include as many predictor variables as possible (Miyama and Managi 2014). The underlying idea is that the more predictor variables we use, the more possible the missing data are correlated with the predictor variables, thus the more likely missing mechanism converts to the MAR from the MNAR. Following the empirical studies in health, medical, environmental and household areas, we assume MAR applies in our research (Van Buuren, Boshuizen and Knook 1999; Schenker et al 2006; Burgette and Reiter 2010; Azur, Stuart, Frangakis and Leaf 2011; White, Royston and Wood 2011; Miyama and Managi 2014).

Given the MAR assumption, MI is one of the most promising methods for dealing with missing data issues, and is outlined in the following steps (van Buuren, Boshuizen, and Knook 1999):

1. Specify the missing variables, the posterior predictive density, and predictor variables given MAR assumption.
2. Draw *m* plausible values for the missing data from the density to generate *m* complete datasets.
3. Conduct *m* complete-data analyses for each of the *m* complete datasets.

4. Combine the *m* data analyses into one estimate with final point and variance estimates.

Rubin (1976) firstly introduce the MI to analyze the nonresponse issue in survey data, and provide the basic reference for MI (Rubin 1987). Numerous statisticians have worked to improve the method (King, Little, Meng, Raghunathan, Rubin, Schafer, Schenker, and van Buurenand), and the MI is a popular choice to tackle missing data in the medical and social sciences in the last two decades. An advantage of the MI is that it considers the true variance of data, because missing values are imputed with different plausible values and are averaged to conclude a final estimate. In this research, we apply the MI method using multivariate imputation by chain equation (MICE) algorithms, introduced by van Buuren, Boshuizen and Knook (1999) and Raghunathan et al. (2001), to impute categorical variables and continuous variables simultaneously without the multivariate normal assumption. The MICE will be introduced in the section of the empirical strategy for missing data.

## Empirical Strategy for Missing Response

Given the MAR assumption, we treat missing data in six scenarios with respect to each BMP practice discussed in the survey. The six scenarios are as followed:

(1) Listwise deletion method: deleting observations with missing responses in the analysis.

(2) Mean imputation method: Replacing missing values of "how-much" by the observed mean.

(3) Using MI to address missing data in follow-up questions: In this case, we only address missing responses in the follow-up questions when respondents answered "yes" to the choice questions but failed to answer the follow-up question. We applied a multivariate MI method referred to as the Multivariate Imputation by Chained Equation (MICE) to replace missing data in the follow-up question.

(4) MI in two-stage: In this case, we consider the issues of missing data both in the choice questions and in the follow-up questions. We first impute missing responses in the choice

questions with “yes”, “no” and “not possible for me”. Then we restrict to respondents who answered or were imputed as “yes”, and impute missing observations of the follow-up questions. The imputation processes for the two missing variables were simultaneous.

(5) MI in two-stage with restriction: This scenario is similar to scenario four but it assumes that missing choices are more likely to be “no” or “not possible for me”. Therefore, we first impute missing choices with “no” and “not possible for me” only, and then impute the missing data in the follow-up questions using the MICE method.

(6) MI in three-stage: In this case, we consider the outcome of “not possible for me” was not the respondents’ preferences but the reality to use a BMP. Under this consideration, we first identified whether it was possible for the respondents who did not answer the choice questions; then for the “possible” group, we impute missing choices with “yes” or “no”; finally for respondents who answered or were imputed as “yes”, we impute missing data in the follow-up questions. These steps were also computed simultaneously using the MICE method.

In the last four scenarios, the basic idea is to decompose the multivariate problem into a series of univariate problems using an iteration algorithm. The procedure is displayed as follows (van Buuren, Boshuizen and Knook 1999; Raghunathan et al. 2001; Schenker et al. 2006; Azur et al. 2011), and is demonstrated in Figure 2 :

1. Let $X$ denote variables fully observed, and $Y^{(1)}, Y^{(2)}, \dots, Y^{(n)}$ denote k variables with missing data, ordered by the amount of missing data from the least to the most.
2. In iteration 1, regress observed $Y^{(1)}$on $X$, and impute the missing values of $Y^{(1)}$ using predictive distribution based on the fitted regression. Then, regress $Y^{(2)}$ on X and $Y^{(1)}$including observed value and recent imputed value, and impute the missing values of $Y^{(2)}$. For $Y^{(k)}$, regress $Y^{(k)}$ on X, $Y^{(1)}$, $Y^{(2)}$, … , $Y^{(k-1)}$ where $Y^{(1)}$, $Y^{(2)}$, … , $Y^{(k-1)}$ include observed value and most recent imputed value, then impute $Y^{(k)}$ using predictive

distribution based on the fitted regression of $Y^{(k)}$. Repeat this procedure until all incomplete variables $Y^{(n)}$ are imputed.

3. In iteration 2, the imputation process is repeated in the same manner as round 1, but predictors in each regression include all variables except for the variable to be imputed. To be specific, regress observed $Y^{(1)}$on $X$ $Y^{(2)}$, $Y^{(3)}$, … , $Y^{(n)}$, where $Y^{(2)}$, $Y^{(3)}$, … , $Y^{(n)}$ consist of imputed values in last round and observed value, and re-impute the missing values of $Y^{(1)}$ using predictive distribution based on the fitted regression. Regress $Y^{(2)}$ on X and $Y^{(1)}$, $Y^{(3)}$, … , $Y^{(n)}$including observed value and imputed value, where $Y^{(1)}$ is the most recent imputed value and $Y^{(3)}$, … , $Y^{(n)}$are imputed in last round; and then re-impute the missing values of $Y^{(2)}$. For $Y^{(k)}$, regress $Y^{(k)}$ on X, $Y^{(1)}$, $Y^{(2)}$, … , $Y^{(k-1)}$, $Y^{(k+1)}$, … , $Y^{(n)}$ where $Y^{(1)}$, $Y^{(2)}$, … , $Y^{(k-1)}$ are the most recent imputed value in current iteration and $Y^{(k+1)}$, … , $Y^{(n)}$ are from the imputed value in last iteration; then re-impute $Y^{(k)}$ using predictive distribution based on the fitted regression of $Y^{(k)}$. This procedure is executed $c$ iterations till the equation chains are converged.

The MICE method allows the use of different models in each regression. If $Y^{(k)}$ is a continuous variable, a normal linear regression is a suitable model; if $Y^{(k)}$ is a binary variable, a logistic regression is a preferable model; if $Y^{(k)}$ is a categorical variable with more than two outcomes, a polytomous regression is a proper model; if $Y^{(k)}$ is a count outcome, a Poisson regression is an appropriate model; if $Y^{(k)}$ is mixed, such as semi-continuous outcome, a two-stage model is applied, such as, zero and non-zero is imputed using logistic regression, and conditional on non-zero group, a normal linear regression model is used to impute non-zero values. All of regressions introduced above are employed in this study, and the computation procedure for each type of regression is introduced in Raghunathan et al (2001).

Define $Y^{(1)}, Y^{(2)}, \ldots$ and $Y^{(n)}$ are variables with missing data;
$X$ are fully observed variables in the dataset;
${Y^{(i)}}_{imp(j)}$ is the $i$ th variable with observed data and imputed data in $j$ th iteration.

Chain Equation Iteration 1:

| Dependent variable | Predictor variables | Imputed variable |
|---|---|---|
| $Y^{(1)}$ | $X$ | ${Y^{(1)}}_{imp(1)}$ |
| $Y^{(2)}$ | $X$ ${Y^{(1)}}_{imp(1)}$ | ${Y^{(2)}}_{imp(1)}$ |
| $Y^{(3)}$ | $X$ ${Y^{(1)}}_{imp(1)}$ ${Y^{(2)}}_{imp(1)}$ | ${Y^{(3)}}_{imp(1)}$ |
| | $\cdots\ \cdots$ | |
| $Y^{(n)}$ | $X$ ${Y^{(1)}}_{imp(1)}$ ${Y^{(2)}}_{imp(1)}$ $\cdots\ \cdots$ ${Y^{(n)}}_{imp(1)}$ | ${Y^{(n)}}_{imp(1)}$ |

Chain Equation Iteration 2:

| Dependent variable | Predictor variables | Imputed variable |
|---|---|---|
| $Y^{(1)}$ | $X$ ${Y^{(2)}}_{imp(1)}$ ${Y^{(3)}}_{imp(1)}$ $\cdots\ \cdots$ ${Y^{(n)}}_{imp(1)}$ | ${Y^{(1)}}_{imp(2)}$ |
| $Y^{(2)}$ | $X$ ${Y^{(1)}}_{imp(2)}$ ${Y^{(3)}}_{imp(1)}$ $\cdots\ \cdots$ ${Y^{(n)}}_{imp(1)}$ | ${Y^{(2)}}_{imp(2)}$ |
| $Y^{(3)}$ | $X$ ${Y^{(1)}}_{imp(2)}$ ${Y^{(2)}}_{imp(2)}$ $\cdots\ \cdots$ ${Y^{(n)}}_{imp(1)}$ | ${Y^{(3)}}_{imp(2)}$ |
| | $\cdots\ \cdots$ | |
| $Y^{(n)}$ | $X$ ${Y^{(1)}}_{imp(2)}$ ${Y^{(2)}}_{imp(2)}$ $\cdots\ \cdots$ ${Y^{(n-1)}}_{imp(2)}$ | ${Y^{(n)}}_{imp(2)}$ |

Chain Equation Iteration $j$:

| Dependent variable | Predictor variables | Imputed variable |
|---|---|---|
| | $\cdots\ \cdots$ | |
| $Y^{(i)}$ | $X$ ${Y^{(1)}}_{imp(j)}$ ${Y^{(2)}}_{imp(j)}$ $\cdots$ ${Y^{(i-1)}}_{imp(j)}$ ${Y^{(i+1)}}_{imp(j-1)}$ $\cdots$ ${Y^{(n)}}_{imp(j-1)}$ | ${Y^{(i)}}_{imp(j)}$ |
| | $\cdots\ \cdots$ | |

Figure 2. Demonstration of the Multivariate Imputation by Chained Equation (MICE) method

**Third Scenario**

The third scenario is to impute missing responses to the follow up questions for each practice. For respondents who answer “No” and “Not possible for me”, the missing value is replaced by zero because they will not adopt the BMPs. For the respondents who answer “Yes” but do not indicate how much they would like to adopt, we impute the missing values with respect to five BMPs simultaneously using the MICE algorithm.

