Journal of Environmental Management (1996) **48**, 375–386

Application of Sample Selection Model in Estimating Response Functions for Nitrate Percolation

Manzoor E. Chowdhury and Ronald D. Lacewell

Department of Agricultural Economics, Texas A&M University, College Station, Texas, U.S.A.

Received 6 *November* 1995; *accepted* 14 *December* 1995

The Seymour aquifer region of Texas has been identified as containing elevated levels of nitrate in ground water. Various state and federal agencies are currently studying policy options for the region by gathering more site-specific information. However, because of lack of sufficient information, cause and effect relationships between water quality and agricultural practices have not been well established for the region. Some recently available biophysical simulation models have impressive capabilities in generating large amounts of data on environmental pollution resulting from agricultural production practices. In this study, the data generated by a biophysical simulation model were used to estimate the nitrate percolation response functions for the Seymour aquifer region. Interestingly, nitrate percolation values obtained from simulation models often comprise a *censored* sample because the non-zero percolation values are only observed under certain climatic events and input levels. It has been shown in the econometric literature that the use of Ordinary Least Squares (OLS) on censored sample data produces biased and inconsistent parameter estimates. Thus, a sample selection model was used in this study to estimate the response functions for nitrate percolation. The study provides some insight into the relationship between nitrate percolation and agricultural production practices. In particular, the study demonstrates the potential of selected design standards in minimizing agricultural nonpoint-source (NPS) pollution for the study area.

Keywords: aquifer, biophysical simulation, environmental policy, ground water, nitrate, nonpoint-source pollution, sample selection model.

1. Introduction

Ground water contamination from agricultural production practices has become an important environmental issue due to real and suspected threats to human health and the economic costs associated with making contaminated water potable. The issue facing the policy makers today is how to select from alternative policies—those which will protect ground water quality while minimizing adverse effects on farm income. A

