Preliminary and Incomplete Draft – Please Do Not Cite

From Farmers' Management Decisions to Watershed Water Quality: A Spatial Economic Model of Land Management Choices

Wendong Zhang

Department of Agricultural, Environmental and Development Economics
The Ohio State University
zhang.822@osu.edu

Elena G. Irwin
Department of Agricultural, Environmental and Development Economics
The Ohio State University
irwin.78@osu.edu

Date: November 11th, 2013

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's

2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

I. Introduction

Non-point source sediment and nutrient runoff from upstream agricultural production is known to impair downstream ecosystem systems and ecosystem services (e.g. Diaz and Rosenberg 2008), which may worsen due to more intense rains with continued climate change (e.g. Groisman et al. 2005). Despite adoption of agricultural best management practices by some farmers, there is a need for greater adoption of conservation treatment. In 2011, USDA's Conservation Effects Assessment Program (CEAP) identified that 16 % (2.8 million acres) and 34% (5 million acres) of the cultivated cropland in Great Lakes Region have a high and moderate level of need for additional conservation treatment to reduce agricultural pollution loadings, respectively (CEAP 2011). The positive environmental impacts of conservation tillage systems are well documented, including reduction in soil erosion, increase in water retention, and reduction in water and nutrient runoff (e.g. Knowler and Bradshaw 2007; Gould et al. 1989). However, significant uncertainty still remains in about the effectiveness of policies that encourage adoption of conservation practices. Much of this uncertainty mainly results from critical gaps in understanding farmers' spatially heterogenous behavioral responses to changes in policies and linkages between farmers' individual decisions and landscape environmental impacts.

Factors determining farmers' adoption decisions of conservation tillage have been extensively studied. Larger farm size, younger age, higher income, better exposure to conservation information, and higher percentage of sloped cropland are found to increase the likelihood of adoption (Norris and Batie 1987; Gould et al. 1989; Featherstone and Goodwin 1993). These studies provide useful insights regarding the importance of spatial information and operator characteristics in determining tillage choices; however, they derived from sample-based surveys of farmers that are not spatially located, and therefore is not possible to link tillage decisions with sediment and nutrient runoffs. To reliably predict nutrient and sediment runoffs, the predictions of agricultural land management outcomes are needed for all land parcels across the entire watershed. In addition, the spatially contiguous field level analysis across the entire watershed is needed to evaluate the effectiveness of spatially targeted policies. Given the limited conservation budget and different environmental impacts across different fields due to their locations, extensive studies have shown that spatially targeted conservation strategy with the objective to maximize the net environmental benefits will enhance economic efficiency (e.g. Wu and Boggess 1999, Lewis et al. 2009).

Recent models of conservation choices (e.g. Wu et al. 2004; Kurkalova et al. 2006) have incorporated the rich spatial heterogeneity of parcels. CEAP watershed studies also attempted to incorporate spatially located Natural Resource Inventory survey points with separate surveys of farmer characteristics; however, the number of survey points is sparsely distributed across multiple river basins in Great Lakes Region (CEAP 2011), which makes it difficult to reliably predict nutrient and sediment runoffs for the entire watershed. Econometrically, previous models such as Wu et al. (2004) typically separately model tillage practice decisions given the choice of crop. This convenient decomposition is correct if the error terms for tillage decisions and the error terms for crop choices are not correlated. However, in reality, these two decisions are jointly influenced by common factors unobserved to the researchers such as the family tradition, which leads to endogeneity bias if left uncontrolled.

The aim of this paper is to develop a spatially explicit model of farmers' crop and tillage choices while accounting for the endogenous linkages between these two choices by treating crop choices as the multinomial endogenous treatment. We adapt the multinomial treatment effects model and its simulated maximum likelihood estimation technique proposed by Deb and Trivedi (2006), and control for the common unobserved characteristics by introducing a latent factor structure into both the treatment (crop choice) and the outcome (tillage choice). We apply this model to a spatially contiguous agricultural field boundaries data in the Maumee River watershed, and we translate the field-level crop and tillage outcomes into nutrient runoffs, especially the dissolved reactive phosphorus (DRP) loadings, using a hydrological watershed model called Soil and Water Assessment Tool (SWAT). Doing so allows us to evaluate different policy scenarios not only in terms of their simulated impacts on farmers' behavior and agricultural management outcomes, but more importantly in terms of their impacts on nutrient runoffs and downstream ecosystem conditions.

The main result provides evidence that there are significant treatment effects between the tillage choices and crop choices, suggesting that there is endogeneity bias if these two decisions are separately modeled. Our analysis also reveal that spatial variations at the field level have a significant impact in determining the crop and tillage choices: for example, farmers are more likely to grow corn on a bigger field with better soil quality. In addition, we find that not every field has the same environmental impacts due to their different spatial locations, and we identified several "hot spot" – environmental sensitive areas by translating crop and tillage outcomes into DRP loadings using SWAT model. Due to computational complexity, we did not run a lot of policy scenarios such as first-best DRP loading taxes, voluntary payment schemes to encourage conservation tillage. However, our little exercise shows that incorporation is crucial to reduce the elevated DRP loadings from no-till, especially for agricultural fields more susceptible to nutrient runoffs.

Overall, this study makes at least two contributions to the literature on the spatial modeling of conservation practices. First, our analysis reveals that separate estimation of crop and tillage choices may suffer from endogeneity, which has to be addressed using methods like our multinomial treatment effects simulated likelihood model. In addition, our spatially contiguous agricultural field data allows for integration with watershed hydrological model and the benefit cost analysis of alternative policy mechanisms such as the spatially targeted policies.