For predictor variables X, we follow a general rule that the number of predictors should be as large as possible (van Buuren, Boshuizen and Knook, 1999). One of the reasons is discussed above in that the more predictors are included, the more possible the MAR condition can be satisfied. Another reason is that using all of the information will increase the precision of prediction, and decrease the bias of imputation. The goal of imputation methods is to predict the distribution of missing variable instead of economic interpretation, and the imputations are drawn from the posterior but do not change the joint distribution (Schafer 1997; King, Honaker, Joseph and Scheve, 2001). In addition, imputation algorithms do not require the causality between predictor variables and imputed variables.

However, White, Royston and Wood (2011) states that if the imputation model includes too many variables, the convergence of such large models is an issue, especially for a complex set of imputation models. Buuren and Groothuis-Oudshoorn (2011) conclude that it is appropriate and suitable to include no more than 15 to 25 variables. Based on previous literature, predictor variables in our imputing models include levels of compensation, land size, rent percentage, having surface water on the farm, percentage of household income from farming, total household income reinvested back to farm, types of farming production, age, gender, education, income, race, water recreation activities, farming experiences, water quality near the farm, participation in government programs, current usage of different types of BMPs, as well as different levels of information explaining the meaning of WQT programs in the survey. Except the variable of using waste storage facility is a count number, all imputed variables are continuous, so the imputation model is a normal linear regression model. Since the missing value of using waste storage facilities is measured by a count number, we apply count variable imputation model based on the Poisson regression to impute the missing data.

**Fourth scenario**

The fourth scenario is to impute missing values both in the yes/no question and the question on the amount of BMPs to be adopted with respect to each practice. The possible response to the yes/no questions is discrete, such as “yes”, “no”, or “not possible for me”. The possible response to the amount of adoption questions is the continuous value when respondents answer “yes” to the yes/no question, and zero when the answer were “no” or “not possible for me”. Therefore, we impute “yes”, “no” or “not possible for me” to the missing value in the yes/no question using a multinomial logit imputation model; then restricting the sample to the “yes” group, we impute missing values in the follow-up question by using a normal linear regression model. The imputation steps are outlined as followed, and is also described in Figure 3:

1. The missing value in the yes/no questions are imputed as discrete values such as “yes”, “no”, or “not possible for me” with respect to each BMP.
2. For respondents who answered “no” and “not possible for me” to the yes/no questions, the missing value in the quantity questions are replaced by zero; for respondents who did not answer these questions but were imputed to answer “no” and “not possible for me”, the missing value are also replaced by zero.
3. For respondents who answered “yes” to the yes/no questions but did not answer how much they would like to implement BMPs, and respondents who did not answer “yes” to the yes/no questions but were imputed to answer “yes”, their missing responses in quantity questions are imputed by MI.

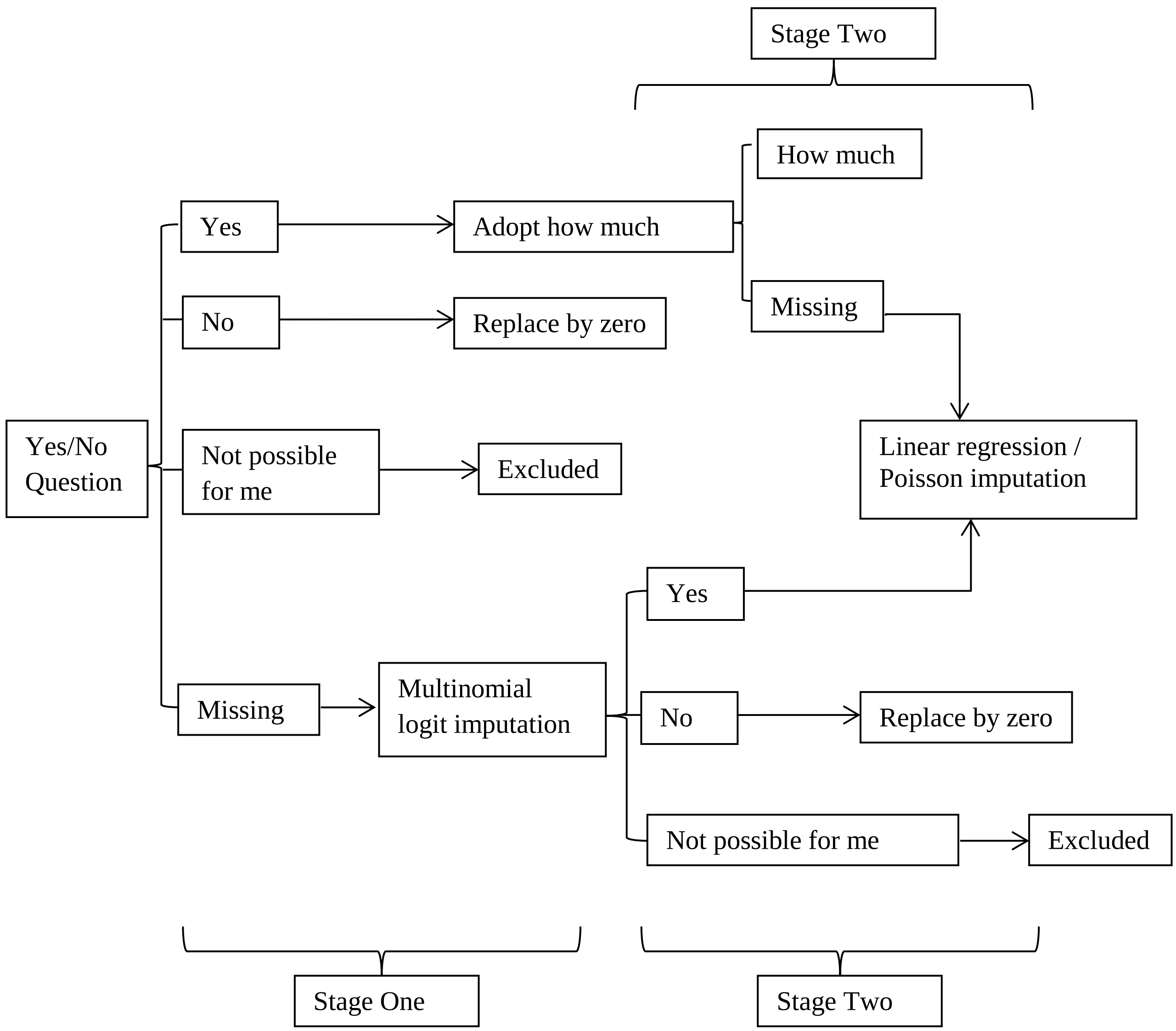


Figure 3. The Fourth Scenario

**Fifth scenario**

The fifth scenario is similar to the fourth scenario that imputes missing values both in the yes/no question and the follow-up question with respect to each practice using a two-stage approach, but we restrict the missing value in the yes/no question as “no” or “not possible for me” only.  In this scenario, we assume that the missing response of yes/no question is more likely to be a “no” or “not possible for me” which is a more plausible and conservative assumption than the last scenario.

Therefore, we first impute the missing value with “no” and “not possible for me” in the yes/no question using a logistic regression model; then restrict the sample to the “yes” group and impute the follow-up question using a linear regression model or Poisson method. The imputation steps are described in Figure 4, and are outlined as followed:

1. The missing value in the yes/no questions is imputed as discrete values such as “no” or “not possible for me” with respect to each BMP using a logistic regression model.
2. For respondents who answered “no” and “not possible for me” to the yes/no questions, the missing value in the quantity questions are replaced by zero; for respondents who did not answer these questions but were imputed to answer “no” and “not possible for me”, the missing value are also replaced by zero.
3. For respondents who answered “yes” to the yes/no questions but did not answer how much they would like to implement BMPs, the missing responses in quantity questions are imputed using linear regression model or Poisson method.

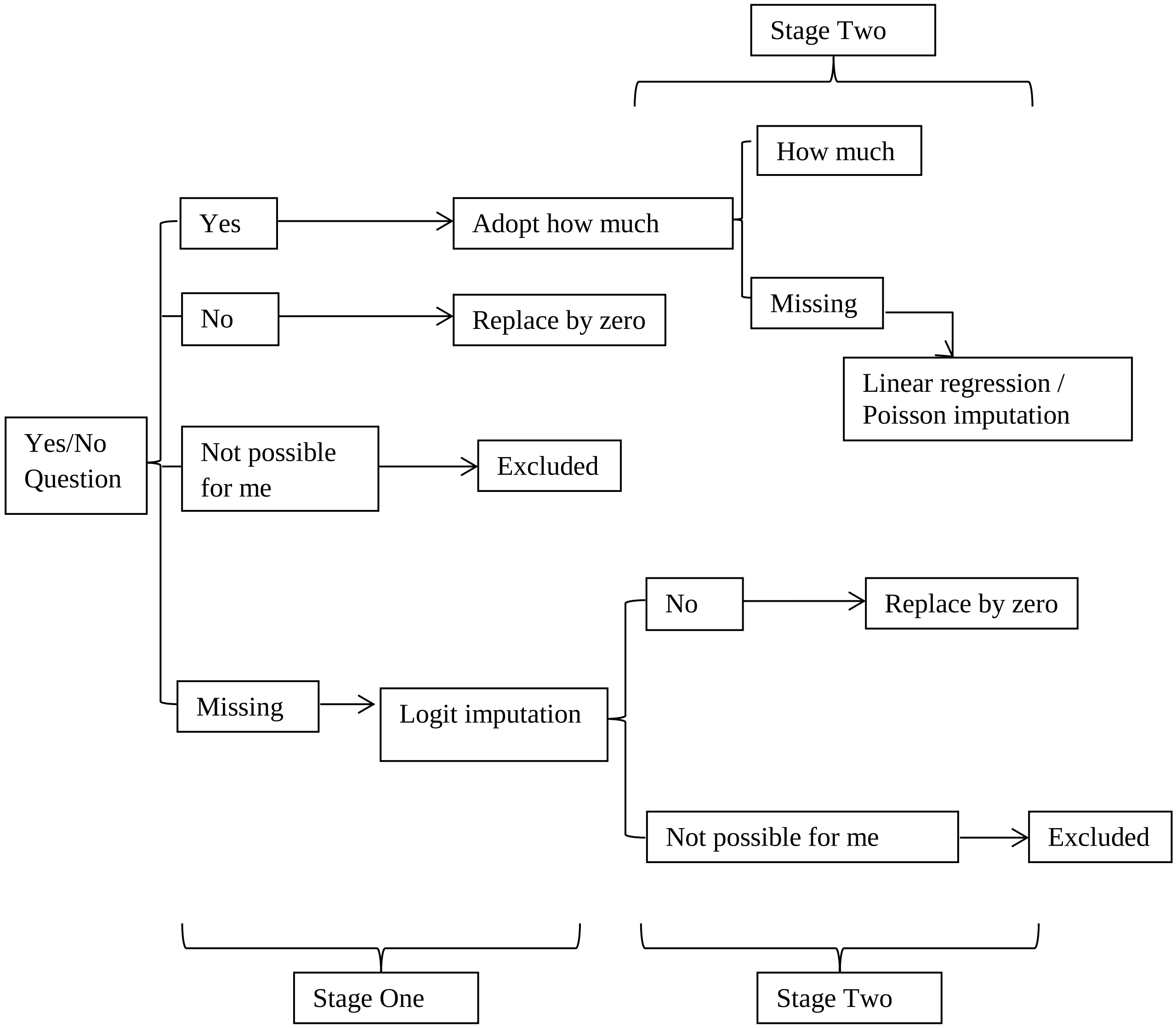


Figure 4. The Fifth Scenario

**Sixth scenario**

The sixth scenario carefully considers the nature of the missing response of the yes/no question. The response “not possible for me” is, in principle, different from the responses of “yes” and “no”. Responses of “yes” and “no” represent personal preference to implement BMPs given the compensation through WQT programs, but the response of “not possible for me” implies whether a farm is able to implement a practice regardless of whether they would like to. If those three responses are treated equally in imputation, the result may be questionable. As a result, this scenario is to firstly determine whether farmers are able to implement BMPs; then, for those who are able to use BMPs, we impute the missing value in the yes/no questions by "yes" or "no". To be specific, we use a logistic regression model to impute the missing response with "possible" and "not possible"; then we focus the sample on those with "possible" group, and impute "yes" and "no" using the logistic regression model again; and for the sample who either answered "yes" initially or were imputed to answer "yes", we impute values in the question on the amount to be adopted. The imputation steps are described in Figure 5, and are outlined as followed:

1. The missing value in the yes/no questions is imputed as discrete values such as "possible" or "not possible for me" with respect to each BMP using a logistic regression model.
2. For respondents who were imputed to be "possible" group, we re-impute the missing responses with "yes" or "no" with respect to each BMP using a logistic regression model.
3. For respondents who answered or were imputed to answer "not possible for me", the missing value in the quantity questions are replaced by zero;

   For respondents who answered or were imputed to answer "no", the missing values in the follow-up questions are replaced by zero;
4. For respondents who answered "yes" to the yes/no questions but did not answer how much they would like to implement BMPs, and the respondents who were imputed to answer "yes", the missing responses in the following up questions are imputed using linear regression model or Poisson method.

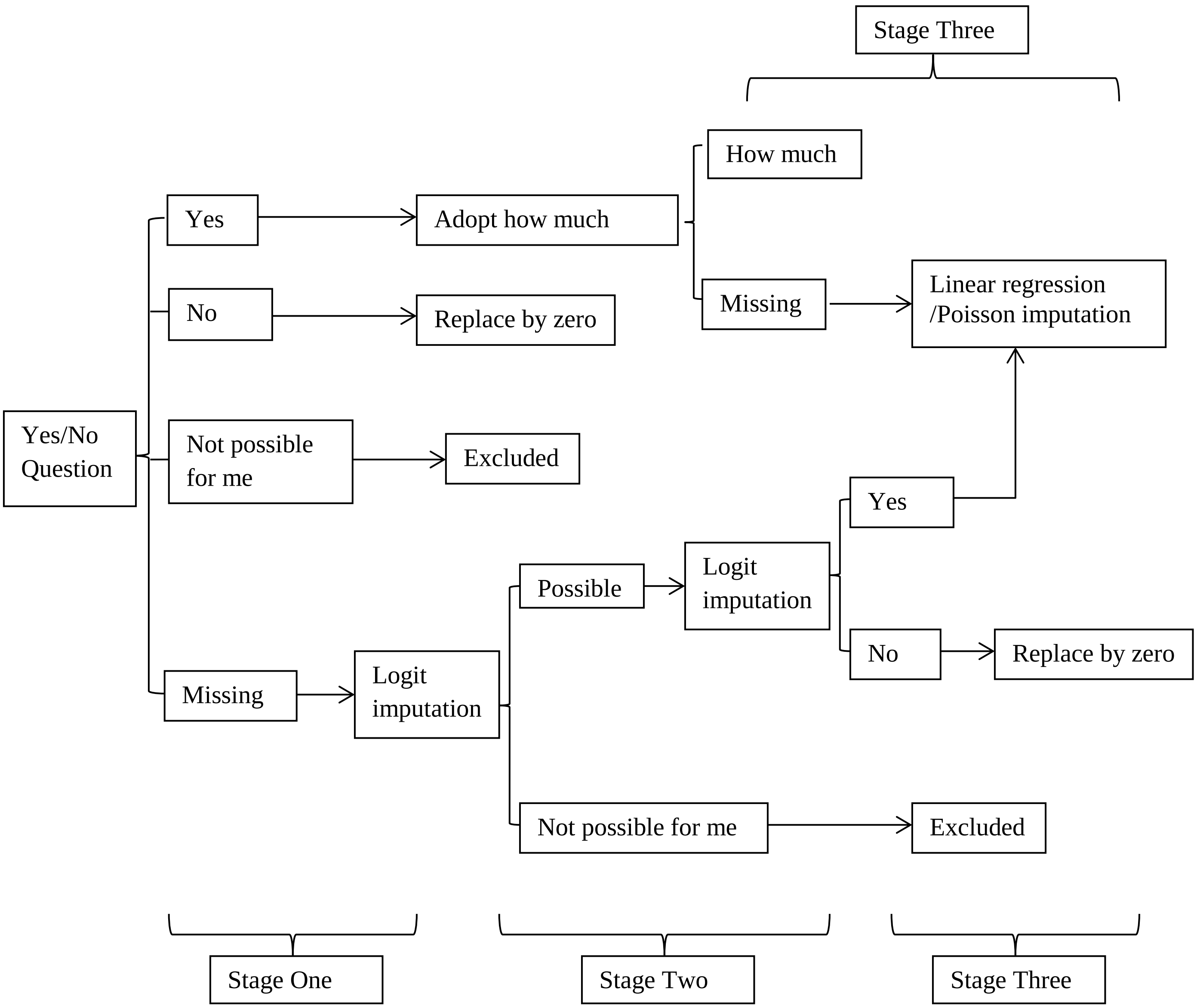


Figure 5. The Six Scenario

## Imputation

### Fitting the imputation model

During the imputation, the normal linear regression model requires the normality assumption for observed value $y_{obs}$ conditional on predicting value X. When the observed values are highly skewed with a relative sample size, normal linear regression model will be invalid. Following Royston and White (2011) paper, we apply a shifted log transformation to the observed value of missing data, $y_{obs}$, in order to satisfy the normality assumption. That is to transform the observed value of missing data, $y_{obs}$, into a log form toward normality using equation (1) where $y_{norm}$ is the log-transformed observed value, $y$ is the observed value of missing data $y_{obs}$, and $k$ is an estimated parameter indicating the skewness. After imputation, we use the inverse transformation equation (2) to compute imputed values of y back to the original scale.

$$y_{norm} = \ln(y - k) \qquad (1)$$

$$y = e^{y_{norm}} - k \qquad (2)$$

Besides, using Poisson model to impute missing data fails to achieve convergence in our convergence checking process. We apply a predictive mean matching (PMM) method that imputes values from the observed values of variable by matching predicted values as closely as possible (Little, 1988), and achieve convergence.

In our research, perfect prediction issues occur in several models. Perfect prediction arises when covariate variables can perfectly predict outcomes of the categorical data (Albert and Anderson, 1984). As a result, the imputation cannot be executed because the estimation has infinite coefficients with infinite standard errors. The imputation with categorical data is more likely to have the perfect prediction issue (White, Daniel, and Royston, 2011), especially for the logit and multinomial logit imputation model. One strategy eliminating the prefect prediction is to diagnose models, and identify and remove the concerning covariate. However, removing a troublesome variable is questionable and may potentially mislead the imputation, because omitting a key determinant leads to a biased result. The alternative strategy is to use the augment approach introduced by White, Daniel, and Royston, (2010). The augment approach is to add several extra observations with weights during estimation. We apply the augment approach in all imputation models with categorical data.

As introduced before, the MICE method is a chained equation method using an iterative process. Before the final value is imputed, iterative equations need to execute burn-in period to converge to a stationary state. Therefore, it is necessary to check the length of the burn-in period to ensure that the MICE algorithm has converged before the value is imputed. White, Royston and Wood (2011) and Van Buuren and Groothuis-Oudshoorn (2011) suggest that ten cycles of iteration are adequate, though there is no harm to conduct some extra iterations to assess the convergence of our imputation models. We conduct 1000 times of iteration to check the convergence for the last four scenarios before the imputation, and find that 10-30 times of burn-in period is sufficient to achieve convergence. As a result, we execute 30 times as burn-in iterations for each imputation in the last four scenarios.

**After the imputation**

After the imputation, we use the upper and lower bound to replace the abnormal value imputed using linear regression model. With respect to each BMP, the upper bound and lower bound are observed maximum and minimum values respectively. Abnormal values that exceed the upper and lower bound account for less than 5% of all of imputation sets across last four scenarios.

For each imputed dataset, we estimate the factors affecting how much farmers may engage in BMPs with respect to each practice. The estimation is specified by equation (3) using Tobit or Poisson regressions. The dependent variables $Y_i$ are how much farmers would like to implement the BMPs. Subscript $i$ denotes different types of BMPs. For the practices of Riparian buffers $Y_1$, Animal fences $Y_2$, No till $Y_3$ and Nutrient management $Y_5$, dependent variables $Y_i$ are continuous if the decision is "yes", and take the value zero if the decision is "no". Due to the fact that the usage of BMPs is censored at zero, we use Tobit model to estimate how much farmers may implement these practices. In addition, since the dependent variable $Y_4$ is a count value, we estimate how many units of waste storage facilities may be adopted using Poisson regression.

$$Y_i = X'\beta + \varepsilon \qquad (3)$$

$$where\ Y_i = continous\ value\ if\ decision\ is\ \text{Yes}, i = 1,2,3,5$$

$$Y_i = count\ number\ if\ decision\ is\ \text{Yes}, i = 4$$

$$Y_i = 0 \; if \; decsion \; is \; "No"$$

Same explanatory variables are used to explain the usage of all five BMPs, and those variables are: compensation, land acre, rent area percentage, whether having surface water on the farm, percentage of household income from farming, total household income reinvested back to farm, income, water quality, participation in the Conservation Reserve Program (CRP) and Working-Land Program (WLP). We examine the cross-effect of adopting BMPs through including the current usage of the five types of BMPs. Finally, for the action to adopt additional BMPs, we also allow the decision to adopt one practice to explain the adoption of the others to examine whether there is synergy between using BMPs in the future.