0301–4797/96/120375 + 12 $18.00/0

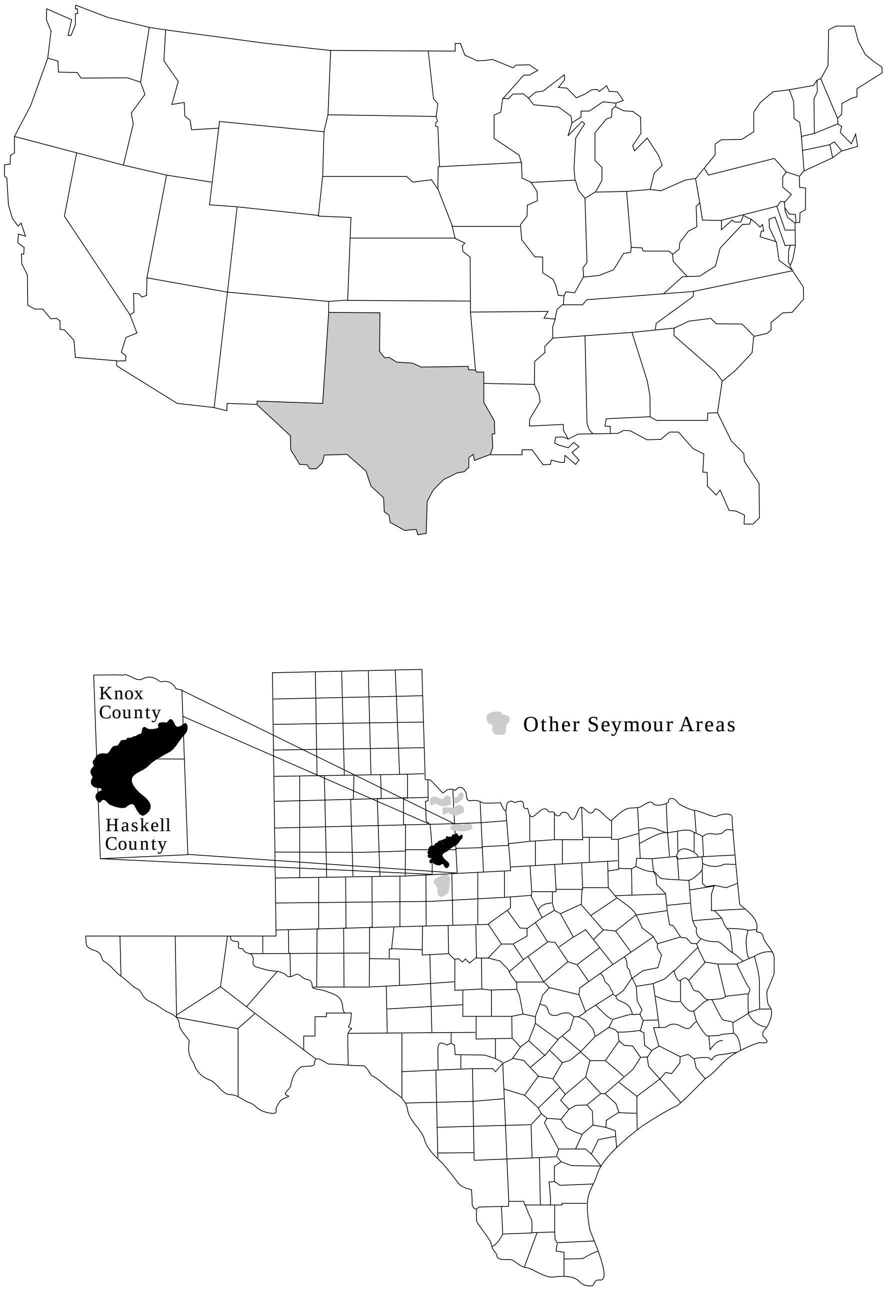

Figure 1. The Seymour Aquifer.

number of policy alternatives have been proposed in the theoretical literature for dealing with nonpoint-source (NPS) pollution. However, the linkages between the theory and application of policy is complicated by the difficulty of identifying sources and measuring individual emissions, their variability over time and space, and the role played by natural processes in determining the ultimate impact of pollutants on the environment.

The policy vacuum in the area of NPS contamination control thus seems largely due to a poor understanding of the cause and effect links between economic activities and environmental damages.

Understanding and quantification of environmental relationships are often hampered by unavailable data. Field testing is prohibitively expensive, and even when some field testing data are available, they are often inadequate for any detailed analysis. These limitations have necessitated the use of sophisticated simulation models to measure environmental impacts of alternative production practices and environmental policies. The simulation models, however, generate a large amount of data. It is advantageous to synthesize these data into clear and concise results by estimating functional relationships. Thus, the predicted value of any dependent variable can be found for any combination of inputs. This advantage is clear when the estimated regression equations are incorporated inside a mathematical programming model. By using only a few regression equations, a large number of production technologies can be generated for analysis.

The purpose of this article is to estimate response functions for nitrate percolation for the Seymour aquifer region of Texas, a Hydrologic Unit Area (HUA) under the President's Water Quality Initiative. The region (Figure 1) covers approximately 274 500 acres in portions of Haskell and Knox Counties in north-west-central Texas, where many water wells exceed the federal safe drinking water standards for nitrate-nitrogen (Kreitler, 1979; Harden and Associates, 1978; Neilsen and Lee, 1987; Aurelius, 1989; Texas Water Commission, 1989). While few cases of death or severe illness in adults are linked directly to nitrate, the most widely recognized human health consequence of nitrate exposure is methemoglobinemia (blue-baby disease) in infants (Bouwer, 1978). Nitrate-nitrogen (NO_3-N) levels greater than 10 ppm (parts per million) make infants more susceptible to this disease. With the predominance of sandy soils in Seymour aquifer area and the shallow depth to ground water, there is significant risk that nitrate and other pollutants may further contaminate the aquifer. Various state and federal agencies have begun to study the best management practices and other policy options to minimize the contribution of agriculture to the pollution of the aquifer. This study is intended to complement these efforts by providing some timely information for policy makers.

Another motivation for this study stems from the opportunity to contribute to applied econometric literature by utilizing an appropriate estimation technique in estimating the response functions for nitrate percolation. Nitrate percolation values obtained from simulation models often comprise a *censored* sample because the non-zero percolation values are only observed under certain climatic events and input levels. A censored sample occurs when some observations for the dependent variable that correspond to known sets of independent variables are zeros. Tobin (1958) had shown that the use of Ordinary Least Squares (OLS) on censored sample data produces biased and inconsistent parameter estimates. Because of the sensitivity to policy implications, careful attention must be given to the choice of an estimation method when estimating response functions using simulated data.