II. Conceptual Framework

In this modeling framework and following Wu et al. (2004), a farmer is assumed to choose a combination of crop and tillage practice decisions at the field level that yields the highest expected utility[1]. These two choices are made simultaneously; the choice of tillage practice may depend on the crop choice or vice versa. Assume the farmer can choose among K crops and M tillage systems, her utility $u_{ij}(X_i, Z_{j|i})$ of choosing crop i and tillage system j can be represented as follows:

$$u_{ij}(X_i, Z_{j|i}) = \upsilon_{ij}(\mathbf{X_i}, \mathbf{Z_{j|i}}) + \varepsilon_{ij}, \text{ i=1, 2,…K and j= 1, 2,… M} \qquad (1)$$

Where $\mathbf{X_i}$ is a vector of variables specific to the crop choice decision, including field-specific characteristics and expected crop prices, and $\mathbf{Z_{j|i}}$ is a vector of variables that influence the farmer's utility from adopting different tillage system, including the cost differential between different tillage system. $\upsilon_{ij}(\mathbf{X_i}, \mathbf{Z_{j|i}})$ captures the deterministic portion of utility that can be explained by $\mathbf{X_i}$ and $\mathbf{Z_{j|i}}$, and is specified as $\upsilon_{ij}(\mathbf{X_i}, \mathbf{Z_{j|i}}) = \mathbf{X_i'\beta_i} + \mathbf{Z_{j|i}'\gamma_{j|i}}$. ε_{ij} is a random error

[1] We assume utility rather than profit maximization to account for other non-economic motives that farmers may have, including family succession and environmental stewardship (Konar et al. 2012).

term that captures other factors that could influence crop and/or tillage choices but unobserved to researchers. A farm operator will choose crop i and tillage practice j, over other crop choices k or tillage practices m if utility is maximized:

$$u_{ij}(X_i, Z_{j|i}) > u_{km}(X_k, Z_{m|k}), \qquad k \neq i \text{ and } m \neq j \qquad (2)$$

The stochastic version of this model for estimation is that the probability of crop choice i and tillage practice j is shown as

$$P_{ij} = Prob\ (crop\ i, tillage\ j) = prob\ (u_{ij} > u_{km}), k \neq i \text{ and } m \neq j$$

$$= \text{prob}\ (\mathbf{X}_\mathbf{i}'\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z}_{\mathbf{j|i}}'\boldsymbol{\gamma}_{\mathbf{j|i}} + \varepsilon_{ij} > \mathbf{X}_\mathbf{k}'\boldsymbol{\beta}_\mathbf{k} + \mathbf{Z}_{\mathbf{m|k}}'\boldsymbol{\gamma}_{\mathbf{m|k}} + \varepsilon_{mk})$$

$$= \text{prob}\ (\varepsilon_{ij} - \varepsilon_{ij} < \mathbf{X}_\mathbf{k}'\boldsymbol{\beta}_\mathbf{k} + \mathbf{Z}_{\mathbf{m|k}}'\boldsymbol{\gamma}_{\mathbf{m|k}} - \mathbf{X}_\mathbf{i}'\boldsymbol{\beta}_\mathbf{i} - \mathbf{Z}_{\mathbf{j|i}}'\boldsymbol{\gamma}_{\mathbf{j|i}}) \qquad (3)$$

III. Econometrics

To estimate equation (3) econometrically, ε_{ij} is typically assumed to follow extreme value I distribution, in which case the probability of choosing crop i and tillage practice j can be represented by a multinomial model (Maddala 1983):

$$P_{ij} = \frac{e^{\mathbf{X}_\mathbf{i}'\boldsymbol{\beta}_\mathbf{i} + \mathbf{Z}_{\mathbf{j|i}}'\boldsymbol{\gamma}_{\mathbf{j|i}}}}{\sum_{k=1}^{K}\sum_m e^{\mathbf{X}_\mathbf{k}'\boldsymbol{\beta}_\mathbf{k} + \mathbf{Z}_{\mathbf{m|k}}'\boldsymbol{\gamma}_{\mathbf{m|k}}}}, \quad i=1, 2,\dots K \text{ and } j= 1, 2,\dots M \qquad (4)$$

Previous literature such as Wu et al. (2004) typically treated the joint probability of crop and tillage choices in equation (4) as a product as two separate components: the probability of choosing crop i, and the conditional probability of choosing tillage system j given the crop choice i[2]:

---

[2] Wu and Babcock (1998) also estimated a multinomial logit model of a nutrient management plan which is a combination of crop rotation, tillage system and soil testing.

$$P_{ij} = \frac{e^{\mathbf{X}_i'\boldsymbol{\beta}_i + \mathbf{Z}_{j|i}'\boldsymbol{\gamma}_{j|i}}}{\sum_m e^{\mathbf{X}_k'\boldsymbol{\beta}_k + \mathbf{Z}_{m|k}'\boldsymbol{\gamma}_{m|k}}} * \frac{\sum_m e^{\mathbf{X}_i'\boldsymbol{\beta}_i + \mathbf{Z}_{m|i}'\boldsymbol{\gamma}_{m|i}}}{\sum_{k=1}^{K}\sum_m e^{\mathbf{X}_k'\boldsymbol{\beta}_k + \mathbf{Z}_{m|k}'\boldsymbol{\gamma}_{m|k}}}$$

$$= \frac{e^{\mathbf{Z}_{j|i}'\boldsymbol{\gamma}_{j|i}}}{\sum_m e^{\mathbf{Z}_{m|k}'\boldsymbol{\gamma}_{m|k}}} * \frac{\sum_m e^{\mathbf{X}_i'\boldsymbol{\beta}_i + \mathbf{Z}_{m|i}'\boldsymbol{\gamma}_{m|i}}}{\sum_{k=1}^{K}\sum_m e^{\mathbf{X}_k'\boldsymbol{\beta}_k + \mathbf{Z}_{m|k}'\boldsymbol{\gamma}_{m|k}}}$$

= prob (tillage j| crop i) * prob (crop i) (5)

This convenient decomposition is correct only if the error terms ε_i for crop choices and the error terms $\varepsilon_{j|i}$ for the conditional tillage choices are not correlated. However, in reality many common factors unobserved to the researchers tend to influence both the crop choices and the tillage choices, including field-level operator characteristics. For example, the farmer may choose a certain crop and tillage system because of family tradition, or may not have enough time to grow corn or conventional tillage due to a large size of the operation and/or the presence of large spring precipitation before planting. In these cases, the crop choice decisions and the tillage choice decisions are jointly determined and a separate estimation of these two decisions would suffer from the endogeneity bias.

We solve this endogeneity bias by adapting Deb and Trivedi (2006)'s multinomial endogenous treatment effects model, in which the crop choices are defined as endogenous multinomial treatment, and the tillage choices are defined as the outcome variables. A latent factor structure is introduced to enter the treatment and outcome equations, which allows for idiosyncratic influences on crop choice to affect tillage outcomes. In this framework, the normalized distributed latent factors have a natural interpretation as proxies for unobserved covariates and the associated factor loadings can be interpreted the same as coefficients on observed covariates. This latent-factor framework is more generalized than the two-step method proposed by Lee (1993) since it is more efficient and could easily be adapted to other statistical structures for treatment and outcome.