The last step of the MI is to integrate the *m* results of estimation using Rubin's method (Rubin 1987). Let $Q$ denote a parameter that need to be estimated, such as a regression coefficient, in each imputed dataset. The point estimate $\overline{Q}$ of $Q$ is the average of the m spate estimate, and is represent by equation (4)

$$\overline{Q} = \frac{1}{m}\sum_{j=1}^{m} Q_j \qquad (4)$$

Let $U_j$ denote the estimated squared standard error of $Q_j$ written as equation (5), and $B$ denote the between-imputation variance across the *m* point estimates written as equation (6). So the estimated variance of point estimate of MI, $T$, is represented by equation (7).

$$\overline{U} = \frac{1}{m}\sum_{j=1}^{m} U_j \qquad (5)$$

$$B = \frac{1}{m-1}\sum_{j=1}^{m}\left(Q_j - \overline{Q}\right)^2 \; (6)$$

$$T = \left(1 + \frac{1}{m}\right)B + \overline{U} \qquad (7)$$

The tests and confidence intervals follow a Student's t-approximation $(\overline{Q} - Q)/\sqrt{T} \sim t_v$ with degrees of freedom $v$ represented as equation (8).

$$v = \left(\frac{1}{m-1}\right)\left[1 + \frac{\overline{U}}{(1+m^{-1})B}\right] \quad (8)$$

Previous studies show that five or ten imputations are sufficient unless the degree of missing data is high. However, White, Royston and Wood (2011) argue that larger numbers of imputation time *m* are preferred due to the efficiency loss and reproducibility. Since the variance of parameters is calculated using equation (7), they propose that the relative efficiency of infinitely many imputations compared to *m* imputations is

$$\lim_{n\to\infty} \frac{\left(1+\frac{1}{m}\right)B + \overline{U}}{\left(1+\frac{1}{n}\right)B + \overline{U}} = \frac{\left(1+\frac{1}{m}\right)B + \overline{U}}{B + \overline{U}} = 1 + \frac{B}{B+\overline{U}} * \frac{1}{m} = 1 + \frac{FMI}{m}$$

Where $\frac{B}{B+\overline{U}}$ is the fraction of missing information (FMI) introduced by Schafer (1997). If we use 5% loss of efficiency in our imputation, $1 + \frac{FMI}{m}$ should be less or equal to 1.05, so $\frac{FMI}{m} \leq 0.05$. FMI is calculated after the analytic model using imputation data, and can be obtained from most of statistical software packages (STATA or R). In our research, we can accept 1% of loss of efficiency, so the imputation times *m* are greater or equal to $(100 * FMI)$. In the estimation, each parameter has its own FMI, so usually there are several FMI values from an analysis. In our research, we believe that the largest FMI value determines the imputation times *m*. Besides, we also want to reproduce our imputation to obtain a robust analysis. Intuitively, the larger *m* is, the less different results we would have when we reproduce the estimation, and the more confident results we would have. After the primary imputation, we accept that *m*=100 times is a reasonable imputation numbers to obtain a robust result, and have conducted 100 times of imputation for each scenario.

## Result

The imputation procedure is executed using Stata 12.0 “rseed” option for reproducing results. The models are estimated excluding responses who answered “not possible for me” with respect to each practice. Besides, variables of percentage of household income from farming, total household income reinvested back to farm, and income level are categorical values, and coefficients of those variables are impossible to be explained. In order to extend the implication of results, those variables are rescaled to the exact percentage value or dollar value before the estimation, and are specified in the Table 8. Table 9-13 display the results of Tobit or Poisson model estimating the factors affecting how much farmers may engage in BMPs with respect to each practice. Each table compares the results of six scenarios for a BMP. The largest FMI values for each model are reported at the bottom of respective tables.

With respect to each BMP, the results show that significance of coefficients is mostly consistent through six scenarios, but the magnitudes vary within a certain range. The results of the deletion method and the mean imputation method are relatively closer to each other than the other four scenarios, but for the statistical significant coefficients, the absolute values of coefficients using the mean imputation are statistically smaller than the ones using the deletion method. This is because replacing the missing value by a constant value will decrease the variability of data, and centralize the distribution of the data. As a result, the mean imputation may potentially distort the efficiency of the estimation, and even lead to underestimated result.

The results of scenarios three, four and five; i.e., the one-stage imputation, the two-stage imputation and the two-stage imputation with restriction, are mostly consistent with each other; but results using the three-stage imputation, the sixth scenario, are not close to any other scenarios. The third, fourth, and fifth scenarios impute missing responses in the yes/no question and the follow-up question directly using the MICE algorithm. The sixth scenario considers the agrarian or geographic issue whether it is feasible for farmers to use BMPs on their farms, so we introduce one more stage to impute whether observations with missing responses are possible to adopt BMPs on their land. On the one hand, the sixth scenario considers the nature of the responses between the respondent's preference and the agrarian issue, and selects the "possible group" carefully and strictly following the theory and procedure; on the other hand, the extra step to determine whether a farm is able to adopt BMPs is also a strong assumption. If variables explaining the "possible group" are not selected carefully, the imputation model will not predict the missing value precisely. As a result, final results may potentially be biased because the extra step is more likely to distort the true distribution of the data. Although we have carefully selected the covariates explaining the "possible group" and conducted convergence check, the three-stage imputation may be less reliable than the other imputation strategies unless we use the extra geographical information to identify the "possible group".

**Additional abatement**

In this analysis, we use results from the third scenario to examine whether farmers have additional ability to reduce agricultural runoff, what the factors affect the ability, and by how much they would like to adopt additional BMPs on their land to generate trading credits. As we introduced before, we suppose that the missing mechanism follows MAR assumption, so the MI method is a more promising and reliable method than the deletion and mean imputation method. In addition, the mechanism of missing responses of Yes/No choices is not clearly determined, but it is straightforward and intuitive to impute the follow-up question using the MICE algorithm based on the linear normal regression. As a result, it is conservative to select results in the third scenario, the one-stage imputation, to conduct analysis.

Table 9-13 show the results of the factors affecting how much farmers may adopt BMPs after using the MICE method to impute missing responses of the follow-up question. Holding other variables constant, increasing 1% of cost coverage for using BMPs will increase the adoption of riparian buffers by 58.44 feet; one more acre on the farm is predicted to decrease the adoption of animal fences by 2.77 feet. On the one hand, if farmers receive more revenue from farms, they are less likely to adopt no till, and results show that increasing 1% of household income from farming will lead farmers to adopt more no till on their land by 1.64 acre. On the other hand, if farmers prefer to invest more assets on their farms, they tend to adopt more riparian buffers and waste storage facilities. Increasing 1% of household income from farming will increase the adoption of riparian buffers by 47.07 feet.

One of the most important findings is that the previous experience of BMPs will significantly affect farmers' adoption of BMPs. Holding other factors constant, if farmers are currently using riparian buffers, they will adopt additional 2831.04 feet of riparian buffers, or 102.53 acres of nutrient management than farmers without using the practice; if farmers have already built up the animal fences on their land, they will use additional animal fences by 1866.46 acres; if farmers currently adopt no till on their land, they are likely to adopt additional animal fences by 2499.37 feet and no till by 129.75 acres; if farmers have already built up the waste storage facility on their land, they are probably to reduce no till by 121.23 acres; if farmers are currently using nutrient management on their farms, they will reduce animal fences by 2618.82 feet, but would like to adopt additional nutrient management by141.46 acres. The results are consistent across six scenarios.

In addition, the results also show that the WQT information featured in the cost saving aspect will encourage farmers to adopt additional practices of animal fences by 1576.89 feet; but the information featured in the environmental aspect will not influence BMPs adoption.

Table 9-13 also show synergy of BMP adoption that certain sets of BMPs often/almost are always practiced together. If farmers would like to use riparian buffers, they are more likely to adopt animal fences, and vice versa. If farmers would like to build up waste storage facilities, they are more likely to implement nutrient management through the WQT program as well, and again vice versa.

## Conclusion

This study explores whether farmers in Kentucky would like to reduce agricultural runoff by adopting additional BMPs, and what factors affect the decision. The study also explains by how much farmers will adopt additional BMPs based on the different levels of compensation generated by the WQT programs. In our survey, about 21.5%, 26.9%, 24.2%, 23.2%, and 18.2% of respondents did not indicate how much they may adopt additional BMPs with respect to five types of investigated in the survey. Therefore, we apply three treatments to address the missing data issues in our study, which include six specific approaches. These approaches are: the deletion method, the mean imputation method, and the one-stage method using MICE, the two-stage method using MICE, the two-stage with restriction method using MICE, and the three-stage method using MICE. Given those treatments, we estimate the factors affecting how much farmers may engage their lands in BMPs using Tobit model and Poisson model.

Our findings show that increasing 1% of the cost coverage for using BMPs is more likely to encourage farmers to adopt additional riparian buffers by 58.44 feet. In addition, land area, percentage of household income from farming, percentage of total household income reinvested back to farm, and current experience of BMPs will affect BMPs adoption. In the end, we observe the synergy of BMP adoption that riparian buffers and animal fences, and waste storage facilities and nutrient management are always adopted together.

Although MI method has been introduced more than 20 years, and become an established method in political science, medical science and behavior science, most of researchers still rely on the deletion method for missing data in surveys of famers. One of the implications from this study is that the MI method may offer a promising way to handle missing data in farmers' decisions. Our research does not intend to offer a normative strategy while dealing with missing data. We are interested in providing a comparison between several popular schemes to address the issue of missing data. Our conclusion is that a conservative strategy to deal with missing data is to provide both the deletion method and the MI method for reference in the analysis. The mean imputation method is not recommended as it may not generate results as reliable as the other methods while the researcher is uncertain about the missing mechanism.

# Reference


Albert, A., and J. Anderson. 1984. "On the existence of maximum likelihood estimates in logistic regression models." *Biometrika* 71:1-10.

Allison, P. 2012. "Handling missing data by maximum likelihood." *Unpublished manuscript. Retrieved from http://www. statisticalhorizons. com/wp-content/uploads/MissingDataByML. pdf*.

Andridge, R.R., and R.J. Little. 2010. "A Review of Hot Deck Imputation for Survey Non-response." *International Statistical Review* 78:40-64.

Azur, M.J., E.A. Stuart, C. Frangakis, and P.J. Leaf. 2011. "Multiple imputation by chained equations: what is it and how does it work?" *Int J Methods Psychiatr Res* 20:40-49.

Burgette, L.F., and J.P. Reiter. 2010. "Multiple imputation for missing data via sequential regression trees." *Am J Epidemiol* 172:1070-1076.

Buuren, S., and K. Groothuis-Oudshoorn. 2011. "MICE: Multivariate imputation by chained equations in R." *Journal of statistical software* 45.