The remainder of the paper provides a critical overview of alternative estimation methods, a description of the Heckman two-step method which has been used in this study, a description of the data, the estimated results and a discussion of their implications.

2. Choice of a sample selection model

2.1 ALTERNATIVE MODELS

As noted earlier, the simulated data on nitrate percolation often comprise a censored sample. One alternative may be to ignore all limiting observations (zeros) and use the OLS estimation procedure. However, ignoring these meaningful limiting values amounts to having a truncated sample, and the OLS parameter estimates in this case are also biased and inconsistent (Maddala, 1983; Amemiya, 1984). Tobin's limited dependent variable model, more commonly known as the Tobit model is customarily used in these situations. The limitation of a Tobit model, however, is that the same set of explanatory variables and parameters determine both the probability of occurrence of pollution and the distribution of pollution (Lin and Schmidt, 1983; Lee and Maddala, 1985; Haines *et al*., 1988). In the context of nitrate percolation, the *occurrence* of percolation and the *amount* of percolation are not so intimately related. In an agricultural setting influenced by stochastic weather, it is not difficult to imagine a scenario where a particular soil or tillage practice might have a lower probability for percolation, but might also have greater average percolation when percolation actually occurred. Thus, the choice of an estimation technique is critical when the relationships involving NPS pollution is measured.

Another limitation ot the Tobit model relates to computational issues. Nitrate percolation is sensitive to weather, soil type, tillage and other relevant production practices. Furthermore, because of the depth of the dewatered (vadose) zone, nitrate requires a certain time to travel through the soil profile and eventually to percolate into the aquifer. Thus, in our study some production practices resulted in high number of limiting values (zeros). As the computation of parameter estimates in a Tobit model is generally performed via some version of Newton's iterative method, the convergence may not occur with a high number of limiting values. Even when convergence occurs, the coefficient estimates of the model may not be accurate or meaningful (Maddala, 1983; Capps, pers. comm.).

Thus, an alternative to the Tobit model is often needed due to the above-mentioned restrictions. A number of consistent alternatives to maximum likelihood estimation have been proposed in the literature. In agricultural economics, these models have been mainly used in consumer demand and recreational demand studies. Cragg (1971) developed several generalizations of the Tobit model that allow the decision process to have two steps. However, because of a truncated distribution used in the second step, Cragg's estimator is computationally more difficult (Haines *et al*., 1988). Fair (1977) suggested an alternative iterative method for obtaining the maximum likelihood estimators of the censored model. Estimates obtained from the Fair's procedure have been shown to be identical to the Tobit model and the criticisms of the Tobit model are also applicable to Fair's model. Based on the method of moments, Greene (1981) proposed an alternative non-iterative method which corrects the bias of least squares estimator of the censored model. The motivation behind such an approach was the prohibitive costs of maximum likelihood estimation which required an iterative procedure. With the current state of computer technology, Greene's procedure does not have an advantage over the Tobit model. An ingenious way of approaching the problem of estimating the Tobit model had been proposed by Heckman (1976, 1979). The Heckman method was used in this study because of its simplicity and appropriate behavioral implications.

2.2. THE HECKMAN MODEL

Heckman devised a relatively simple two-stage estimation process that yields consistent parameter estimates. In his method, the censored sample problem is treated as a specification error or omitted variable problem. According to Heckman, it is possible to correct for the above problem by first estimating the omitted variable λ_i. Using Probit analysis, λ_i is consistently estimated as the inverse of Mill's ratio, $f(x_i'b/s)/F(x_i'b/s)$, where $f(\cdot)$ and $F(\cdot)$ are the density and distribution function of the standard normal distribution, respectively. Though not directly observable, λ_i can be consistently formed by a likelihood function of the binary variable,

$$Z_i \quad = 1 \quad \text{if } Y^*_i > 0,$$
$$\quad = 0 \quad \text{if } Y^*_i \leq 0.$$

where Y^*_i is nitrate percolation or dependent variable. The first stage of Heckman's two-stage procedure is to obtain the consistent estimates of the parameters, $x_i'b/s$ and λ_i, by maximizing the log-likelihood function,

$$L = \sum_{i=1}^{n} [1 - z_i) \log F(-x_i'b/s) + z_i \log F(x_i'b/s)]$$

The second stage involves using the estimated λ_i as an additional regressor, and applying least squares only on those observations where $Y_i > 0$. The parameters obtained from the second stage are consistent and asymptotically normal.