For estimation purposes, the crop choices and the tillage outcomes are treated as separate components although they are estimated jointly. The probability of the multinomial crop choices can be represented by

$$\mathrm{prob}(\mathbf{d_i}|\mathbf{X_i'},\mathbf{l_i}) = \frac{\sum_m e^{\mathbf{X_i'\beta_i}+\mathbf{Z'_{m|i}\gamma_{m|i}}+\mathbf{l_i}}}{\sum_{k=1}^{K}\sum_m e^{\mathbf{X_k'\beta_k}+\mathbf{Z'_{m|k}\gamma_{m|k}}+\mathbf{l_k}}}$$

$$=\mathbf{g}(\mathbf{X_i'\beta_i}+\mathbf{Z'_{m|i}\gamma_{m|i}}+\mathbf{l_i}) \qquad \mathrm{i}=1,2,\ldots \mathrm{K} \qquad (6)$$

where $\mathbf{g}$ is an appropriate multinomial logit distribution, $\mathbf{d_i}$ is a binary dummy variable indicating the choice of crop i and $\mathbf{l_i}$ is a vector of latent factors that incorporates unobserved characteristics common to the farmer's crop and tillage choices specific to crop choice i.

The outcome is the tillage choices denoted as $\mathrm{y_j}$. The expected outcome equation is formulated as

$$\mathrm{E}(\mathrm{y_j}|\mathbf{Z'_{j|k}},\mathbf{d_k},\mathbf{l_k}) = \mathbf{Z'_{j|i}\gamma_{j|i}}+\sum_{\mathbf{k}}\gamma_{\mathrm{k}}\mathbf{d_k}+\sum_{\mathbf{k}}\lambda_{\mathrm{k}}\mathbf{l_k} \qquad (7)$$

Where γ_{k} denotes the treatment effects relative to the control, and when λ_{k}, the factor loading parameter, is positive when treatment and outcome are positively correlated through unobserved characteristics. Denote the probability distribution of tillage outcome as $\mathrm{f}(\mathrm{y_j}|\mathbf{Z'_{j|k}},\mathbf{d_k},\mathbf{l_k})$.

The joint distribution of treatment (crop choice) and outcome variables (tillage choice), conditional on the common latent factors, can be written as the product of the marginal density of crop choices and the conditional density of tillage choices given crop choices.

$$\mathrm{prob}\left(y_j|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{X}'_{\mathbf{i}},\mathbf{d_i},\mathbf{l_i}\right) = \mathrm{f}(y_j\left|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{d_k},\mathbf{l_k}\right) * \mathbf{g}(\mathbf{X}'_{\mathbf{i}}\boldsymbol{\beta}_{\mathbf{i}} + \mathbf{Z}'_{\mathbf{m|i}}\boldsymbol{\gamma}_{\mathbf{m|i}} + \mathbf{l_i}) \qquad (8)$$

The problem in estimation arises because the latent factors $\mathbf{l_i}$ are unknown. We assume that the $\mathbf{l_i}$ are independently and identically distributed draws from the standard normal distribution so their joint distribution **h** can be integrated out of the joint density:

$$\mathrm{prob}\left(y_j|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{X}'_{\mathbf{i}},\mathbf{d_i}\right) = \int \mathrm{f}(y_j\left|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{d_k},\mathbf{l_k}\right) * \mathbf{g}(\mathbf{X}'_{\mathbf{i}}\boldsymbol{\beta}_{\mathbf{i}} + \mathbf{Z}'_{\mathbf{m|i}}\boldsymbol{\gamma}_{\mathbf{m|i}} + \mathbf{l_i})\mathbf{h}(\mathbf{l})d\mathbf{l} \qquad (9)$$

The main computational problem, given suitable specifications for f, g, and h, is that the integral (9) does not have, in general, a closed-form solution. This difficulty is addressed by using simulation-based estimation:

$$\mathrm{prob}\left(y_j|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{X}'_{\mathbf{i}},\mathbf{d_i}\right) = \mathrm{E}\{\mathrm{f}(y_j\left|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{d_k},\mathbf{l_k}\right) * \mathbf{g}(\mathbf{X}'_{\mathbf{i}}\boldsymbol{\beta}_{\mathbf{i}} + \mathbf{Z}'_{\mathbf{m|i}}\boldsymbol{\gamma}_{\mathbf{m|i}} + \mathbf{l_i})\}$$

$$\approx \frac{1}{S}\sum_S \mathrm{f}(y_j\left|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{d_k},\widetilde{\mathbf{l_{ks}}}\right) * \mathbf{g}(\mathbf{X}'_{\mathbf{i}}\boldsymbol{\beta}_{\mathbf{i}} + \mathbf{Z}'_{\mathbf{m|i}}\boldsymbol{\gamma}_{\mathbf{m|i}} + \widetilde{\mathbf{l_{is}}}) \qquad (10)$$

Where $\widetilde{\mathbf{l_{is}}}$ is the *s*th draw of a pseudorandom number based on Halton sequences from the density h. The simulated log-likelihood function for the data is given by

$$\ln l\left(y_j|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{X}'_{\mathbf{i}},\mathbf{d_i}\right) \approx \sum \ln[\frac{1}{S}\sum_S \mathrm{f}(y_j\left|\mathbf{Z}'_{\mathbf{j|k}},\mathbf{d_k},\widetilde{\mathbf{l_{ks}}}\right) * \mathbf{g}(\mathbf{X}'_{\mathbf{i}}\boldsymbol{\beta}_{\mathbf{i}} + \mathbf{Z}'_{\mathbf{m|i}}\boldsymbol{\gamma}_{\mathbf{m|i}} + \widetilde{\mathbf{l_{is}}})] \quad (11)$$

The simulated log-likelihood function relies on an aggregate of probabilities across all farmers in this application. Provided that S is sufficiently large, maximization of the simulated log likelihood (11) is equivalent to maximizing the log likelihood.

