Estimation of a Discrete Choice Model When Individual Choices Are Not Observable

Lyubov A. Kurkalova* and Sergey Rabotyagov

Center for Agricultural and Rural Development, Iowa State University

*Mailing Address:
560A Heady Hall
CARD
Iowa State University
Ames, IA 50011-1070
Telephone: (515) 294-7695
Fax: (515) 294-6336
E-Mail: lyubov@iastate.edu

DRAFT
07/15/2004
Please do not cite without authors' permission

Estimation of a Discrete Choice Model When Individual Choices Are Not Observable

Abstract

This paper presents an econometric technique for circumventing the lack of individual choice data in a framework of binary choice model by utilizing aggregate choice data. The probability of observing a certain number of individuals making choice A out of the total number of individuals in a group is presented as a sum of probabilities of disjoint events, in which some individuals are picked to make choice A, and others are not. These probabilities are then used to form a likelihood function. The model, which is estimated using the method of maximum likelihood, performs favorably in an application to real discrete choice data.

Estimation of a Discrete Choice Model When Individual Choices Are Not Observable

1. Introduction

In many areas of economics, interest centers on the estimation of discrete choice models and related issues. However, the individual choice data may be costly to collect. Also, the information on the choices made by individuals may be not available due to confidentiality reasons. This is a common situation in many areas of economics. In agricultural economics, it often happens in modeling land use decisions (e.g., Miller and Plantinga, 1999, Plantinga and Wu, 2003). This paper proposes an econometric technique for circumventing the lack of individual choice data in a framework of a discrete choice model by utilizing aggregated choice data.

The motivation for the study comes from the problem of modeling tillage choices made by Midwestern farmers. Understanding farmers' decisions on adoption of conservation practices is important for predicting the effectiveness of agricultural and environmental policies such as the conservation component of the recently passed 2002 Farm Bill. To model the policies on a large scale, it is desirable to use spatially detailed data representing large regions. One such data set is the National Resource Inventory (NRI) (Nusser and Goebel, 1997), which contains information on cropping history, conservation practices, and soil characteristics for some 800,000 physical points in the U.S. The 1992 NRI data have been successfully used to econometrically estimate a model of farmers' choices between conventional and conservation tillage in Iowa and to calculate the necessary adoption subsidy to induce farmers to switch to conservation tillage practices (Kurkalova et al., 2003). A similar model could become a basis for updating the analysis of the conservation tillage adoption decisions as newer data becomes available. However, estimation of a similar model using the newest 1997 NRI is not possible because individual tillage choices are not reported in the 1997 NRI.

In the present paper, we derive an econometric method, which allows for the discrete choice model to be estimated when the information on the attributes of the choices and decision-makers is available at the individual level, but the information on the choices made by the individuals is aggregated across groups of individuals. We derive the likelihood of the model and estimate the model using the method of maximum likelihood. We illustrate the proposed model by applying it to estimation of the determinants of the choice of conservation tillage.

The remainder of the paper is organized as follows: The second section presents and motivates the basic model used in the paper. The third section discusses the results of estimation of the model and compares them to those obtainable when the individual choice information is available to the researchers. The forth section concludes and comments on possible extensions.

2. Model

We consider a set of N observations corresponding to binary choices made by N different individuals. The choice is described by the variable Y_i, which takes on the value of 1 or 0 depending of which of the two alternatives available is chosen, i.e.

$$Y_i = \begin{cases} 1 & if\ \ alternative\ A\ is\ chosen, \\ 0 & if\ \ alternative\ B\ is\ chosen, \end{cases} \quad (i = 1, \ldots, N)$$

We also assume that an economic model postulates that the choice is a function of K covariates, representing the attributes of the decision maker and/or the choices themselves. The covariates are given by the vector $\mathbf{x}_i = (x_{1i},...,x_{Ki})$, $i = 1,...,N$. In the standard econometrics setting, the exact relationship between Y_i and $\mathbf{x}_i$ is assumed known to the individuals making the choice, but unobservable to the researcher. As a consequence, the probability of choosing the alternative A from the researcher's perspective is specified as