The heckman model is attractive for several reasons. Unlike the Tobit model, the Heckman model is better suited for situations where the probability of occurrence of nitrate percolation and the amount of percolation is not intimately related. It is also free of computational difficulties when the number of zeros are large—a problem encountered with the likelihood function of the standard Tobit model. Another attractive feature of the Heckman two-step procedure is that it allows the researchers to statistically test for sample selection bias. If the estimated coefficient associated with the Mill's ratio is significantly different from zero, then there is no sample selection bias that arises from using non-randomly selected samples to estimate behavioral relationships. Moreover, since the final coefficients are estimated using OLS, all the usual goodness of fit criteria can be applied.

Heckman's model is also more appropriate when the limiting values of the dependent variable (zeros) are unknown and inexplainable while the Tobit or Cragg's model applies when the values of dependent variable are known to equal zero (Lin and Schmidt, 1983). In the context of agricultural NPS pollution, zero values of nitrate percolation (for a particular scenario) generated by a simulation model are not entirely explainable because of the complexity of biophysical process. Thus, Heckman's model is more appropriate for such a sample where the limiting values do not have an obvious interpretation or meaning.

3. The data

The data required for estimating the response functions for nitrate percolation were generated from EPIC-WQ (Erosion Productivity Impact Calculator-Water Quality), a biophysical simulation model. EPIC-WQ is a sophisticated process model which is composed of physically-based components for processes of soil erosion, plant growth, weather, hydrology, nutrient cycling, tillage, soil temperature and economics (Sharpley and Williams, 1990). The model has been employed successfully for numerous sites in the U.S. as well as in other countries (Jones and O'Toole, 1987; Cabelguenne *et al*., 1990).

EPIC-WQ's soil database maintains a 2-m soil profile data for different regions of the U.S. This precludes simulation of nitrate percolation into the aquifer through the dewatered (vadose) zone. To modify the EPIC-WQ model by incorporating the dewatered zone, 36 well logs were selected from a Texas Department of Water Resource's study on Seymour aquifer (Harden and Associates, 1978). The average depth of these wells is about 26–27 feet, the same as the average depth to water across the Seymour aquifer area. With the help of EPIC-WQ model developers (Benson, pers. comm.; Williams, pers. comm.), the dewatered zone was divided into five layers and these layers were added below the top soil in the EPIC-WQ model. This enables the EPIC-WQ model to simulate nitrate percolation through the dewatered zone into the aquifer.

Results from the U.S. Geological Survey (USGS) well testing (two wells drilled in 1992) in the Gilliland/Truscott segment of the Seymour aquifer were used to validate EPIC-WQ for nitrate percolation to the aquifer. The Gilliland/Truscott segment is small, isolated from the main segment of the Seymour aquifer, and has no intensive agricultural production activities. By using the well logs and by simulating native pasture production for 50 years, nitrate leaching results were obtained from EPIC-WQ and compared with well testing results. EPIC-WQ simulation of native pasture scenario for two wells were 6·2 ppm and 5·7 ppm compared to actual well tests of 9·3 ppm and 8·4 ppm, respectively. This suggests that the EPIC-WQ generated value may be slightly lower than actual, but the relationship of one value to another is in the right direction. Furthermore, since the USGS results were obtained by single well tests only (instead of repeated testing of the same site), the discrepancy between the well tests and simulated values does not necessarily reflect any weakness in EPIC-WQ.