Due to computational complexity, the framework illustrated above can only work with binary, continuous, or negative binomial outcomes. In our application on tillage outcomes, we have to model the tillage outcomes as binary (conventional tillage or conservation tillage) or as continuous variable by converting discrete tillage choice into continuous crop residue percentage variable. Detailed procedure on this transformation is introduced later in the following data section. The model is implemented using *mtreatreg* package in Stata.

IV. Data

We apply this integrated model to the Maumee River Watershed, which is the largest in the Great Lakes Region and contributing by far the largest volume of sediment into Lake Erie (Zmijewski and Becker 2010). Eight percent of land in the Maumee River Watershed is in agricultural land use, from which the nutrient and sediment loadings are contributing to excessive, harmful algal blooms and other water quality problems in Lake Erie (Reutter et al. 2011). A recent study reveals that of the entire Great Lakes Region, the Western Lake Erie drainage, including the Maumee River, has the largest amount of under-treated acres – 2.3 million acres (48% cropped acres) with a high or moderate need for additional conservation treatment to reduce nitrogen loss in subsurface flows (CEAP 2011). This makes the Maumee River Watershed an ideal laboratory to study how farmers' tillage choices under different policy scenarios are impacting the nutrient/sediment runoffs and water quality in Lake Erie.

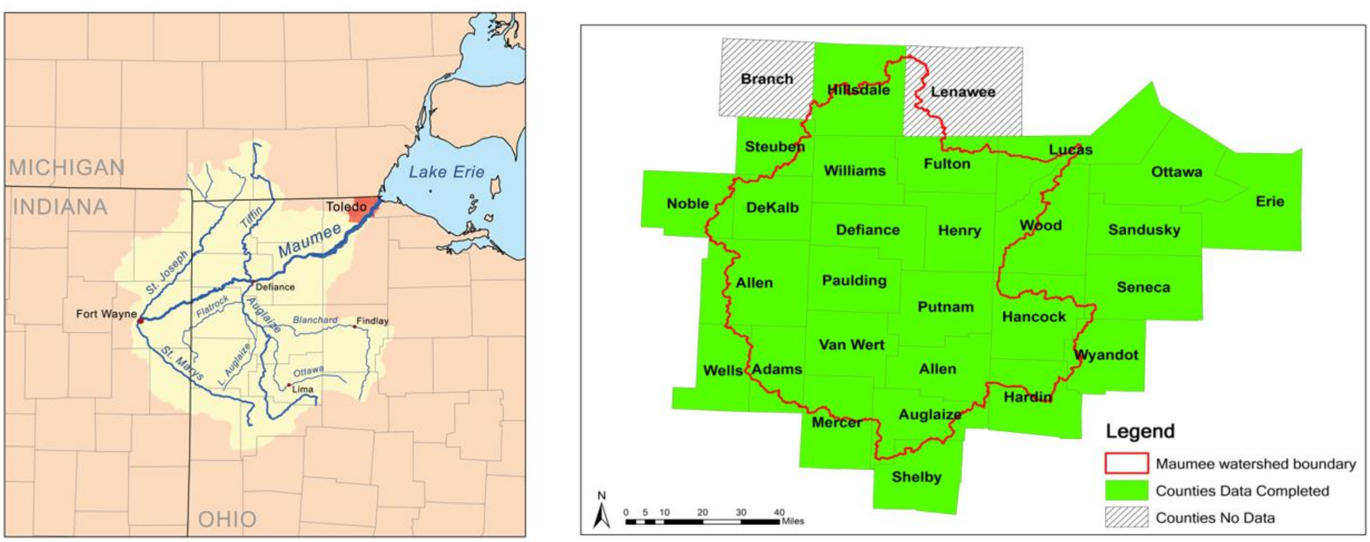

Figure 1: Study region and field boundaries data availability

We obtain the Common Land Unit (CLU) field boundaries data for almost all counties in the Maumee River watershed from USDA – Farm Service Agency. These data is comprised of

186,453 agricultural fields that span a total of six million spatially contiguous acres located either within or adjacent to the Maumee River watershed in 28 counties in Ohio, Indiana, and Michigan. The CLU field boundaries data are overlaid with NASS Cropland Data Layer 2006 to 2012 at 30m*30m or 56m*56m resolution to identify the field-level dominant land use and crop choice which has the largest area within a specific field. This yields field-level crop choices for each year from 2006 to 2012, the land use trajectories from 2006 to 2012, and the crop rotation patterns 2006-2012. We also obtain the tillage choices from overlaying the CLU data with the remote-sensed data on tillage in 2006, 2007, and 2008 from collaborators from the University of Toledo. In this paper, we mainly use crop choice and tillage choice in the year 2008, and the following table 1 summarizes the data.

		Tillage System			
		No-till	Mulch/ reduced	Conventional	Total Acres
Crop Choice	Corn	18.1%	5.5%	76.5%	962603
	Soybean	53.6%	11.3%	35.1%	1803851
	Wheat	24.6%	70.7%	4.7%	457494
	Hay & Other	68.1%	21.6%	10.4%	258936
	Total Acres	1429917	635737.2	1417230	3482885

Table 1. Total acres by crop and tillage system

After identifying the dominant tillage choice, ideally we could estimate a multinomial logit model of tillage outcomes. However, the framework illustrated in above sections can only handle binary or continuous outcome variables due to coding and computational complexity. As a result, two different tillage outcome variables are used in this paper, one is the binary variable indicating the farmer is choosing conventional till or conservation till, which includes no-till, mulch-till, and reduced-till. Another outcome variable is a continuous crop residue percentage variable. For example, if after overlaying the CLU data with remote-sensed tillage choice data, we find that 60%, 25%, and 15% of the acreage for one particular field is in no-till, mulch-till, and conventional-till respectively, we create the crop residue variable by taking a weighted average approach, where the weights are the different percentage of acreage for different tillage system. The crop residue constructed for this field is 60% * 90 + 25% * 30 + 15% *10 = 54, assuming the crop residue for these three tillage systems are 90%, 30% and 10% respectively. The limitation of this approach is that it ignores the fact that on one particular field, it is most likely there is only one tillage system, and the existence of multiple tillage systems in one field may result from the inaccuracy and resolution of the remote-sensed data.