$$Pr[Y_i = 1] = Pr[\varepsilon_i < h(\boldsymbol{\beta}, \mathbf{x}_i)],$$

where ε_i represents researcher's error arising from measurement errors, specification errors due to functional form choice, and the effect of the omitted variables, $h(.)$ is a specified function of its parameters, and vector $\boldsymbol{\beta}$ represents the parameters of interest. Alternative choices on the distribution of ε_i lead to alternative models of discrete choice, e.g., logit and probit. In this model, we assume ε_i to be logistically distributed and independent across i's, and a linear functional form for $h(.)$, i.e. $h(\boldsymbol{\beta}, \mathbf{x}_i) = \boldsymbol{\beta}'\mathbf{x}_i$ and

$$Pr[Y_i = 1] = \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta}'\mathbf{x}_i)}. \quad (1)$$

As it will be seen later, neither of these two assumptions is crucial for the proposed technique. In the discussion that follows, we always assume that the data on the covariates $\mathbf{x}_i$ are available for all $i = 1,...,N$, and consider alternative assumptions on the availability of the data y_i on the choices Y_i, $i = 1,...,N$. When the data y_i on the choices Y_i are observed for all i's, the likelihood function (for the i'th observation) for the model (1) can be written as

$$L(\boldsymbol{\beta} \mid y_i, \mathbf{x}_i) = \left(\frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{1 + \exp(\boldsymbol{\beta}'\mathbf{x}_i)} \right)^{y_i} \left(\frac{1}{1 + \exp(\boldsymbol{\beta}'\mathbf{x}_i)} \right)^{1-y_i}, \quad (2)$$

and the model (2) can be easily estimated using the method of maximum likelihood.

Our study pertains to the situation when less information is available to researchers: specifically, we assume that $\mathbf{x}_i$'s are still observed for all i's, but the observations on the Y_i are not observed at the individual level. Rather, the averages of the Y_i's over certain groups of individuals are observed. That is, the observations $\overline{y}^{G_j}$ are available on the random variable $\overline{Y}^{G_j} \equiv \frac{1}{N^{G_j}} \sum_{i \in G_j} Y_i$, where G_j's are mutually exclusive non-empty subsets of $\{1,...,N\}$ such that $\bigcup_j G_j = \{1,...,N\}$, and N^{G_j} is the number of observations in group G_j, j=1,…, J.

If the model we were considering were linear in the parameters of interest $\boldsymbol{\beta}$, this structure of the data would not create a serious problem for recovering the parameters. Indeed, if a linear counterpart of model (1) is given by

$$Y_i = \boldsymbol{\beta}'\mathbf{x}_i + \eta_i, \tag{1-lin}$$

where η_i are i.i.d. error terms, then one can estimate the parameters $\boldsymbol{\beta}$ by fitting the model with aggregate data, i.e. by fitting the model

$$\overline{Y}^{G_j} = \boldsymbol{\beta}'\overline{\mathbf{x}}^{G_j} + \overline{\eta}^{G_j},$$

where $\overline{\mathbf{x}}^{G_j} \equiv \frac{1}{N^{G_j}} \sum_{i \in G_j} \mathbf{x}_i$ and $\overline{\eta}^{G_j} \equiv \frac{1}{N^{G_j}} \sum_{i \in G_j} \eta_i$.

The nonlinear counterpart of the aggregated data model considered in the literature (e.g., Wu and Plantinga) is given by

$$\overline{Y}^{G_j} = \frac{\exp\left(\boldsymbol{\alpha}'\overline{\mathbf{x}}^{G_j}\right)}{1+\exp\left(\boldsymbol{\alpha}'\overline{\mathbf{x}}^{G_j}\right)} + \xi_j \quad , \tag{3}$$

where ξ_j are i.i.d. error terms. While this model is useful for explaining and predicting *grouped* choices, it is not immediately useful for explaining and predicting *individual* choices. This is so due to the nonlinearity of postulated relationship: the parameters $\boldsymbol{\alpha}$ in (3) cannot be interpreted as parameters $\boldsymbol{\beta}$ in (1). In contrast, we propose and implement a new conceptual model that allows recovery of the parameters $\boldsymbol{\beta}$ explaining the choices by individuals corresponding to model (1).