After validating EPIC-WQ, the data on the concentration of nitrate percolation were generated by EPIC-WQ simulations for various input combinations and management practices for cotton and wheat production. The EPIC-WQ simulations were performed for selected levels of pre-plant and post-plant nitrogen fertilizer application, irrigation water applications across three periods (early to mid-July, mid-July to early August, and early August to late August), two soil types (Miles and Abilene), and two tillage practices (conventional and minimum). Seven levels of pre-plant and post-plant fertilizer (20 to 80 pounds at 10-pound intervals) and three levels of irrigation application in each period (3, 5 and 7 inches) were used for irrigated cotton. Combinations of seven levels of pre-plant and post-plant fertilizer (20 to 80 pounds at 10-pound intervals) were used for dryland wheat. For dryland cotton, only five levels of fertilizer (20 to 60 pounds at 10-pound intervals) were used. These same input levels were used to generate the data under alternative soils and tillage practices which were used as binary variables. The total number of simulated data points or observations were 4900, 700, and 700 for irrigated cotton, dryland cotton, and dryland wheat, respectively. The simulations were conducted for a 25-year period. Corresponding rainfall data for 25 years were generated by applying the stochastic weather generator in EPIC-WQ.

4. Estimation

4.1 MODEL SPECIFICATION AND CHOICE OF FUNCTIONAL FORM

Three nitrate percolation response functions were estimated for irrigated cotton, dryland cotton, and dryland wheat by using the econometric computer package SHAZAM (White, 1993). The following nitrate percolation response function was considered for irrigated cotton. The specification for dryland cotton and wheat was identical except irrigation was not included.

$$Y = f(N_1, N_2, W_1, W_2, W_3, R, D_1, D_2)$$

where Y is the concentration of nitrate percolated into the aquifer, N_1 is the amount of pre-plant nitrogen fertilizer used, N_2 is the amount of post-plant nitrogen used, W_1, W_2, and W_3 are the irrigation water used during period 1 (early to mid-July), period 2 (mid-July to early August), and period 3 (early August to late August), respectively, R is rainfall during the growing season (June to September for cotton, and February to May for Wheat), D_1 is a binary variable taking on a value of one for the alternative soil (Abilene soil) and zero otherwise (Miles soil) and D_2 is another binary variable taking on a value of one for minimum tillage and zero otherwise (conventional tillage).

The main criterion used for choosing the appropriate functional form for nitrate percolation function was the need to include nitrate leaching even when no inputs were used. This choice of criterion is particularly important for the Seymour aquifer area where a portion of nitrate contamination had been argued to have originated from natural soil nitrates (Harris, pers. comm.). EPIC-WQ simulation of native pasture for 50 years showed a positive (but low) percolation of nitrates. Allowing an intercept for the leaching function resulted in the initial selection of six functional forms: linear, semi-log, cubic, quadratic, square-root, and three-halves. Other functional forms were not considered as they did not have the characteristic imposed by the choice criterion. The linear form showed a poor fit supporting the hypothesis that nitrate percolation is a complicated and highly non-linear process. The final selection was made from semi-log, cubic, quadratic, square-root and three-halves. Using the highest R^2 and the lowest Schwarz criterion, the semi-log (log of dependent variable) form resulted in the best statistical fit for all crops. A non-nested testing procedure was also used to identify the preferred functional form among linear, semi-log, cubic, quadratic, square-root, and three-halves. The procedure is the likelihood dominance criterion of Pollak and Wales (1991) where the model with the highest likelihood value is preferred. The likelihood dominance criterion also selected the semi-log form over other forms. The semi-log form implies that concentration of nitrate percolating to the aquifer is relatively low for small fertilizer application but gradually increases at an increasing rate with higher fertilizer application.

4.2 RESULTS AND IMPLICATIONS

The coefficient estimates of nitrate percolation functions for the semi-log form are reported in Table 1. The t-values of the estimated coefficients show that they are all significant at the 10% level. As the function is in log form, the coefficients must be translated by taking the exponential. Because of the presence of interaction terms, a Wald test (Kmenta, 1986, p. 492) was conducted to test the significance of variables