Dempster, A.P., N.M. Laird, and D.B. Rubin. 1977. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)*:1-38.

Grittner, U., G. Gmel, S. Ripatti, K. Bloomfield, and M. Wicki. 2011. "Missing value imputation in longitudinal measures of alcohol consumption." *Int J Methods Psychiatr Res* 20:50-61.

Groves, R.M. 2006. "Nonresponse rates and nonresponse bias in household surveys." *Public Opinion Quarterly* 70:646-675.

Groves, R.M., and E. Peytcheva. 2008. "The impact of nonresponse rates on nonresponse bias a meta-analysis." *Public Opinion Quarterly* 72:167-189.

Horton, N.J., and K.P. Kleinman. 2007. "Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models." *Am Stat* 61:79-90.

Kenward, M.G., and J. Carpenter. 2007. "Multiple imputation: current perspectives." *Statistical Methods in Medical Research* 16:199-218.

King, G., J. Honaker, A. Joseph, and K. Scheve (2001) "Analyzing incomplete political science data: An alternative algorithm for multiple imputation." In *American Political Science Association.* Cambridge Univ Press, pp. 49-69.

Lin, I.-F., and N.C. Schaeffer. 1995. "Using survey participants to estimate the impact of nonparticipation." *Public Opinion Quarterly* 59:236-258.

Little, R.J., and D.B. Rubin. 1989. "The analysis of social science data with missing values." *Sociological Methods & Research* 18:292-326.

Miyama, E., and S. Managi. 2014. "Global environmental emissions estimate: application of multiple imputation." *Environmental Economics and Policy Studies* 16:115-135.

Raghunathan, T.E., J.M. Lepkowski, J. Van Hoewyk, and P. Solenberger. 2001. "A multivariate technique for multiply imputing missing values using a sequence of regression models." *Survey methodology* 27:85-96.

Ribaudo, M.O., and J. Gottlieb. 2011. "Point-Nonpoint Trading - Can It Work?1." *JAWRA Journal of the American Water Resources Association* 47:5-14.

Royston, P. 2004. "Multiple imputation of missing values." *Stata Journal* 4:227-241.

Royston, P., and I.R. White. 2011. "Multiple imputation by chained equations (MICE): implementation in Stata." *Journal of statistical software* 45:1-20.

Rubin, D.B. 1976. "Inference and missing data." *Biometrika* 63:581-592.

---. 1987. *Multiple imputation for nonresponse in surveys*: John Wiley & Sons.

Schafer, J.L. 1999. "Multiple imputation: a primer." *Statistical Methods in Medical Research* 8:3-15.

Schenker, N., T.E. Raghunathan, P.-L. Chiu, D.M. Makuc, G. Zhang, and A.J. Cohen. 2006. "Multiple Imputation of Missing Income Data in the National Health Interview Survey." *Journal of the American Statistical Association* 101:924-933.

Shortle, J. 2013. "Economics and Environmental Markets: Lessons from Water-Quality Trading." *Agricultural and Resource Economics Review* 42:57-74.

Van Buuren, S., and K. Oudshoorn. 1999. "Flexible multivariate imputation by MICE." *Leiden, The Netherlands: TNO Prevention Center*.

Weber, J.G., and D.M. Clay. 2013. "Who does not Respond to the Agricultural Resource Management Survey and does it Matter?" *American Journal of Agricultural Economics*:aas171.

White, I.R., P. Royston, and A.M. Wood. 2011. "Multiple imputation using chained equations: Issues and guidance for practice." *Stat Med* 30:377-399.

# Table

**Table 1. Variable Summary Statistics (N=357)**

| Variable | Definition of Variables | Mean | Std. Dev. |
|---|---|---|---|
| **Current BMPs adoption:** | | | |
| $y_1$ | Currently using any BMPs (=1); otherwise (=0) | 0.739 | 0.44 |
| $y_2$ | Currently using riparian buffers (=1); otherwise (=0) | 0.367 | 0.483 |
| $y_3$ | Currently using animal fences (=1); otherwise (=0) | 0.465 | 0.499 |
| $y_4$ | Currently using no-till (=1) ; otherwise (=0) | 0.311 | 0.464 |
| $y_5$ | Currently using waste storage facilities (=1) ; otherwise (=0) | 0.067 | 0.251 |
| $y_6$ | Currently using nutrient management (=1) ; otherwise (=0) | 0.241 | 0.428 |
| **Cost coverage compensation:** | | | |
| Offer | The percentage that treatment plant or factory will cover the cost of implementing the BMPs if the farmer uses the additional BMPs, there are ten different levels of compensation. Those levels are 75%, 80%, 85%, 90%, 95%, 100%, 105%, 110%, 115% and 120%. | 0.97 | 0.15 |
| **Explanatory variables:** | | | |
| Land size | Land size includes rented and owned land for operating. (unit: 1000 acre) | 0.282 | 0.537 |
| Rent percent | Rented land for operating / Total land for operating | 0.142 | 0.275 |
| Surface water | Surface water on farmland (=1) ; otherwise (=0) | 0.86 | 0.348 |
| Percentage of household income from farming | Share of pre-tax household income from farming (see table 4) | 2.417 | 1.815 |
| Total household income reinvested back to farm | Share of pre-tax household income back to farming (see table 4) | 2.529 | 1.542 |
| Farms with crop | Farms earning revenue from crop or farmers planting crop on their land (=1) ; otherwise (=0) | 0.423 | 0.495 |
| Farms with livestock | Farms earning revenue from livestock or raising livestock (=1) ; otherwise (=0) | 0.798 | 0.402 |
| Age | Farmer's age | 60.154 | 11.908 |
| Male | Male =1; otherwise (=0) | 0.857 | 0.35 |
| Education | Farmer's education level (see table 4) | 4.078 | 1.92 |
| Income level | Household annual pre-tax income level (see table 4) | 4.359 | 1.499 |
| Farming experience | Farming experience (year) | 32.22 | 15.307 |

(Continued)

**Table 1. Continued**

| Variable | Definition of Variables | Mean | Std. Dev. |
|---|---|---|---|
| Water recreation | Participating in water related recreation at least once a year (=1) ; otherwise (=0) | 0.661 | 0.474 |
| CRP | Currently participating in Conservation Reserve Program (CRP) (=1) ; otherwise (=0) | 0.118 | 0.323 |
| WLP | Currently participating in Working-Land Program (WLP) (=1); otherwise (=0). WLP includes Conservation Stewardship Program (CSP), Environmental Quality Incentives Program (EQIP), Wildlife Habitat Incentives Program (WHIP) | 0.204 | 0.404 |
| Water quality | Discrete levels from 1 to 7 indicating the poorest to the best water quality nearest to farmers' properties | 5.038 | 1.365 |
| Concern of environmental issue | Respondents' awareness of issues concerning the environment Self-rated with seven levels. Level seven is very aware, and level one is unaware. | 4.947 | 1.556 |
| **Target farmers:** | | | |
| Beginning farmers | Farming less than ten years (=1) ; otherwise (=0) | 0.12 | 0.326 |
| Socially disadvantage farmers (Non-white) | Operator's race is not white (=1) ; otherwise (=0) | 0.045 | 0.207 |
| **Infeasible to implement BMPs** | | | |
| $z_1$ | Answer "not possible for me" to all BMPs (=1) ; otherwise (=0) | 0.345 | 0.476 |
| $z_2$ | Answer "not possible for me" to riparian buffers (=1) ; otherwise (=0) | 0.583 | 0.494 |
| $z_3$ | Answer "not possible for me" to animal fences (=1) ; otherwise (=0) | 0.49 | 0.501 |
| $z_4$ | Answer "not possible for me" to no-till (=1) ; otherwise (=0) | 0.501 | 0.501 |
| $z_5$ | Answer "not possible for me" to waste storage facilities (=1) ; otherwise (=0) | 0.577 | 0.495 |
| $z_6$ | Answer "not possible for me" to nutrient management (=1) ; otherwise (=0) | 0.507 | 0.501 |
| **Information:** The survey was designed with 4 levels of information explaining the meaning of WQT programs | | | |
| Level 1 | The least detailed information level (=1); otherwise (=0) | 0.235 | 0.425 |
| Level 2 | The less detailed information level(=1); otherwise (=0) | 0.261 | 0.44 |
| Level 3 | The more detailed information level(=1); otherwise (=0) | 0.21 | 0.408 |
| Level 4 | The least detailed information level(=1); otherwise (=0) | 0.294 | 0.456 |

Note: Discrete levels in table are interpreted in table 4.

**Table 2. Frequency Distribution of Discrete Variables**

| Level | Percentage of household income from farming | Frequency | Percent |
|---|---|---|---|
| 1 | 0-15% | 162 | 45.38% |
| 2 | 16-30% | 77 | 21.57% |
| 3 | 31-45% | 36 | 10.08% |
| 4 | 46-60% | 28 | 7.84% |
| 5 | 61-75% | 17 | 4.76% |
| 6 | 75-90% | 17 | 4.76% |
| 7 | above 90% | 20 | 5.6% |
| Level | Total household income reinvested back to farm | Frequency | Percent |
| 1 | 0-15% | 106 | 29.69% |
| 2 | 16-30% | 116 | 32.49% |
| 3 | 31-45% | 48 | 13.45% |
| 4 | 46-60% | 45 | 12.61% |
| 5 | 61-75% | 20 | 5.6% |
| 6 | 75-90% | 13 | 3.64% |
| 7 | above 90% | 9 | 2.52% |
| Level | Income ($) | Frequency | Percent |
| 1 | 0 to 14999 | 14 | 3.92% |
| 2 | 15000 to 24999 | 21 | 5.88% |
| 3 | 25000 to 49999 | 60 | 16.81% |
| 4 | 50000 to 74999 | 110 | 30.81% |
| 5 | 75000 to 99999 | 64 | 17.93% |
| 6 | 100000 to 149999 | 56 | 15.69% |
| 7 | above 150000 | 32 | 8.96% |
| Level | Education | Frequency | Percent |
| 1 | Not a high school graduate | 17 | 4.76% |
| 2 | High school graduate | 88 | 24.65% |
| 3 | Some college, no degree | 64 | 17.93% |
| 4 | Associate degree | 14 | 3.92% |
| 5 | Bachelor degree | 83 | 23.25% |
| 6 | Master degree | 51 | 14.29% |
| 7 | Professional degree | 26 | 7.28% |
| 8 | Doctorate | 14 | 3.92% |

**Table 3 Response Frequency of Willingness to Adopt Riparian buffers**

| Category | Options | Freq. | Quantitative questions | Freq. | Mean |
|---|---|---|---|---|---|
| Answer the question | yes | 69 | Say yes and answer the amount | 37 | 839.32 |
| | | | Say yes without answering the amount | 32 | |
| | no | 80 | | | |
| | not possible for me | 70 | | | |
| Missing | | 138 | | | |
| Total | | 357 | | | |