We begin by noting that the probability that $\overline{Y}^{G_j}$ is equal to $\overline{y}^{G_j}$ can be interpreted as the probability that $NA^{G_j} \equiv \overline{y}^{G_j} N^{G_j}$ individuals in group G_j choose the alternative A. Then this probability can be represented as the sum of the probabilities of disjoint events, in each of which exactly NA^{G_j} of N^{G_j} individuals choose alternative A. Most importantly, the setup of the model (1) can be preserved, and this probability can be expressed in terms of the original parameters $\boldsymbol{\beta}$ and data $\mathbf{x}_i$'s.

To illustrate, consider the case of $N^{G_j} = 3$, $\overline{y}^{G_j} = 1/3$, and thus $NA^{G_j} = 1$. Then

$$
\begin{aligned}
&\Pr\left[\overline{Y}^{G_j} = 1/3\right] \\
&\quad = \Pr\left[1 \text{ } \textit{out of } 3 \text{ } \textit{individuals in group } G_j \textit{ choose } A\right] \\
&\quad = \Pr\left[1\textit{st chooses } A,\ 2\textit{nd choses } B,\ \textit{and } 3\textit{nd chooses } B\right] \\
&\quad + \Pr\left[1\textit{st chooses } B,\ 2\textit{nd choses } A,\ \textit{and } 3\textit{nd chooses } B\right] \\
&\quad + \Pr\left[1\textit{st chooses } B,\ 2\textit{nd choses } B,\ \textit{and } 3\textit{nd chooses } A\right] \\
&= \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_1)}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_1)} \bullet \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_2)} \bullet \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_3)} \\
&+ \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_1)} \bullet \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_2)}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_2)} \bullet \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_3)} \\
&+ \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_1)} \bullet \frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_2)} \bullet \frac{\exp(\boldsymbol{\beta}'\mathbf{x}_3)}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_3)}.
\end{aligned}
$$

As can be easily seen, the resulting probability is expressed in terms of the individual choice model parameters $\boldsymbol{\beta}$, individual choice covariates data $\mathbf{x}_i$, and grouped choice data $\bar{y}^{G_j}$.

For generic values of $\bar{y}^{G_j}$ and N^{G_j}, the corresponding probability is

$$\Pr\left[\bar{Y}^{G_j} = \bar{y}^{G_j}\right] =$$

$$\Pr\left[\begin{matrix}\bar{y}^{G_j} N^{G_j} \ \textit{out of } \ N^{G_j} \\ \textit{individuals chose } A\end{matrix}\right] = \sum_{\sum_{i=1}^{N^{G_j}} \delta_i = \bar{y}^{G_j} N^{G_j}} \prod_{i=1}^{N^{G_j}} \left(\frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_i)}\right)^{\delta_i} \left(\frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_i)}\right)^{1-\delta_i},$$

where the δ_i's take on the values of 0 or 1 and play the role of the unobserved information on the individual choices y_i. The last expression provides the basis for forming the likelihood function (for the j-th group of observations) corresponding to model (1) with the modified data availability structure:

$$L\left(\boldsymbol{\beta} \mid \bar{y}^{G_j}, N^{G_j}, \mathbf{x}_i\left(i \in G_j\right)\right) = \sum_{\sum_{i=1}^{N^{G_j}} \delta_i = \bar{y}^{G_j} N^{G_j}} \prod_{i=1}^{N^{G_j}} \left(\frac{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_i)}\right)^{\delta_i} \left(\frac{1}{1+\exp(\boldsymbol{\beta}'\mathbf{x}_i)}\right)^{1-\delta_i}. \quad (4)$$

Next section presents estimation of the resulting model using the method of maximum likelihood.

3. **Examples of estimation on simulated data**

In order to test the method performance we conduct two simulation exercises.

3.1. Example 1: simulated data

To construct the data for simulation exercise, first the "independent variables" are drawn from a series of uniform distributions. Second, a logistically distributed random errors

ε_i's are drawn. Finally, the binary response variable is simulated for a simple linear model of the following form: $y_i = \beta' x_i + \varepsilon_i$. Then