TABLE 1. Parameter estimates of nitrate leaching functions

Parameter	Estimated value*		
	Irrigated cotton	Dryland cotton	Dryland wheat
Intercept	1·9013	1·40	1·756
	(6·74)	(2·26)	(6·76)
N_1	0·0273	0·0414	0·0146
	(8·12)	(2·89)	(2·04)
N_2	0·0183	0·0337	0·0135
	(7·68)	(2·07)	(2·45)
N_1N_2	0·000145	0·00025	−0·00019
	(4·18)	(2·98)	(−3·02)
R	−0·222	−0·0219	−0·0240
	(−24·31)	(−4·14)	(−11·31)
RN_1	−0·00041	−0·00013	
	(−1·693)	(−2·41)	
RN_2	−0·00052	−0·00015	
	(−3·46)	(−2·37)	
W_1	0·0439		
	(2·42)		
W_2	0·0278		
	(1·65)		
W_3	0·0498		
	(2·66)		
W_1W_2	0·00556		
	(2·67)		
W_1W_3	−0·00324		
	(−1·95)		
W_2W_3	0·00428		
	(2·04)		
D_1	0·10363	0·0970	0·113
	(3·11)	(2·62)	(3·09)
D_2	−0·0016	−0·0014	−0·041
	(−0·321)	(−1·97)	(−0·252)
D_1N_1	0·00221	0·0021	0·00263
	(3·92)	(1·71)	(1·89)
D_1N_2	0·00243	0·0017	0·00221
	(3·47)	(1·92)	(1·76)
D_2N_1	−0·00231	−0·0019	−0·00649
	(−1·35)	(−0·378)	(−0·183)
D_2N_2	−0·00319	−0·00193	−0·00632
	(−0·479)	(−0·523)	(−0·386)
R^2	0·5928	0·4902	0·5540
Schwarz criterion	−0·4176	−0·428	−0·429
Mill's ratio	3·8945	4·2054	1·8848
	(1·279)	(1·022)	(0·8945)
Selected functional form	Semi-Log	Semi-Log	Semi-Log

*t-statistics are given in parentheses.

which showed that all variables are significant at the 10% level. The coefficients of N_1, N_2, and N_1N_2 are all highly significant implying that nitrogen fertilizer is one of the primary contributors of percolated nitrate. The negative signs of coefficients for rainfall variable (R) and interaction terms for rainfall and nitrogen (RN_1 and RN_2) show the effect of rainfall on nitrate concentration, i.e. rainfall reduces the concentration of nitrate. Interestingly, the concentration of nitrate associated with irrigation has a positive relationship because of the presence of some nitrate in irrigation water. However, the magnitude of the coefficients (W_1, W_2, and W_3) is small. The intercept and slope shifters of the Abilene soil (D_1, D_1N_1, and D_1N_2) are positive and significant, implying that the concentration of nitrate percolation would be higher in Abilene soil. Also of interest is the relationship between tillage practice and nitrate percolation and the results in this case are mixed. Coefficients for minimum tillage under irrigated and dryland cotton (D_2 and D_2N_1) indicate that concentration of nitrate would decrease under minimum tillage although the magnitude is rather small. Other coefficients with respect to minimum tillage are shown to be insignificant.

Figure 2 illustrates the estimated relationships for three selected scenarios to provide some qualitative insight. It must be noted that the graphs show the relationship between nitrate percolation and nitrogen application only. Because of this two-dimensional nature, variables such as rainfall or irrigation had to be fixed at their average level. The top panel of Figure 2 shows the relationship between nitrate percolation and pre-plant nitrogen application under Miles and the Abilene soil for dryland cotton. The plot shows that the Abilene soil has a higher percolation than Miles soil and the difference increases with higher levels of nitrogen application. With 40 lb of pre-plant nitrogen application, the Abilene soil shows a percolation of 21·08 ppm compared with a 17·59 ppm for the Miles soil. However, with 60 lb of pre-plant nitrogen application, the difference in percolation is shown to be 10 ppm (49·44 ppm in the Abilene soil vs. 39·01 ppm in Miles soil). The major policy implication is that it might be possible to attain a percolation limit by allocating crops to different soils. The reduction in farmers' net income, however, may be significant because of yield difference associated with different soils.

Figure 2 (middle panel) shows the relationship between nitrate percolation and pre-plant nitrogen application under conventional and minimum tillage for irrigated cotton. With a 60 lb pre-plant nitrogen application, nitrate percolation was 13·15 ppm and 11·45 ppm, respectively, for conventional and minimum tillage. Although the reduction in nitrate percolation was not substantial for minimum tillage, with little or no difference in crop yield, adoption of minimum tillage can help reduce the percolation of nitrates.