**Table 4 Response Frequency of Willingness to Adopt Animal fences**

| Category | Options | Freq. | Quantitative questions | Freq. | Mean |
|---|---|---|---|---|---|
| Answer the question | yes | 120 | Say yes and answer the amount | 71 | 1531.014 |
| | | | Say yes without answering the amount | 49 | |
| | no | 62 | | | |
| | not possible for me | 60 | | | |
| Missing | | 115 | | | |
| Total | | 357 | | | |

**Table 5 Response Frequency of Willingness to Adopt No till**

| Category | Options | Freq. | Quantitative questions | Freq. | Mean |
|---|---|---|---|---|---|
| Answer the question | yes | 111 | Say yes and answer the amount | 68 | 71.33088 |
| | | | Say yes without answering the amount | 43 | |
| | no | 67 | | | |
| | not possible for me | 49 | | | |
| Missing | | 130 | | | |
| Total | | 357 | | | |

**Table 6 Response Frequency of Willingness to Adopt Waste storage facility**

| Category | Options | Freq. | Quantitative questions | Freq. | Mean |
|---|---|---|---|---|---|
| Answer the question | yes | 70 | Say yes and answer the amount | 45 | 1.42 |
| | | | Say yes without answering the amount | 25 | |
| | no | 81 | | | |
| | not possible for me | 69 | | | |
| Missing | | 137 | | | |
| Total | | 357 | | | |

**Table 7 Response Frequency of Willingness to Adopt Nutrient management**

| Category | Options | Freq. | Quantitative questions | Freq. | Mean |
|---|---|---|---|---|---|
| Answer the question | yes | 110 | Say yes and answer the amount | 78 | 98.75 |
| | | | Say yes without answering the amount | 32 | |
| | no | 66 | | | |
| | not possible for me | 38 | | | |
| Missing | | 143 | | | |
| Total | | 357 | | | |

**Table 8. Rescaled Categorical Variables**

| Categorical value | Rescaled value<br>Percentage of household income from farming | Total household income reinvested back to farm | Income (1000 dollars) |
|---|---|---|---|
| 1 | 8% | 8% | 0.5 |
| 2 | 23% | 23% | 20 |
| 3 | 38% | 38% | 37.5 |
| 4 | 53% | 53% | 62.5 |
| 5 | 68% | 68% | 87.5 |
| 6 | 82% | 82% | 125 |
| 7 | 97% | 97% | 233.3 |

**Table 9. Tobit Regression for Factors Affecting Farmers' Riparian Buffer Adoption**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| Offer | 148.39 | 1220.14 | 5844.58* | 2285.81 | 3701.85 | 6412.12** |
| | (1774.11) | (1033.83) | (3254.79) | (3214.25) | (3027.58) | (3056.01) |
| Land acre | -456.93 | -276.9 | -926.19 | -737.95 | -583.55 | -125.84 |
| | (856.22) | (325.95) | (1196.95) | (1084.06) | (946.81) | (767.96) |
| Rent percentage | 580.55 | -71.66 | 814.53 | 1138.14 | 324.07 | -25.76 |
| | (1077.33) | (595.01) | (1814.07) | (1623.91) | (1589.47) | (1645.31) |
| Surface water | 1202.77 | 117.48 | 664.56 | 116.78 | 147.55 | 344.27 |
| | (935.51) | (481.15) | (1257.14) | (1180.68) | (1192.79) | (1119.89) |
| Percentage of household | -1895.26 | -1149.35 | -3013.64 | -2667.28 | -3300.94* | -1980.11 |
| income from farming | (1354.15) | (741.73) | (2160.84) | (1942.95) | (1899.53) | (1899.19) |
| Total household income | 2085.75 | 1677.21* | 4707.3* | 3943.3* | 4499.54* | 2739.09 |
| reinvested back to farm | (1423.19) | (864.28) | (2485.1) | (2124.31) | (2346.65) | (2325.5) |
| Income | 5.49 | -2.59 | -2.73 | -3.82 | -3.34 | -5.01 |
| | (4.26) | (2.44) | (7.24) | (6.61) | (6.58) | (7.03) |
| Water quality | -217.65 | -134.7 | -58.25 | 143.91 | -247.15 | 13.89 |
| | (187.78) | (115.17) | (309.29) | (257.06) | (289.95) | (279.94) |
| CRP | -452.92 | 237.62 | 1274.44 | 544.59 | 1242.17 | 1019.14 |
| | (772.65) | (454.84) | (1253.61) | (1177.28) | (1207.5) | (1317.42) |
| WLP | 1065.28 | 276.65 | -743.54 | -81.66 | -89.58 | -497.1 |
| | (643.69) | (352.12) | (1116.72) | (1016.35) | (1072.98) | (1005.95) |
| **Current usage of other BMPs:** | | | | | | |
| Riparian buffers | 1577.12*** | 1267.72*** | 2831.04*** | 2121** | 2754.91*** | 2095.24** |
| | (593.6) | (335.24) | (967.91) | (908.09) | (884.11) | (910.39) |
| Animal fence | -1204.81* | -695.18* | -1173.39 | -1269.11 | -1324.06 | -248.89 |
| | (611.44) | (369.24) | (1050.72) | (1008.1) | (911.06) | (1026.49) |
| No till | -731.74 | 184.27 | 1119.03 | 995.1 | 1308.1 | 1612.13 |
| | (756.87) | (412.41) | (1299.82) | (1108.48) | (1100.15) | (1132.05) |
| Waste storage | -2419.11 | -1370.3** | -2363.07 | -1991.94 | -2142.06 | -3440.22* |
| facility | (1484.21) | (662.39) | (1943.83) | (1960) | (1903.22) | (1860.99) |
| Nutrient | -275 | -170.88 | -449.41 | -823.09 | -3.8 | -789.36 |
| management | (622.43) | (372) | (1125.2) | (962.36) | (966.26) | (1041.01) |

(Continued)

**Table 9. Continued**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| **Choices of other BMPs:** | | | | | | |
| Animal fences | 3428.09*** | 2004.32*** | 4010.42*** | 4031.86*** | 3896.6*** | 2300.1*** |
| | (679.85) | (383.24) | (1209.22) | (1275.85) | (1122.78) | (858.18) |
| No till | 283.72 | 290.42 | 1098.67 | 1366.6 | 1365.59 | 421.93 |
| | (661.76) | (382.08) | (1119.14) | (983.97) | (1019.47) | (888.47) |
| Waste storage facilities | -1037.53 | -697.42* | -1359.29 | -1241.41 | -1702.55 | -57.96 |
| | (719.85) | (415.1) | (1308.26) | (1087.33) | (1248.52) | (1060.42) |
| Nutrient management | 601.73 | 218.56 | 427.57 | 800.82 | 966.3 | 452 |
| | (693.89) | (344.1) | (1133.87) | (1010.98) | (1108.74) | (937.01) |
| **Information about WQT:** | | | | | | |
| Cost saving information | -1467.2** | -494.17 | -582.07 | -268.14 | -179.45 | -539.6 |
| | (732.95) | (425.22) | (1210.91) | (1006.5) | (1129.82) | (1147.45) |
| Environmental aspect Info | -127.62 | 164.53 | 664.12 | 550.4 | 1188.01 | 187.72 |
| | (782.43) | (452.72) | (1468.57) | (1353.66) | (1320.7) | (1256.05) |
| Combined Information | -448.68 | -321.88 | -346.48 | -139.22 | -292.91 | -271.78 |
| | (624.82) | (415.97) | (1213.4) | (991.47) | (1012.88) | (1069.11) |
| Constant | -3139.38 | -2052.21 | -10156.97** | -7200.64* | -8027.06** | -9291.22** |
| | (2372.58) | (1375.97) | (4304.47) | (3945.22) | (4005.27) | (3902.3) |
| Sigma | 1690.19*** | 1361.95*** | 3308.4*** | 3025.59*** | 3309.37*** | 3459.34*** |
| | (203.13) | (120.2) | (682.62) | (586.06) | (588.63) | (567.43) |
| N | 119 | 149 | 149 | 225 | 199 | 218 |
| | | | | 256 | 237 | 251 |
| Largest FMI | - | - | 0.8199 | 0.8859 | 0.7452 | 0.8211 |

Note:

1. The "yes/no" choices are imputed in the last three scenarios, so numbers of observations used in the estimation are varied across different imputation datasets. We report the largest and the smallest numbers of observation for the last three scenarios.
2. Standard errors in parentheses; *, **, and *** imply significant at the 10%, 5%, and 1% significance levels, respectively.

**Table 10. Tobit Regression for Factors Affecting Farmers' Animal Fences Adoption**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| Offer | 1469.02 | 1057.82 | 3186.06 | 2615.7 | 2059.35 | 3207.67 |
| | (1640.03) | (1035.47) | (2015.51) | (2066.06) | (1802.69) | (2140.37) |
| Land acre | -1685.91* | -957.62* | -2771.64** | -2691.97** | -2491.41** | -2715.92*** |
| | (857.87) | (549.12) | (1129.81) | (1117.19) | (1037.2) | (1024.52) |
| Rent percentage | 106.51 | 486.85 | 993 | 839.47 | 953.47 | 1266.29 |
| | (956.73) | (580.61) | (1270.19) | (1388.32) | (1102.85) | (1306.33) |
| Surface water | -179.94 | -283.98 | -788.63 | -1176.3 | -1006.45 | -460.42 |
| | (855.98) | (514.83) | (1106.91) | (1178.09) | (965.1) | (875.08) |
| Percentage of household income from farming | 2182.16 | 974.57 | 2551.81 | 5307.78*** | 1642.65 | 4006.3 |
| | (1536.51) | (881.1) | (1761.24) | (1963.67) | (1755.76) | (1976.92) |
| Total household income reinvested back to farm | -1047.69 | -317.66 | -55.61 | -4159.16* | 851.64 | -2383.9 |
| | (1726.26) | (1040.45) | (2021.52) | (2156.11) | (1878.66) | (2080.62) |
| Income | 2.59 | 0.39 | 4.75 | 5.12 | 2.36 | 1.37 |
| | (4.51) | (2.68) | (5.44) | (5.99) | (5.18) | (5.79) |
| Water quality | -323.7* | -297.78** | -346.65 | -274.07 | -360.54 | -392.36* |
| | (192.44) | (124.3) | (233.67) | (235.4) | (220.94) | (229.11) |
| CRP | 49.14 | -2.44 | -1035.6 | -197.15 | -226.37 | -232 |
| | (824.82) | (513.16) | (964.84) | (1001.32) | (1002.23) | (987.01) |
| WLP | -742.25 | -190.82 | 44.61 | -757.05 | -313.56 | -454.49 |
| | (668.87) | (385.01) | (869.01) | (804.03) | (775.26) | (842.22) |
| **Current usage of other BMPs:** | | | | | | |
| Riparian buffers | 1058.09** | 445.21 | 790.18 | 724.08 | 1080.87* | 939.95 |
| | (518.25) | (343.92) | (650.75) | (709.87) | (616.99) | (669.21) |
| Animal fence | 2108.27*** | 1100.09*** | 1866.46*** | 1900.53*** | 1906.62*** | 1935.93*** |
| | (555.56) | (347.88) | (654.19) | (688.08) | (645.01) | (701.09) |
| No till | 1316.6* | 678.88 | 2499.37** | 1227.24 | 1573.23* | 1508.3 |
| | (751.35) | (437.38) | (999.12) | (940.03) | (931.89) | (913.87) |
| Waste storage facility | 1900.86* | 375.21 | 1323.54 | 2373.2* | 1621.97 | 1906.9 |
| | (1131.87) | (679.77) | (1305.82) | (1360.33) | (1245.28) | (1547.76) |
| Nutrient management | -1957.04*** | -933.43** | -2618.82*** | -2003.13** | -2148.27** | -2249.92*** |
| | (695.79) | (428.86) | (856.41) | (863.8) | (900.82) | (855.37) |