$\Pr[y_i = 1] = \Pr[\varepsilon \le \beta' x_i] = \dfrac{\exp(\beta' x_i)}{1 + \exp(\beta' x_i)}$. Having picked (arbitrary) values for the parameter vector β, we compute the above probability and assign y_i a value of 1 if the resulting probability is greater than 0.5 and a value of zero otherwise. The choice of the parameter vector values, although arbitrary in principle, in this case was made so that the average probability of y_i being 1 was roughly 30 percent. This was done in keeping in mind the intended application of the proposed method, that is, to estimate a model of conservation tillage adoption, and the average adoption rate in real data is quite close to 30 percent. The simulated data set was then used to estimate two econometric models. The first model is a standard logit model that was estimated using the correct model specification. The second model estimated was the aggregated data model. The observations were grouped and we pretend that we do not know the response variable value for each observation but instead know the number of observations in a group and the number of observations for which the response variable y_i is equal to 1 (e.g., 3 out of 5 observations). The standard logit model was used as a benchmark for judging the estimates obtained using the aggregated data method. Table 1 presents the results of this simulation exercise.

This result demonstrates that even when the individual response variables are not directly available, parameter estimates of a discrete choice model can still be recovered provided that the aggregated data is based on the actual individual-level data and that the number of observations is sufficiently large.

3.2: Example two: grouped real data

To illustrate the proposed technique in a real-world setting we apply it to the 1992 NRI data set used by Kurkalova et al. (2003). We begin by briefly introducing the economic model of conservation tillage adoption.

The original model of Kurkalova et al.(2001) assumes that a farmer will adopt conservation tillage if the expected annual net return to using conservation tillage (π_1) exceeds the expected net return from using conventional tillage (π_0) plus the premiums associated with uncertainty P. A farmer typically demands a premium for adopting conservation tillage due either to risk aversion or to real options. The model of Kurkalova et al. (2003) both (a) explicitly incorporates an adoption premium to reflect risk aversion and real options, and (b) allows recovery of the individual parameter values, overcoming the usual difficulty associated with standard discrete choice models, i.e., that only the ratios of the coefficients to the standard deviation can be recovered. Assuming an additive error and that the expected net returns from conventional tillage is known to the farmer, they write the probability of adoption as

$$\begin{aligned}\Pr[adopt] &= \Pr[\pi_1 \geq \bar{\pi}_0 + P + \sigma\varepsilon] \\ &= \Pr[\boldsymbol{\gamma}'\mathbf{x} \geq \bar{\pi}_0 + P + \sigma\varepsilon] \\ &= \Pr\left[\varepsilon \leq \frac{\boldsymbol{\gamma}'\mathbf{x}}{\sigma} - \frac{\bar{\pi}_0}{\sigma} - \frac{P}{\sigma}\right],\end{aligned}$$

where P represents the premium as a function of its explanatory variables and the bar on $\bar{\pi}_0$ denotes that this variable is known.

In this formulation, recovery of the standard deviation multiplier σ is straightforward as it will simply be the inverse of the coefficient estimated on $\bar{\pi}_0$. Thus, by adding information to the model in the form of the expected net profits from conventional tillage, it is possible to estimate the standard error, in turn allowing recovery of the specific parameter values for $\boldsymbol{\gamma}$.

Assuming a linear form for the net returns to conservation tillage, linearly multiplicative form for the premium, and logistic distribution for the standard econometric stochastic component, Kurkalova et al. (2003) estimate the model using the method of maximum likelihood. Specifically, they estimate

$$\Pr\left[adopt\right] = \Pr\left[\pi_{1,j} \geq \overline{\pi}_{0,j} + P_j\right], \quad j = cn, sb, oth,$$

where

$$\pi_{1,j} = \gamma_{0,cn} \cdot I_{cn} + \gamma_1 \cdot SLOPE + \gamma_2 \cdot PM + \gamma_3 \cdot AWC + \gamma_4 \cdot TMAX + \gamma_5 \cdot TMIN + \gamma_6 \cdot PRECIP$$
$$+ \gamma_7 \cdot TENANT + \sigma_\varepsilon \cdot \varepsilon ,$$

and

$$P_j = \sigma_{precip}(\alpha_{1,j} + \alpha_{2,j} \cdot \overline{\pi}_{0,j}$$
$$+ \alpha_{3,j} \cdot OFFFARM + \alpha_{4,j} \cdot TENANT + \alpha_{5,j} \cdot AGE + \alpha_{6,j} \cdot MALE),$$

and the definitions of the variables are provided in Table 2. Details on interpretation of the effects of the explanatory variables are provided in Kurkalova et al. (2003).