Figure 2 (bottom panel) shows nitrate percolation for irrigated cotton for pre-plant and split application of fertilizer under a 4-in irrigation level in each of the three periods. The split application is applied in 50–50 proportion (50% of the fertilizer in pre-plant application and 50% in post-plant application). In this scenario, nitrate percolation exceeded the U.S. Environmental Protection Agency (EPA) standard of 10 ppm at 48 lb of pre-plant nitrogen application. A split application (24 lb pre-plant and 24 lb post-plant) reduces percolation to 7·18 ppm. This suggests that the split application of nitrogen fertilizer is a viable management practice for controlling nitrate percolation in the Seymour aquifer.

A few other findings, while not illustrated with graphs, deserve mention. The response function for wheat shows that the difference in nitrate percolation between all pre-plant and split application becomes more prominent at higher levels of fertilizer use. In the case of dryland cotton, the estimated response function suggests that the

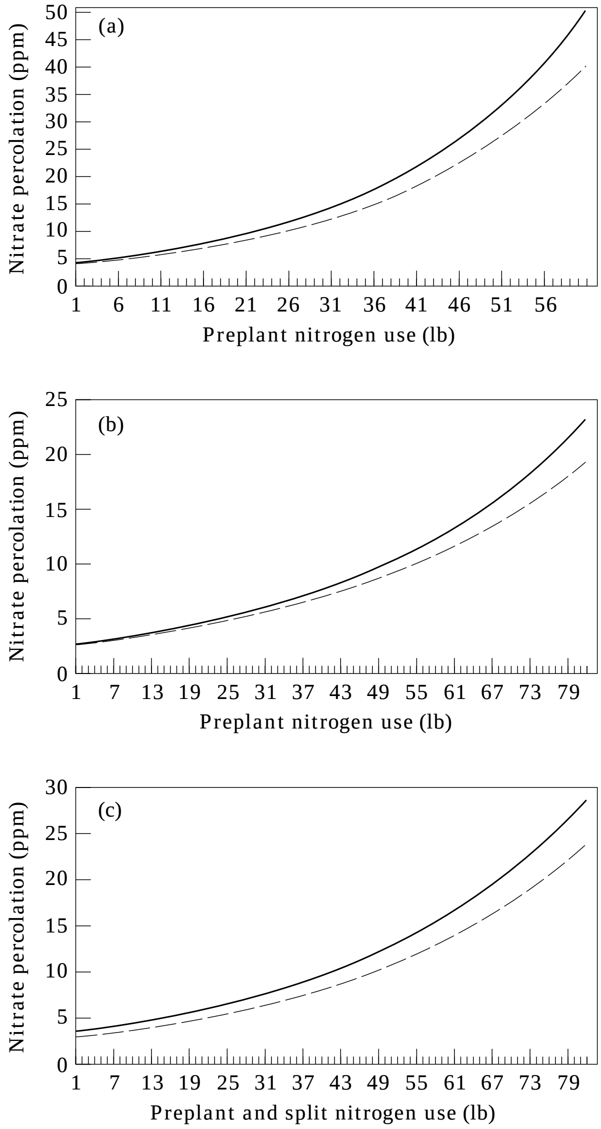

Figure 2. (a) Nitrate percolation in Miles (---) and Abilene Soil (—) for dryland cotton; (b) Nitrate percolation in convention (—) and minimum tillage (---) for irrigated cotton; (c) Nitrate percolation with pre-plant (—) and split nitrogen application (---) for irrigated cotton.

concentration of nitrate percolation is higher than irrigated cotton. This finding, contrary to the popular belief, is open to several interpretations. One possible explanation for higher nitrate concentration under dryland cotton production is that in the absence of soil moisture, the crop uptake of applied nitrogen is not as high as irrigated cotton. Thus, with occasional high rainfall, particularly during non-growing season, the applied nitrogen percolates into the aquifer with higher concentration. There may exist other agronomic explanations, the investigation of which is beyond the scope of this paper.

The most noteworthy policy implication that can be drawn from the above analysis is that design standards may prove to be successful in controlling nitrate percolation in the Seymour aquifer region of Texas. In contrast to ambient-based policies or input taxes and subsidies which are difficult to enforce as well as costly, design standards are easy to enforce and can be successfully used to minimize NPS pollution. Design standards such as use of specific tillage practice or soil are observable and this can be easily monitored.