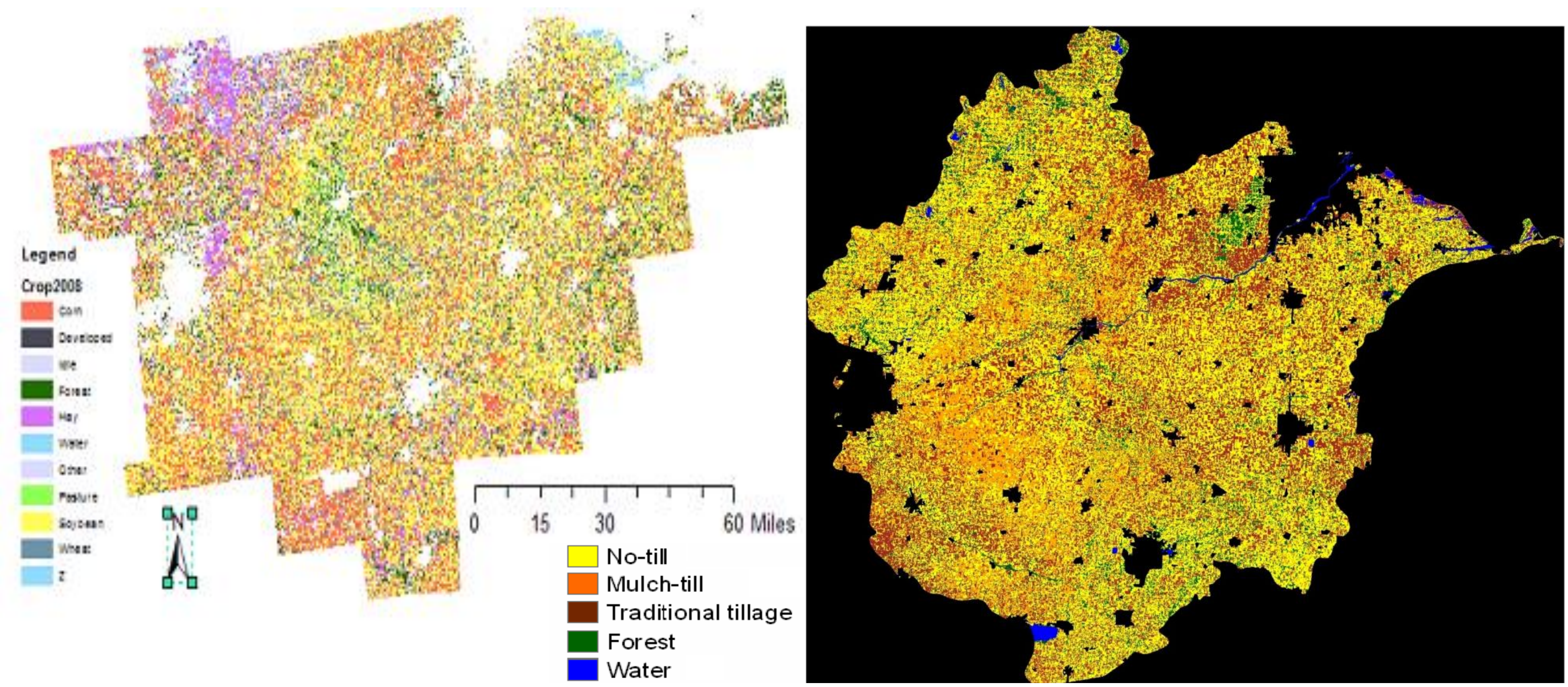

Figure 2. Crop choice and tillage choice in the Maumee River watershed

Our variable on *crop prices* come from two sources: the average futures prices for corn, soybean, and wheat were calculated as the average closing prices in February at the Chicago Board of Trade for December corn, November soybean and November wheat. In addition, we obtain average basis prices for these three crops in October for all 146 grain elevators, ethanol plants, soybean crushing facilities, and other agricultural output terminals in our study region from a private company, GeoGrain. By summing these average future prices and basis prices, we obtain a proxy for cash forward prices in October from 2006 to 2012 for these 146 spatially located agricultural markets for December corn, November soybean and November wheat. We combine data used in Wu et al. (2004) and the Ohio production budget data developed by Ward (2013) to obtain site-specific production costs, by crop and tillage practices by state, crop and previous year crop, from which we constructed another key variable *additional production costs for conventional till over conservation till*.

To capture the yield and environmental differences among different fields, physical variables reflecting land quality at individual fields are included as independent variables in the models. *Parcel size* captures the acres of agricultural fields, and *slope* is a continuous variable measured as a percentage. The differences in soil quality are captured by three technical variables: *soil available water capacity (soil awc)* is the availability water capacity of the soil layer measured in mmH_2O/mm soil; *soil organic carbon content (soil cbn)* is defined percentage of soil weight; and *soil loss potential* is defined as the soil erodibility (K) factor used in USLE equation measured in (metric ton m^2 hr)/(m^3-metric ton cm). Dummies for previous crops such as *previous crop is corn* capture the potential effects of crop rotation.

In addition, historical weather data at a 4km*4km grid level from 1980 to 2011 were obtained from Wolfram Schlenker's website. We constructed several historical average variables based on temperature and precipitation: we assume farmers' expectations of weather conditions were assumed to be constant and thus can be represented by *mean daily maximum temperature 1981 to 2011* and *mean daily minimum temperature 1981 to 2011* of the nearest weather grid. The *mean daily spring precipitation* from 1981 to 2011 is the mean daily spring precipitation from March to April, which is designed to capture the potential negative impacts of heavy spring rains on the reduction in crop planting windows, especially for corn. The *standard deviation of spring precipitation* on the other hand captures the long-run variability and seasonality of the precipitation. Another variable of interest is the *degree days over 20^oC*, a non-linear transformation of the temperature variable suggested by agronomic experiments to be a better predictor of plant growth. Roughly this variable captures the heat stress or intensity of the high temperature on plant growth, for example, a day with a maximum temperature at 32^{o}C contributes to 12 *degree days over 20^oC*. Please see Schlenker et al. (2007) for details in the construction of the variable. In addition, we have regional dummies at the crop reporting district level to capture the idiosyncratic factors specific to one specific region or state.