To illustrate our proposed technique, we completely preserve the described economic model setup. However, on the econometrics side, we assume that less information is available to researchers. Specifically, we treat the data set as if individual choices are not observable, but the grouped ones are. To construct the "observable" choice data, we grouped the original 1,339 observations into 240 groups by county and crop information available in the data set. The results of estimation of this "reduced information" model are provided in Table 3 together with the original results of Kurkalova et al. (2003) for the fully observed individual choice data.

Estimates of all the variables in the model are very similar between our "reduced information" and "full information" cases. The effects of all variables are same in sign and of similar magnitude. Moreover, no estimates reported in Kurkalova et al. (2003) that were statistically significant lost significance during the new estimation procedure. Again, the aggregated data model performs well in a situation when the data is 'artificially' aggregated on the basis of the actual individual choice data.

4. Application

A potentially very useful application of the aggregated data model is to the 1997 NRI data, as this would utilize the most recent NRI data available and make the model estimates more valuable for policy analysis. As with the 1992 version, the 1997 version of the NRI provides information on the natural resource characteristics of the land (soil properties and slope), and the crop grown (1997 and 1996 seasons). To complete the data set, we follow the procedures of Kurkalova et al. (2003), and supplement the data with constructed net returns, weather and climate data, and county average indicators of farm operator characteristics. However, there exists a significant problem in attempting to use the 1997 NRI survey to estimate a model of conservation tillage adoption – point level information on the adoption of conservation tillage is not available in the 1997 NRI.

The grouped information on conservation tillage choices, the proportion of cropland in conservation tillage, comes from the data assembled by the Conservation Tillage Information Center (CTIC, 2000) as well as from the ARMS data (http://www.ers.usda.gov/Briefing/ARMS). These proportions are used to infer how many farms that produce a specific crop in a county do so employing conservation tillage practices. The potential problem with such an approach, however, is that the aggregated adoption data is not directly computed from the underlying individual choice observations, but instead is constructed based on county-wide estimates from CTIC and ARMS datasets. In a sense there is no guarantee anymore that the specified likelihood function is going to be well-behaved. This concern unfortunately did not fail to materialize. While we continue working on the estimation, so far we have not been successful in estimating the full adoption model using 1997 NRI data with the same degree of parameter precision. However, some reduced models were relatively more successful. The last two columns of Table 3 show the estimated coefficients for the conservation tillage adoption model when the aggregated data was constructed using the 1997 NRI survey, CTIC and ARMS conservation tillage estimates.

4. Conclusions

In this paper, we propose an econometric technique for recovering the parameters describing individual choices when only grouped data are available on the choices made. The model, which is estimated using the method of maximum likelihood, performs encouragingly well in an application to simulated data, as well as to simulated grouped 1992 NRI data on adoption of conservation tillage.

The technique developed allows not only for estimation of a discrete choice model of conservation tillage adoption using the newer 1997 NRI data, but can also be used in other situations of discrete choice modeling when information on the characteristics of the decision-makers and the determinants of their choices is available at the individual level, but the choice information is aggregated because of, for example, confidentiality reasons.

Acknowledgements

The authors acknowledge helpful discussions with Wayne Fuller, Michael W. McCracken, Ivan Jeliazkov, and Joe Herriges.

References

Conservation Tillage Information Center (CTIC). 2000. *2000 Crop Residue Management Survey*, National Association of Conservation Districts, West Layfayette, IN http://www.ctic.purdue.edu/CTIC/CTIC.html. (Accessed May 2001.)

Kurkalova, L.A., C.L. Kling, and J. Zhao. 2003. "Green Subsidies in Agriculture: Estimating the Adoption Costs of Conservation Tillage from Observed Behavior," *CARD Working Paper 01-WP 286*, March 2003 http://www.card.iastate.edu/publications/DBS/PDFFiles/01wp286.pdf.

Miller, D.J., and A.J. Plantinga. 1999. "Modeling Land Use Decisions with Aggregate Data." *American Journal of Agricultural Economics* 81(2): 180-194.

Nusser, S. M., and J.J. Goebel. 1997. "The National Resources Inventory: A Long-Term Multi-Resource Monitoring Programme." *Environmental and Ecological Statistics* 4: 181-204.