5. Concluding comments

This study has estimated nitrate percolation response functions for the Seymour aquifer area of Texas by using a sample selection model and also has demonstrated the need for using the appropriate estimation technique because of the sensitivity of estimated relationships for policy analysis. While the study provides some insight into the relationship between nitrate percolation and selected agricultural management practices, it does not attempt to investigate economic trade-offs and alternative policies by conducting a farm-level or regional-level analysis. Perhaps the most promising direction for future research on this issue would be developing economic models for policy analysis where the response functions estimated in this study could be incorporated. Such research would seem particularly relevant because of the recently announced proposals for the re-authorization of the Clean Water Act which emphasize voluntary watershed-level planning by developing water quality guidelines and best management practices tailored to local conditions.

This research was supported in part by the United States Geological Survey, through the Texas Water Resources Institute, and the United States Department of Agriculture–Natural Resource Conservation Service, Strategic Planning and Policy Analysis Division. The views expressed are those of the authors and not of their institution or of the funding institutions.

References

Amemiya, T. (1984). Tobit models: a survey. *Journal of Econometrics* **24**, 3–61.

Aurelius, L. A. (1989). *Testing for Pesticide Residues in Texas Well Water.* Austin, TX: Texas Department of Agriculture.

Bouwer, H. (1978). *Groundwater Hydrology.* New York: McGraw-Hill Book Co.

Cabelguenne, M., Jones, C. A., Marty, J. R., Dyke, P. T. and Williams, J. R. (1990). Calibration and validation of EPIC for crop rotations in southern France. *Agricultural Systems* **33**, 153–171.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* **39**, 829–844.

Fair, R. C. (1977). A note on the computation of the tobit estimator. *Econometrica* **45**, 1732–1727.

Greene, W. H. (1981). On the asymptotic bias of the ordinary least squares estimator of the tobit model. *Econometrica* **49**, 505–513.

Haines, P. S., Guilkey, D. K. and Popkin, B. M. (1988). Modeling food consumption decisions as a two-step process. *American Journal of Agricultural Economics* **70**, 543–552.

Harden, R. W. *et al*. (1978). *The Seymour aquifer: groundwater quality and availability in Haskell and Knox Counties, Texas. Report 226,* volumes 1 and 2. Texas Department of Water Resources.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* **5**, 475–492.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica* **47**, 153–161.

Jones, C. A. and O'Toole, J. C. (1987). Application of crop production models in agroecological characterization: simulation models for specific crops. In *Agricultural Environments: Characterization, Classification, and Mapping*. (A. H. bunting, ed.), Wallingford, U.K.: C. A. B. International.

Kmenta, J. (1986). *Elements of Econometrics,* 2nd edn. New York: Macmillan Publishing Company.

Kreitler, C. W. (1979). Nitrogen-isotope ratio studies of soils and groundwater nitrate from alluvial fan aquifers in Texas. *Journal of Hydrology*, **42**, 147–170.

Lee, L.-F. and Maddala, G. S. (1985). The common structure of tests for selectivity bias, serial correlation, heteroskedasticity and non-normality in the tobit model. *International Economic Review* **26**, No. 1, 1–20.

Lin, T.-F. and Schmidt, P. (1983). A test of the tobit specification against an alternative suggested by Cragg. *Review of Economics and Statistics* **65**, 174–177.

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

Nielsen, E. G. and Lee, L. K. (1987). *The Magnitude and Costs of Groundwater Contamination from Agricultural Chemicals*. Washington, D.C.: Economic Research Service, U.S. Department of Agriculture.

Pollak, R. A. and Wales, T. J. (1991). The likelihood dominance criterion: a new approach to model selection. *Journal of Econometrics* **47**, 227–242.

Sharpley, A. N. and Williams, J. R. (1990). *EPIC–Erosion Productivity Impact Calculator: 1. Model documentation*. Temple, Texas: U.S. Department of Agriculture.

Texas Water Commission (1989). *Groundwater quality of Texas—an overview of natural and man-affected conditions*. *Report 89–01*, Austin, Texas: Texas Water Commission.

Tobin, J. (1958) Estimation of relationships for limited dependent variables. *Econometrica* **26**, 24–36.

White, K. J. (1993). *SHAZAM Econometrics Computer Program: User's Reference Manual, Version 7.0*. New York: McGraw-Hill Book Company.