V. Results and Discussion

Table 2 presents the results of our treatment effects crop choice model. Note that only the coefficients of the models are shown in Table 2 and the significant ones are highlighted in bold. These are just coefficients not marginal effects, so only the signs have economic meanings. Most of the coefficients are intuitive and the base outcome is hay. The positive and significant coefficients on crop prices reveal that the higher the expected crop prices this October, the higher the probability of choosing this particular crop. The negative and significant coefficients on previous crop dummies show signs of crop rotation: in other words, if the previous year's crop is soybean, farmers are more likely to grow corn. These results reflect the fact that corn-soybean and corn-soybean-wheat rotation are the most popular cropping systems in our study region, while continuous corn and continuous soybean are not widely practiced.

Regarding the differential impacts of land quality on crop choices, our analysis show that high-quality land indicated by higher values of *soil awc* are more likely to be planted to high

	Corn	Soybean	Wheat
Crop price	**5.67E-04**	4.32E-06	**1.91E-04**
Previous crop is corn	**-8.6E-07**	2.25E-07	2.62E-07
previous crop is soybean	**2.39E-08**	**-4.64E-08**	4.18E-08
previous crop is wheat	8.23E-08	8.55E-08	**-2.01E-08**
previous crop is hay	5.97E-06	2.99E-07	-2.99E-07
Parcel size	**8.42E-07**	1.44E-08	-2.04E-08
slope	**-3.91E-07**	**-2.90E-09**	**3.33E-09**
soil water capacity	**4.21E-08**	-2.42E-08	-0.12E-08
soil organic content	-1.49E-08	-1.14E-07	4.67E-07
soil loss potential	**-9.2E-06**	**-2.87E-06**	2.19E-06
mean max temp 1986-2005	2.28E-06	2.31E-06	-3.6E-07
mean min temp 1986-2005	2.79E-06	2.60E-06	1.21E-05
degrees days over 20 C	-1.10E-05	3.58E-06	2.68E-06
mean spring precipitation	**-1.26E-05**	2.62E-07	**1.05E-06**
std dev spring precipitation	9.56E-07	8.69E-06	-9.8E-08
ag district - NE Indiana	-1.02E-08	4.48E-08	-2.18E-08
ag district - NW Ohio	4.53E-08	-7.39E-08	**4.09E-08**
ag district - W Ohio	2.65E-07	-7.65E-07	3.16E-07

Table 2: Regression results of multinomial treatment – crop choices

valued crops such as corn than to hay. Land with steeper slopes is more likely to be allocated to hay and wheat than to corn and soybeans because they are erosion-prone crops. Agricultural fields with a larger acreage are more likely to be planted in corn. Corn typically demands higher soil quality than other crops: land with higher *soil awc* and lower *soil runoff potential* are more likely to be planted in corn. Regarding the temperature and precipitation variables, our analysis reveals that heavy spring rains might shorten the window for planting and it would significantly lower the probability of growing corn. In addition, wheat is spatially concentrated in the

Northwest corner of Ohio, which comparably has a lower soil quality than some other agricultural districts such as Western Ohio.

	Tillage (binary)
Additional costs of conventional-till over conservation-till	**1.77**
slope	0.956
soil water capacity	**-5.71**
soil organic content	-0.086
soil loss potential	**0.718**
mean max temp 1986-2005	2.844
mean min temp 1986-2005	0.078
degrees days over 20 C	0.0012
mean spring precipitation	**-0.7498**
std dev spring precipitation	0.0549
ag district - NE Indiana	0.477
ag district - NW Ohio	2.178
ag district - W Ohio	-0.935
Intercept	72.17
Category_corn parcels	**-0.53**
Category_soybean parcels	**5.63**
Category_wheat parcels	0.049
Idiosyncratic latent factors_corn	**10.23**
Idiosyncratic latent factors_soybean	0.123
Idiosyncratic latent factors_wheat	**-10.65**

Table 3: Regression results of tillage outcomes

Table 3 shows the regression results of binary tillage outcomes. The significant coefficients on these category variables $\mathbf{d_i}$ for crop parcels reveal that there are significant treatment effects between tillage and crop choices. For example, they suggest that agricultural fields planted in corn are less likely to choose conservation till than the reference crop group – hay. These significant treatment effects show that if left uncontrolled, the separate estimation suffers from the endogeneity bias. The interpretations of other variables are relatively intuitive. One variable of particular interest is the *additional production costs of conventional-till over conservation-till*. Since the tillage decisions do not have an immediate impact on crop yield, these production cost differentials serve as the same as the expected crop prices in the crop choice model. Our analysis suggests that lower the costs of conservation-till compared to conventional-till, the farmers are more likely to adopt conservation-till. In addition, land with better soil quality is less likely to adopt the conservation tillage, maybe reflecting a lesser concern of the long-term soil health. The significant coefficients on idiosyncratic factors show that farmers who are more likely to grow corn relative to hay, on the basis of unobserved characteristics, are more likely to adopt the conservation-till. These significant coefficients are consistent with previous findings by Norris and Batie (1987) and Konar et al. (2012) that find individual operator characteristics are important determinants of farmers' agricultural best management practices.