Plantinga, A.J., and J. Wu. 2003. "Co-Benefits from Carbon Sequestration in Forests: Evaluating Reductions in Agricultural Externalities from an Afforestation Policy in Wisconsin." *Land Economics*, 79(1): 74-85.

Table 1. Estimation of discrete choice model with complete and grouped choice data

	True parameter value	Logit estimates (1000 obs)	Aggregate model estimates (1000 obs)	Logit estimates (10000 obs)	Aggregate model estimates (10000 obs)
β_0	-1.0	-1.1129	-0.9705	-0.9612	-0.9165
		(0.147)	(0.3069)	(0.0366)	(0.0555)
β_1	3.0	4.6532	8.8613	2.3622	2.3939
		(1.4498)	(2.7083)	(0.2994)	(0.3947)
β_2	2.0	2.6307	2.3968	1.9642	1.7409
		(0.3005)	(0.9109)	(0.0452)	(0.1074)
β_3	3.0	2.8904	2.9484	2.8139	2.669
		(0.2974)	(0.8886)	(0.0678)	(0.1453)
β_4	-1.0	-1.4141	-2.5936	-0.9103	-0.7026
		(0.2853)	(0.9211)	(0.0514)	(0.0911)
β_5	-2.0	-2.2995	-2.3694	-1.8091	-1.7655
		(0.3071)	(0.7264)	(0.0576)	(0.102)

All estimates are significant at 1% level of significance.
Standard errors are in parenthesis

Table 2. Definition of variables and summary statistics [i]

Notation	Description	Units	Sample Mean	Sample St. Dev.
	Conservation tillage (1-yes, 0-no)	Number	0.63	0.48
I_{cn}	Corn (1-corn, 0-soybeans or other crop)	Number	0.57	0.50
$\bar{\pi}_{0,cn}$	Net returns to conventional tillage, corn [ii]	$ per acre	145	23
$\bar{\pi}_{0,sb}$	Net returns to conventional tillage, soybeans [iii]	$ per acre	109	14
$\bar{\pi}_{0,oth}$	Net returns to conventional tillage, other crops [iv,v]	$ per acre	93	43
SLOPE	Land slope	Percent	4.1	3.9
PM	Soil permeability	Inches per Hour	1.7	2.2
AWC	Soil available water capacity	Percent	18.5	2.8
TMAX	Mean of daily maximum temperature during the corn growing season	Fahrenheit	78.7	1.8
TMIN	Mean of daily minimum temperature during the growing season	Fahrenheit	55.6	2.0
PRECIP	Mean of daily precipitation during the growing season	Inches	0.141	0.012
σ_{precip}	Standard deviation of daily precipitation during the growing season	Inches	0.331	0.027
OFFFARM	Proportion of operators working off-farm to the total number of farm operators in the county	Number	0.471	0.055
TENANT	Proportion of harvested cropland operated by tenants to the total county harvested cropland	Number	0.199	0.050
AGE	County average farm operator age	Years	50.2	1.8
MALE	Proportion of male operators to the total number of farm operators in the county	Number	0.9774	0.0096

[i] 1,339 observations
[ii] 762 observations
[iii] 475 observations
[iv] wheat, or hay
[v] 102 observations

TABLE 3. Maximum Likelihood Estimates of the Adoption Model[i]