(Continued)

**Table 10. Continued**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| **Choices of other BMPs:** | | | | | | |
| Riparian Buffers | 1073.98* | 476.89 | 1509.14** | 1906.97*** | 1653.1** | 1106.79* |
| | (568.51) | (344.96) | (683.06) | (698.16) | (687.26) | (625.37) |
| No till | -46.95 | 64.76 | -506.06 | 302.38 | 611.49 | 181.41 |
| | (643.94) | (406.13) | (767.88) | (877.25) | (772.81) | (657.55) |
| Waste storage facilities | 739.87 | 418.54 | 1053.9 | 1373.8* | 976.21 | 709.18 |
| | (691.64) | (408.82) | (825.54) | (825.95) | (827.23) | (783.18) |
| Nutrient management | -275.52 | 9.3 | -396.68 | -167.82 | -135.29 | 115.84 |
| | (636.64) | (408.15) | (786.34) | (781.51) | (788.94) | (723.61) |
| **Information about WQT:** | | | | | | |
| Cost saving information | 1557.11** | 844.07** | 1576.89* | 1298.35 | 1898.13** | 1157.19 |
| | (714.19) | (431.34) | (905.43) | (909.81) | (833.31) | (876.46) |
| Environmental aspect Info | 503.06 | 105.31 | 478.65 | 412.75 | 260.87 | 33.14 |
| | (724.79) | (477.44) | (955.09) | (878.6) | (851.38) | (937.62) |
| Combined Information | -437.96 | -305.87 | -519.48 | -973.81 | -151.72 | -1091.03 |
| | (682.44) | (428.89) | (839.71) | (860.9) | (744.84) | (863.53) |
| Constant | -2095.28 | 15.3 | -2940.03 | -2559.48 | -2889.84 | -2288.85 |
| | (2221.91) | (1443.59) | (2644.73) | (2582.91) | (2520.92) | (3010.75) |
| Sigma | 2245.07*** | 1834.23*** | 2766.97*** | 2771.55*** | 2829.42*** | 2863.08*** |
| | (199.15) | (124.74) | (340.89) | (300.06) | (329.99) | (343.59) |
| N | 134 | 182 | 182 | 249 | 216 | 253 |
| | | | | 276 | 255 | 276 |
| Largest FMI | - | - | 0.6987 | 0.7581 | 0.65 | 0.7721 |

Note:

1. The “yes/no” choices are imputed in the last three scenarios, so numbers of observations used in the estimation are varied across different imputation datasets. We report the largest and the smallest numbers of observation for the last three scenarios.
2. Standard errors in parentheses; *, **, and *** imply significant at the 10%, 5%, and 1% significance levels, respectively.

**Table 11. Tobit Regression for Factors Affecting Farmers' No Till Adoption**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| Offer | -97.79 | -56.97 | -16 | -43.71 | -30.39 | -27.39 |
| | (76.81) | (51.43) | (113.79) | (94.7) | (105.9) | (87.48) |
| Land acre | 18.81 | 18.48* | 21.37 | 25.25 | 28.38 | 19.07 |
| | (13.14) | (10.22) | (19.96) | (19.74) | (20.22) | (18.05) |
| Rent percentage | 99.47*** | 48.84* | 110.67 | 129.45* | 104.32* | 77.76 |
| | (37.44) | (25.51) | (73.79) | (68.73) | (62.9) | (56.1) |
| Surface water | -5.35 | -16.35 | 10.39 | 25.26 | 33.82 | 11.04 |
| | (37.15) | (25.51) | (52.16) | (49) | (53.11) | (37.62) |
| Percentage of household income from farming | 115.43** | 106.63*** | 164.24* | 168.35** | 157.03** | 156.13** |
| | (57.8) | (39.55) | (85.23) | (84.27) | (79.2) | (74.69) |
| Total household income reinvested back to farm | -71.86 | -70.91 | -62.77 | -90.46 | -111.54 | -18.53 |
| | (65.56) | (47.42) | (93.81) | (86.34) | (92.73) | (78.51) |
| Income | 0.41** | 0.25 | 0.35 | 0.31 | 0.37 | 0.33 |
| | (0.18) | (0.11) | (0.26) | (0.24) | (0.25) | (0.21) |
| Water quality | -8.46 | -6.54 | -9.71 | -9.25 | -6.5 | -12.44 |
| | (8.01) | (5.75) | (11.46) | (9.97) | (11.29) | (10.14) |
| CRP | -27.27 | -22.36 | -74.46 | -74.19 | -57.17 | -71.03* |
| | (31) | (22.17) | (51.04) | (48.67) | (47.18) | (42.78) |
| WLP | 19.19 | 18.9 | 5.34 | 20.63 | 0.8 | 20.31 |
| | (29.66) | (18.97) | (47.91) | (40.53) | (41.26) | (33.94) |
| **Current usage of other BMPs:** | | | | | | |
| Riparian buffers | -4.82 | 9.14 | 5.46 | -12.85 | -12.99 | 1.71 |
| | (24.65) | (16.6) | (36.73) | (31.83) | (33.85) | (32.28) |
| Animal fence | 2.67 | 2.51 | 7.39 | 1.09 | -10.64 | 17.79 |
| | (27.66) | (18.48) | (38.82) | (35.82) | (35.31) | (31.43) |
| No till | 103.8*** | 72.72*** | 129.75*** | 119.5*** | 128.53*** | 121.21*** |
| | (25.57) | (17.25) | (43.2) | (38.59) | (41.3) | (37.55) |
| Waste storage facility | -106.39** | -63.42** | -121.23* | -116.29** | -120.54* | -115.84** |
| | (43.22) | (27.39) | (66.23) | (58.2) | (63.48) | (53.2) |
| Nutrient management | -20.66 | -16.84 | -8.3 | -14.95 | 10.64 | -16.31 |
| | (28.08) | (19.24) | (41.04) | (36.83) | (40.08) | (33.98) |

(Continued)

**Table 11. Continued**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| **Choices of other BMPs:** | | | | | | |
| Riparian Buffers | 19.51 | 10.11 | 47.8 | 38.17 | 66.86 | 16.69 |
| | (28.25) | (17.64) | (47.96) | (37.88) | (44.2) | (31.13) |
| Animal fences | 6.05 | 16.75 | 8.09 | 12.11 | 27.57 | -4.66 |
| | (31.25) | (20.17) | (43.96) | (43.22) | (39.41) | (30.98) |
| Waste storage facilities | 34.81 | 14.34 | 38.46 | 30.88 | 68.39* | 28.86 |
| | (30.54) | (19.51) | (41.95) | (38.44) | (42.69) | (33.34) |
| Nutrient management | 43.79 | 37.33** | 45.66 | 76.74* | 77.35* | 22.12 |
| | (27.8) | (17.81) | (37.77) | (39.44) | (40.4) | (29.42) |
| **Information about WQT:** | | | | | | |
| Cost saving information | 25.67 | 28.23 | 52.48 | 33.87 | 26.1 | 31.51 |
| | (30.05) | (20.74) | (45.96) | (39.22) | (44.53) | (36.13) |
| Environmental aspect Info | 14.09 | 20.73 | 70.8 | 48.88 | 51.54 | 25.83 |
| | (35.74) | (22.39) | (62.6) | (51.42) | (51.61) | (44.97) |
| Combined Information | 13.9 | 10.01 | 14.58 | -18.06 | -6.36 | 0.81 |
| | (28.65) | (20.98) | (41.11) | (40.27) | (39.32) | (35.2) |
| Constant | -21.02 | -1.54 | -152.98 | -130.48 | -206.82 | -76.64 |
| | (104.2) | (66.57) | (160.8) | (121.15) | (154.09) | (116.22) |
| Sigma | 98.78*** | 85.23*** | 141.83*** | 136.51*** | 146.96*** | 128.51*** |
| | (8.75) | (5.93) | (26.25) | (21.52) | (26.11) | (20.89) |
| N | 136 | 178 | 178 | 254 | 226 | 254 |
| | | | | 285 | 264 | 283 |
| Largest FMI | - | - | 0.8684 | 0.8709 | 0.8501 | 0.8825 |

Note:

1. The "yes/no" choices are imputed in the last three scenarios, so numbers of observations used in the estimation are varied across different imputation datasets. We report the largest and the smallest numbers of observation for the last three scenarios.
2. Standard errors in parentheses; *, **, and *** imply significant at the 10%, 5%, and 1% significance levels, respectively.