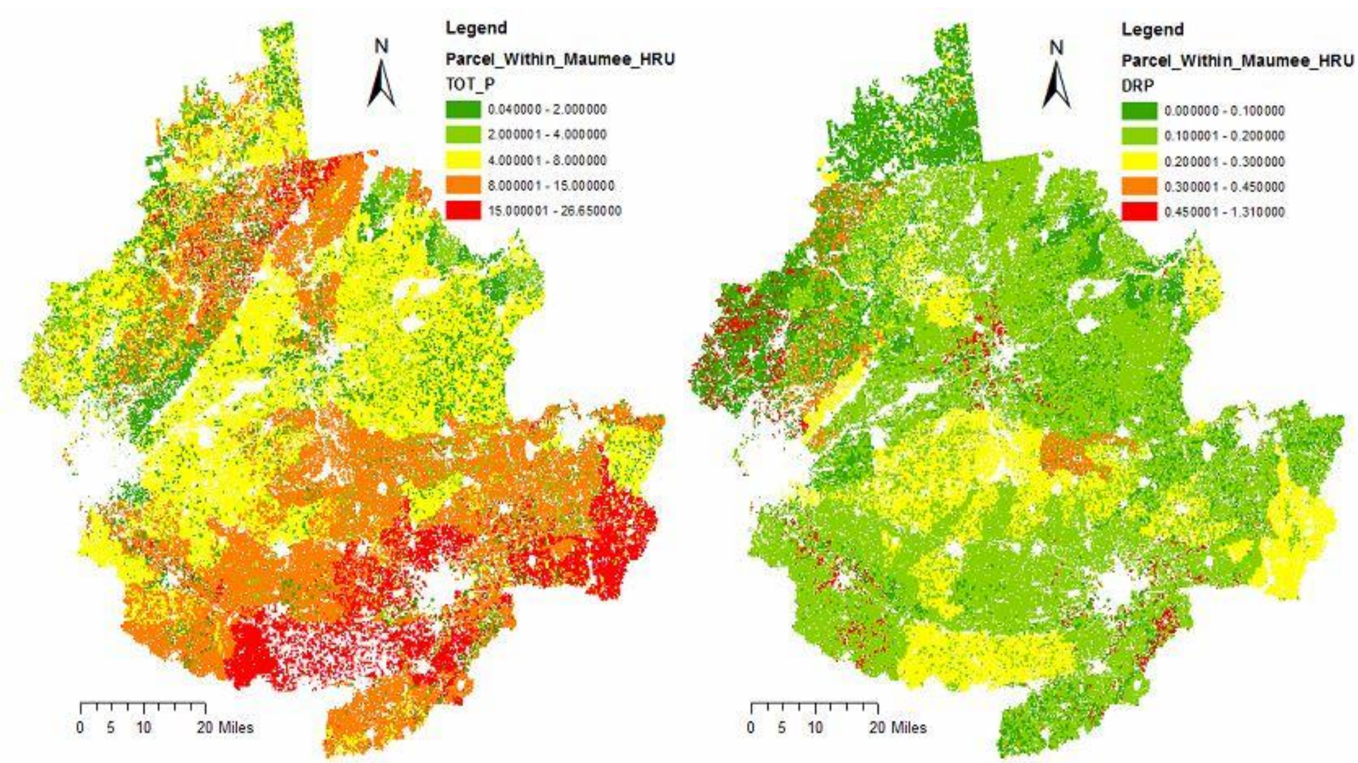

Figure 3: Field-level total phosphorus and dissolved reactive phosphorus (DRP) loadings

To analyze the environmental impacts of alternative nutrient management policies, we use a hydrological model – Soil and Water Assessment Tool (SWAT) – to translate farmers' crop and tillage outcomes to environmental impacts such as DRP loadings. SWAT is a continuous watershed model that was developed by the U.S. Department of Agriculture's (USDA) Agricultural Research Service (ARS) for assessing efficacy of agricultural and watershed management practices in reducing sediment and nutrient losses from croplands to receiving water bodies (Arnold et al. 1998). Being a direct outgrowth of various crop growth, hydrology and chemical transport models developed by ARS over the last 30 years, SWAT currently embodies a wide range of capabilities to quantify agricultural nonpoint source pollution under various cropland management practices. Figure 3 presents the per-acre total phosphorus and DRP loadings from SWAT based on the corresponding field-level crop and tillage choices. Based on these per-acre phosphorus loadings by crop and tillage systems in each of the 1200 spatially delineated hydrological resource unit (HRU), we obtain the HRU-specific export coefficients in terms of phosphorus loadings that make the translation of agricultural land management outcomes into environmental quality variables much easier.

Our model could analyze the impacts of alternative policies such as emission taxes – a first best tax scheme based on field-specific phosphorus loadings assuming the tillage outcomes can be accurately inferred from remote-sensed data; or voluntary payments scheme to encourage adoption of conservation tillage; or fertilizer tax which has an indirect effect on cost-savings yet yield-reducing conservation-till given rises in production costs and potential yield reductions due to decreased fertilizer application. Due to time constraints and computational complexity, we were not able to analyze these policies. However, we did run a small exercise which analyzes the impacts of incorporation for small no-till fields on DRP loadings. Typically phosphorus fertilizer application is broadcast, which leaves a uniform distribution of phosphorus on the soil surface, and is easy to apply. However, this application method can result in those nutrients stratifying and accumulating in the top 2 or 3 inches of the soil, prone to phosphorus runoff in events of heavy rains. Agronomists advocate incorporation instead, which incorporates the fertilizer into the soil through strip till or banding and thus alleviate the stratification problem and reduce the nutrient runoffs (Sharpley et al. 2006). Figure 4 shows the map of DRP loadings across the watershed if all the agricultural fields currently adopting no-till and with acreage of less than 20 acres all switch from broadcast application to incorporation. Comparing Figure 3(b) and Figure 4 reveals a noticeable reduction in the DRP loadings at least in the Blanchard River region (in the middle of the graph) just through incorporation. This is reasonable because we assume the changes only occur in small parcels which do not require long operation times.

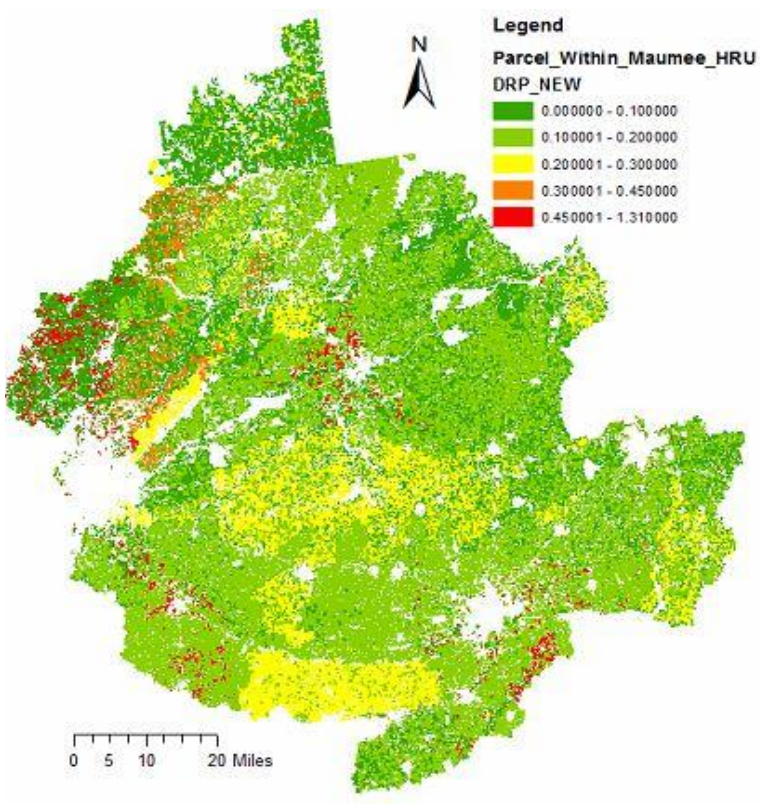

Figure 4: DRP loadings if all no till, <20 acre fields switch from broadcast to incorporation