Variable(s)	Parameter	Aggregated data model (4), 1992 NRI data	Kurkalova et. al., model (2)	Aggregated data model, 1997 NRI data	Reduced model, 1997 NRI data
Net returns to conservation					
I_{cn}	$\gamma_{0,cn}$	44.7*	41*	1.28	0.59
		(6.9)	(11)	(4.96)	(2.31)
SLOPE	γ_1	0.56*	0.22***	0.60	0.44
		(0.18)	(0.12)	(1.07)	(0.81)
PM	γ_2	0.85**	0.63**	7.33	15.32
		(0.37)	(0.31)	(15.55)	(12.32)
AWC	γ_3	0.87*	0.73**	-79.97	-66.49
		(0.32)	(0.29)	(268.14)	(177.22)
TMAX	γ_4	2.32*	2.57*	-2.01	-0.53
		(0.48)	(0.68)	(3.53)	(2.25)
TMIN	γ_5	-2.42*	-2.48*	4.80	2.68
		(0.54)	(0.72)	(5.25)	(2.96)
PRECIP	γ_6	107**	76	-5.32	-2.98
		(47)	(69)	(7.05)	(4.40)
TENANT	γ_7	223*	194**	0.58	0.61
		(58)	(92)	(4.59)	(0.65)
	σ_ε	5.69*	6.0*	4.55	3.97**
		(0.12)	(1.6)	(3.85)	(1.93)
Premium					
$\sigma_{precip} \cdot I_{cn}$	$\alpha_{1,cn}$	1410*	1400*	-665.07	_
		(298)	(411)	(937.11)	_
$\sigma_{precip} \cdot I_{sb}$	$\alpha_{1,sb}$	1032**	1123*	-1530.85	_
		(285)	(432)	(1321.36)	_
$\sigma_{precip} \cdot I_{oth}$	$\alpha_{1,oth}$	1145***	770	718.69	_
		(646)	(557)	(1162.09)	_
$\sigma_{precip} \cdot \overline{\pi}_{0,cn}$	$\alpha_{2,cn}$	-2.812*	-2.79*	-7.88**	-7.66*
		(0.071)	(0.11)	(3.49)	(1.86)
$\sigma_{precip} \cdot \overline{\pi}_{0,sb}$	$\alpha_{2,sb}$	-3.27*	-3.32*	-7.66*	-8.02*
		(0.13)	(0.19)	(2.50)	(0.91)
$\sigma_{precip} \cdot \overline{\pi}_{0,oth}$	$\alpha_{2,oth}$	-3.52*	-3.00*	-11.48**	-9.93*

		(0.38)	(0.22)	(4.90)	3.42
$\sigma_{precip} \cdot TENANT \cdot I_{cn}$	$\alpha_{3,cn}$	688*	607**	273	_
		(175)	(274)	(4837)	_
$\sigma_{precip} \cdot TENANT \cdot I_{sb}$	$\alpha_{3,sb}$	751*	682*	-764	_
		(167)	(264)	(4484)	_
$\sigma_{precip} \cdot TENANT \cdot I_{oth}$	$\alpha_{3,oth}$	671*	442	1028	_
		(224)	(339)	(5121)	_
$\sigma_{precip} \cdot OFFFARM \cdot I_{cn}$	$\alpha_{4,cn}$	-107*	-103**	125	_
		(33)	(47)	(1257)	_
$\sigma_{precip} \cdot OFFFARM \cdot I_{sb}$	$\alpha_{4,sb}$	-135*	-131*	-276	_
		(45)	(59)	(1203)	_
$\sigma_{precip} \cdot OFFFARM \cdot I_{oth}$	$\alpha_{4,oth}$	-94	-53	-20.11	_
		(112)	(94)	(36.46)	_
$\sigma_{precip} \cdot AGE \cdot I_{cn}$	$\alpha_{5,cn}$	-5.2*	-5.1*	0.54	0.25
		(1.3)	(1.8)	(0.61)	(239488)
$\sigma_{precip} \cdot AGE \cdot I_{sb}$	$\alpha_{5,sb}$	-3.7*	-4.0**	1.16	-0.51
		(1.4)	(2.0)	(1.01)	(37343)
$\sigma_{precip} \cdot AGE \cdot I_{oth}$	$\alpha_{5,oth}$	-8.8***	-2.9	0.40	0.41
		(5.4)	(4.1)	(0.87)	(211979)
$\sigma_{precip} \cdot MALE \cdot I_{cn}$	$\alpha_{6,cn}$	-774*	-763**	43.20	_
		(208)	(302)	(69.91)	_
$\sigma_{precip} \cdot MALE \cdot I_{sb}$	$\alpha_{6,sb}$	-550**	-605***	101.47	_
		(223)	(338)	(89.10)	_
$\sigma_{precip} \cdot MALE \cdot I_{oth}$	$\alpha_{6,oth}$	-350	-301	-82.25	_
		(551)	(469)	(105.35)	_
Fraction of correct		_	0.70	_	_
Log (likelihood)		-1.75	-779.3	-1.27	-1.27

[i] Standard errors are reported in parenthesis; they are computed from analytic second derivatives; *, **, and *** indicate statistical significance at the 1%, 5%, and 10% levels respectively.