**Table 12 . Poisson Regression for Factors Affecting Farmers' Waste Storage Facilities Adoption**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| Offer | 0.091 | -0.17 | -0.633 | -0.459 | -0.281 | -0.127 |
| | (1.03) | (0.854) | (0.84) | (0.725) | (0.827) | (0.751) |
| Land acre | 0.034 | -0.012 | 0.024 | 0.016 | 0.1 | 0.033 |
| | (0.264) | (0.2) | (0.172) | (0.187) | (0.156) | (0.171) |
| Rent percentage | -0.743 | -0.389 | -0.199 | -0.119 | -0.404 | -0.251 |
| | (0.577) | (0.468) | (0.446) | (0.364) | (0.442) | (0.377) |
| Surface water | -0.197 | -0.104 | 0.056 | 0.181 | 0.058 | 0.055 |
| | (0.514) | (0.426) | (0.439) | (0.403) | (0.429) | (0.362) |
| Percentage of household income from farming | -0.254 | -0.388 | -0.524 | -0.63 | -0.682 | -0.448 |
| | (0.69) | (0.598) | (0.584) | (0.516) | (0.592) | (0.537) |
| Total household income reinvested back to farm | 1.746** | 1.381** | 1.11* | 0.87 | 1.211* | 1.179** |
| | (0.754) | (0.613) | (0.602) | (0.55) | (0.625) | (0.556) |
| Income | -0.003 | -0.001 | -0.002 | -0.002 | -0.002 | -0.002 |
| | (0.003) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Water quality | -0.035 | -0.06 | -0.072 | -0.059 | -0.099 | -0.092 |
| | (0.121) | (0.095) | (0.092) | (0.088) | (0.097) | (0.084) |
| CRP | 0.393 | 0.249 | 0.175 | 0.197 | 0.219 | 0.271 |
| | (0.425) | (0.31) | (0.314) | (0.294) | (0.313) | (0.294) |
| WLP | 0.188 | 0.173 | 0.197 | 0.211 | 0.266 | 0.071 |
| | (0.325) | (0.261) | (0.26) | (0.251) | (0.266) | (0.241) |
| **Current usage of other BMPs:** | | | | | | |
| Riparian buffers | 0.344 | 0.291 | 0.2 | 0.032 | 0.111 | 0.241 |
| | (0.304) | (0.259) | (0.265) | (0.256) | (0.27) | (0.233) |
| Animal fence | -0.062 | -0.088 | -0.127 | -0.196 | -0.148 | -0.044 |
| | (0.332) | (0.283) | (0.278) | (0.247) | (0.265) | (0.261) |
| No till | -0.425 | -0.239 | -0.001 | 0.026 | -0.071 | 0.016 |
| | (0.434) | (0.325) | (0.327) | (0.272) | (0.329) | (0.297) |
| Waste storage facility | 0.542 | 0.295 | 0.209 | 0.333 | 0.345 | 0.159 |
| | (0.417) | (0.343) | (0.339) | (0.312) | (0.338) | (0.305) |
| Nutrient management | -0.065 | -0.158 | -0.233 | -0.249 | -0.178 | -0.179 |
| | (0.361) | (0.298) | (0.287) | (0.252) | (0.288) | (0.264) |

(Continued)

**Table 12. Continued**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| **Choices of other BMPs:** | | | | | | |
| Riparian Buffers | -0.07 | -0.105 | -0.12 | -0.041 | -0.154 | -0.063 |
| | (0.333) | (0.278) | (0.275) | (0.264) | (0.284) | (0.245) |
| Animal fences | 1.146*** | 0.843*** | 0.757** | 0.723** | 0.841*** | 0.457* |
| | (0.38) | (0.296) | (0.296) | (0.328) | (0.3) | (0.247) |
| No till | -0.22 | 0.013 | 0.108 | 0.135 | 0.394 | 0.046 |
| | (0.354) | (0.285) | (0.294) | (0.303) | (0.307) | (0.26) |
| Nutrient management | 1.049*** | 0.799*** | 0.84*** | 0.993*** | 1.041*** | 0.575** |
| | (0.361) | (0.291) | (0.291) | (0.357) | (0.305) | (0.251) |
| **Information about WQT:** | | | | | | |
| Cost saving information | 0.541 | 0.316 | 0.293 | 0.41 | 0.359 | 0.418 |
| | (0.437) | (0.365) | (0.358) | (0.326) | (0.356) | (0.334) |
| Environmental aspect Info | 0.34 | 0.272 | 0.35 | 0.351 | 0.35 | 0.346 |
| | (0.49) | (0.391) | (0.397) | (0.34) | (0.391) | (0.358) |
| Combined Information | 0.376 | 0.228 | 0.182 | 0.263 | 0.248 | 0.332 |
| | (0.476) | (0.39) | (0.386) | (0.362) | (0.388) | (0.371) |
| Constant | -2.364 | -1.473 | -0.767 | -1.157 | -1.555 | -0.887 |
| | (1.443) | (1.176) | (1.137) | (1.045) | (1.155) | (1.054) |
| N | 128 | 151 | 151 | 211 | 200 | 223 |
| | | | | 243 | 231 | 253 |
| Largest FMI | - | - | 0.1766 | 0.5681 | 0.2519 | 0.4706 |

Note:

1. The "yes/no" choices are imputed in the last three scenarios, so numbers of observations used in the estimation are varied across different imputation datasets. We report the largest and the smallest numbers of observation for the last three scenarios.
2. Standard errors in parentheses; *, **, and *** imply significant at the 10%, 5%, and 1% significance levels, respectively.

**Table 13. Tobit Regression for Factors Affecting Farmers' Nutrient Management Adoption**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| Offer | 61.76 | 54.31 | 106.64 | 77.08 | 141.67 | 149.16 |
| | (143.74) | (107.83) | (167.03) | (152.04) | (157.92) | (146.53) |
| Land acre | 29.77 | 27.63 | 31.16 | 30.92 | 16.75 | 32.4 |
| | (28.44) | (22.32) | (33.48) | (31.65) | (33.91) | (33.7) |
| Rent percentage | 25.6 | 33.87 | 89.45 | 115.43 | 63.95 | 111.09 |
| | (74.73) | (55.06) | (93.72) | (101.12) | (92.54) | (97.19) |
| Surface water | -41.14 | -34.5 | -38.88 | -57.29 | -74.49 | -3.14 |
| | (64.49) | (49.06) | (70.14) | (63.54) | (68.87) | (58.55) |
| Percentage of household income from farming | 44.41 | 46.31 | 21.74 | 66.86 | -2.73 | 30.66 |
| | (104.19) | (75.85) | (117.28) | (128.78) | (109.4) | (103.02) |
| Total household income reinvested back to farm | 49.36 | -7.46 | 27.45 | -2.1 | -2.85 | 76.11 |
| | (136.86) | (89.67) | (148.35) | (137.03) | (146.52) | (132.95) |
| Income | 0.29 | -0.05 | 0.24 | 0.27 | 0.55 | 0.02 |
| | (0.33) | (0.23) | (0.41) | (0.4) | (0.39) | (0.38) |
| Water quality | -5.17 | -0.22 | -2.02 | 8.92 | 8.87 | -3.63 |
| | (16.12) | (12.84) | (20.09) | (17.95) | (19.39) | (17.58) |
| CRP | 38.41 | 50.49 | 64.71 | 57.02 | 8.74 | 52.53 |
| | (71.02) | (49.47) | (89.46) | (75.94) | (84.9) | (71.21) |
| WLP | -72.23 | -84.61** | -104.71 | -120.37* | -98.82 | -74.71 |
| | (54.59) | (41.16) | (74.43) | -66.82 | (67.13) | (62.91) |
| **Current usage of other BMPs:** | | | | | | |
| Riparian buffers | 97.14** | 82.94** | 102.53* | 102.82* | 93.87* | 93.28* |
| | (49.59) | (34.77) | (55.79) | (54.79) | (52.56) | (50.56) |
| Animal fence | -10.72 | 1.12 | 20.63 | 19.33 | 37.66 | 38.77 |
| | (47.49) | (35.7) | (55.81) | (49.98) | (50.38) | (45.38) |
| No till | 73.57 | 60.67 | 96.59 | 76.78 | 101.53 | 120.15* |
| | (57.06) | (40.9) | (66.18) | (56.61) | (65.89) | (65.77) |
| Waste storage facility | -139.59 | -91.76 | -133.35 | -125.8 | -128.2 | -154.27 |
| | (84.85) | (58.52) | (105.39) | (98.53) | (96.93) | (103.49) |
| Nutrient management | 147.99*** | 108.79*** | 141.46** | 128.97** | 172.09*** | 106.36** |
| | (48.59) | (35.15) | (57.02) | (51.25) | (56.71) | (48.36) |

(Continued)

**Table 13. Continued**

| | Scenario 1 (Deletion method) | Scenario 2 (Mean imputation) | Scenario 3 (One-stage) | Scenario 4 (Two-stage) | Scenario 5 (Two-stage with restriction) | Scenario 6 (Three-stage) |
|---|---|---|---|---|---|---|
| **Choices of other BMPs:** | | | | | | |
| Riparian Buffers | 24.7 | 35.08 | 44.46 | 29.59 | 58.44 | 33.95 |
| | (53.18) | (37.93) | (61.12) | (55.54) | (59.96) | (52.97) |
| Animal fences | 15.97 | 9.01 | 13.11 | 15.7 | 20.63 | 16.46 |
| | (55.77) | (40.02) | (63.24) | (60.13) | (57.57) | (55.34) |
| No till | 70.64 | 78.6** | 95.7 | 122.54* | 144.86** | 48.36 |
| | (53.65) | (38.96) | (63.11) | (64.6) | (62.54) | (53.22) |
| Waste storage facilities | 138.57*** | 104.59*** | 145.52** | 157.73** | 164.39** | 89.14* |
| | (52.41) | (38.19) | (62.49) | (65.53) | (64.16) | (52.36) |
| **Information about WQT:** | | | | | | |
| Cost saving information | 4.9 | 16.9 | 14.65 | 2.42 | 37.08 | 27.16 |
| | (65.32) | (47.48) | (74.55) | (67.87) | (77.55) | (74.95) |
| Environmental aspect Info | -13.23 | -19.75 | 15.38 | -39.63 | 25.14 | -9.91 |
| | (66.26) | (48.52) | (81.61) | (68.75) | (78.53) | (77.17) |
| Combined Information | 60.32 | 40.49 | 44.75 | 40.85 | 35.28 | 44.21 |
| | (60.92) | (46.88) | (68.69) | (61.65) | (64.98) | (65.98) |
| Constant | -310.09 | -225.54 | -395.18 | -416.3* | -566.63** | -395.25* |
| | (200.46) | (145.83) | (242.38) | (229.73) | (249.89) | (212.5) |
| Sigma | 209.46*** | 180.13*** | 236.07*** | 226.54*** | 245.81*** | 229.56*** |
| | (17.12) | (12.41) | (44.13) | (49.68) | (47.38) | (46.48) |
| N | 145 | 176 | 176 | 254 | 239 | 264 |
| | | | | 288 | 272 | 290 |
| Largest FMI | - | - | 0.8718 | 0.9376 | 0.8747 | 0.9274 |

Note:

1. The "yes/no" choices are imputed in the last three scenarios, so numbers of observations used in the estimation are varied across different imputation datasets. We report the largest and the smallest numbers of observation for the last three scenarios.
2. Standard errors in parentheses; *, **, and *** imply significant at the 10%, 5%, and 1% significance levels, respectively.