VI. Conclusion and Next Steps

The record-setting harmful algal bloom in Lake Erie in 2011 has renewed the interests among policy makers, researchers and the public in the linkages between non-point agricultural nutrient runoffs and downstream ecosystem conditions (Michalak et al. 2013). In this paper, with a focus on the largest watershed in the Great Lakes region, we developed a spatially-explicit model of farmers' crop and tillage decisions and analyzed the impacts of nutrient management policies in phosphorus loadings. Using a multinomial treatment effects simulated maximum likelihood model, we controlled for the endogeneous linkages between crop choices and tillage choices which were not accounted for in previous separate estimation of these two decisions. Our spatially contiguous parcel data allows for integration with watershed hydrological model which facilitates the benefit-cost analysis of alternative policies such as spatial-targeting policy in terms of their impacts on nutrient loadings.

Due to computational complexity imbedded in the simulated-maximum likelihood estimation, we were not able to analyze many relevant policy scenarios. The model is difficult to converge; as a result the results presented here are very preliminary. As the next steps, first we plan to investigate the possibility of estimating a simpler model of nutrient management plan, which treats crop choice and tillage choice as a bundle. Secondly, we plan to analyze the impacts of other nutrient management policies, such as first-best spatially targeted DRP emission taxes and voluntary payment schemes.

References:

Arnold, J. G., Allen, P. M., 1999. Automated methods for estimating baseflow and ground water recharge from streamflow records. *Journal of American Water Resource Assocaition* 35, 411–424.

Conservation Effects Assessment Project (CEAP). 2011. "Assessment of the Effects of Conservation Practices on Cultivated Cropland in the Great Lakes Region."

Diaz, R.J., and R. Rosenberg. 2008. "Spreading Dead Zones and Consequences for Marine Ecosystems." *Science*, 321(5891): 926-929.

Deb, P., and P. K. Trivedi. 2006. "Maximum Simulated Likelihood Estimation of a Negative Binomial Regression Model with Multinomial Endogenous Treatment." *The Stata Journal*, 6(2): 246-255.

Featherstone, A. M. and B. K. Goodwin. 1993. "Factors Influencing a Farmer's Decision to Invest in Long-term Conservation Improvements." *Land Economics,* 69 (1), 67–81.

Gould, B.W., Saupe, W.E., Klemme, R.M., 1989. "Conservation Tillage: the Role of Farm and Operator Characteristics and the Perception of Soil Erosion." *Land Economics*, 65 (2), 167–182.

Groisman, P., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, V. N. Razuvaev, 2005. "Trends in Intense Precipitation in the Climate Record." *Journal of Climate*, 18, 1326–1350.

Knowler, D., and B. Bradshaw. 2007. "Farmers' Adoption of Conservation Agriculture: A Review and Synthesis of Recent Research." *Food Policy*, 32, 25-48

Konar, A. , B. E. Roe, and E. G. Irwin. 2012. "Do Farmers Have Heterogeneous Preferences for the Environment and Does It Matter? A Latent-Class Approach to Explaining Field-Level Tillage Choices." Paper selected for presentation at the Agricultural & Applied Economics Association 2012 Annual Meeting, Seattle, WA, August 12-14, 2012.

Kurkalova, L., C. L. Kling, and J. Zhao. 2006. "Green Subsidies in Agriculture: Estimating the Adoption Costs of Conservation Tillage from Observed Behavior." *Canadian Journal of Agricultural Economics*, 54, 247–67.

Lewis, J. L., A. J. Plantinga, and J. Wu. 2009. "Targeting Incentives to Reduce Habitat Fragmentation." *American Journal of Agricultural Economics* 91(4): 1080-96.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Michalak, A. M., Anderson, E. J., Beletsky, D., Boland, S., Bosch, N. S., Bridgeman, T. B., Chaffin, J. D., Cho, K., Confesor, R., Daloglu, I., DePinto, J. V., Evans, M. A., Fahnenstiel, G. L., He, L., Ho, J. C., Jenkins, L., Johengen, T. H., Kuo, K. C., LaPorte, E., Liu, X., McWillaims,

M. R., Moore, M. R., Posselt, D. J., Richards, R. P., Scavia, D., Steiner, A. L., Verhamme, E., Wright, D. M., and Zagorski, M. A. 2013. Record-Setting Algal Blooms in Lake Erie Caused by Agricultural and Meteorological Trends Consistent with Expected Future Conditions, *Proceedings of National Academy of Sciences USA* 110(16): 6448-52.

Norris, P. E. and S. S. Batie. 1987. "Virginia Farmers' Soil Conservation Decisions: An Application of Tobit Analysis." *Southern Journal of Agricultural Economics*, 19, 79–90.

Schlenker, W., W. M. Hanemann, and A. C. Fisher. 2007. "Water Availability, Degree Days, and the Potential Impact of Climate Change on Irrigated Agriculture in California." *Climate Change* 81 (1): 19-38.

Sharpley, A.N., T. Daniel, G. Gibson, L. Bundy, M. Cabrera, T. Sims, R. Stevens, J. Lemunyon, P. Kleinman, and R. Parry. 2006. "Best Management Practices to Minimize Agricultural Phosphorus Impacts on Water Quality," USDA-Agricultural Research Service, ARS–163.

Wu, J., R. M. Adams, C. L. Kling, and K. Tanaka. 2004. "From Microlevel Decisions to Landscape Changes: An Assessment of Agricultural Conservation Policies." *American Journal of Agricultural Economics*, 86, 26-41.

Wu, J., and W. G. Boggess. 1999. "The Optimal Allocation of Conservation Funds." Journal of Environmental Economics and Management 38: 302-21.