Papers of the 1991 Annual Meeting

Western Agricultural Economics Association

Portland, Oregon
July 7 – 10, 1991

Chung L. Huang, University of Georgia
Sukant K. Misra, University of Georgia

Modeling Consumer Preference for Organic Produce with Selectivity Bias

The study postulated a sequential probit model to estimate consumer preference for organic produce. The first equation estimates the probability of respondent's intention to purchase organic produce. The second equation corrects for sample selectivity bias and estimates the probability that a consumer would buy organic produce with apparent sensory defects.

Food habits and dietary patterns of American consumers are changing. More convenience foods are being purchased and more meals are being eaten away from home, particularly among two-income households. Today's consumers are also more informed about health and nutrition, and are asking more questions and expressing concerns about food quality and safety. Despite the obvious merits of the United Sates agriculture, and food system; consumer concerns about food safety and environmental quality, continue to increase. Evidence of environmental degradation and health risks associated with pesticides use have made food safety a priority issue on the public policy agenda. Following the Alar scare and the Chilean grapes incident in 1989, public concerns about the potentially adverse effects of pesticides or their residues on human health have arisen to an unprecedented level.

Penner, Kramer, and Frantz found that Kansas consumers ranked pesticide residues as the third most important food safety concerns following environmental contaminants and disease-causing organisms. Similarly, in a 1986 survey of Pennsylvanian households, 71.1% of respondents showed a great deal or some concern with eating fruits and vegetables sprayed or dusted with pesticides (Sachs, Blair, and Richter). This compares to 41.5% of Pennsylvanian respondents expressing concern in 1965 (Bealer and Willits). In a study of food shoppers among four U.S. cities, Zellner and Degner reported that about 83% of the respondents expressed a high or medium level of concern about pesticides and chemicals. Most recently, Zind reported that 86% of survey respondents expressed concern about chemical residues on fresh produce.

Responding to consumers' pesticide fears, some retailers have initiated private residue-monitoring programs and are advertising their fresh produce as being specifically tested for chemical residues. Others are promoting the sale of organic foods. The growing interest in nutrition and food safety has contributed to increasing demand for organic foods, particularly in states like California and Washington. Franco estimated that sales of organically grown produce in California soared 41% from 1986 to 1987. The farm value for organic produce in California amounted to $50 million in 1987 as compared with $15.5 billion for all produce grown in that state. Based on a study of focus group, Hammitt suggested that consumers who purchase organically grown produce are a small self-selected group, which is younger, wealthier and better educated than average. In a recent survey of California consumers, Jolly et al. found that approximately one-forth of the respondents indicated they would look for organic foods when they shop, and about 30% of the respondents plan to buy organic foods in the next month. Approximately 57% of the respondents considered organic foods to be better than conventional foods. Concerns for safety, freshness, general health benefits, nutritional value, and effect on the environment were the most important factors that motivated consumer's demand for organic foods (Jolly et al.).

Research related to consumers' preference and demand for organic produce is sparse. Most studies are primarily descriptive in nature and few has focused on analyzing the factors that explain consumer's demand for organic produce. The first objective of this study is to develop a theoretical framework suitable for analyzing consumer choice behaviors. The second objective is to estimate a system of probabilistic choice models of demand for organic produce based on data collected from a survey of Georgia consumers. This study seeks to assess consumer preference and purchase intention for organic produce, and to identify important socioeconomic characteristics that influence consumer's choice for organic produce.

Theoretical Framework

Probabilistic choice models (PCM) are increasingly popular in applied econometric studies because many important economic decisions involve choice among discrete alternatives. Examples are decisions on labor force participation, travel mode, and brands of commodity purchases. The common thread of these examples is the postulate of a causal link between explanatory variables and a response variable and the objective of predicting the impact on responses of changes in explanatory variables. The application of PCM is particularly attractive for marketing research. Market planners and product developers frequently need to assess the market potential of a product that is not yet available in the marketplace.

In general, a PCM has structural microeconomic interpretations as demand functions can be derived from stochastic utility maximization (McFadden). To begin, consider a sample of T consumers, each facing a finite discrete set of I alternatives. Each alternative i provides utility, U_{it}, to consumer t. Utility U_{it} consists of a deterministic component, D_{it}, and an additive disturbance, ε_{it}. Thus,

$$U_{it} = D_{it} + \varepsilon_{it}, \qquad i = 1, \ldots, I, \text{ and } t = 1, \ldots, T.$$

The deterministic component is usually defined in terms of a linear combination of K characteristics, which characterize the choice set and decision maker. Thus, the utility function is formally defined as:

(1) $$U_{it} = u(w_{i.}^{m}, r_{.t}^{n}) + \varepsilon_{it}, \; m = 1, \ldots, K\text{-}M+1, \text{ and } n = M+1, \ldots, K,$$

where $w_{i.}^{m}$ is a vector of characteristics, which measure the real or intrinsic M attributes of alternative i; and $r_{.t}^{n}$ is a vector of measured N socio-demographic characteristics of individual t.

In the decision process, a consumer is said to compare the maximum utility attainable among the set of I alternatives subject to some budget constraints and selects one alternative for which U_i is a maximum. More specifically, equation (1) suggests that an individual t will prefer alternative i and choose that alternative over alternative j if and only if

(2) $$U_{it} \geq U_{jt}, \text{ for all } j \neq i.$$

Equations (1) and (2) imply that the probability individual t choosing alternative i from the set of all I alternatives is:

(3) $$\Pr(i \mid I, w_{i.}^{m}, r_{.t}^{n}) = \Pr(U_{it} \geq U_{jt}; \text{ for all } j \neq i).$$

In practice, U_{it} represents a latent variable that is unobservable and only the outcome of the decision process is observed. Thus, consider a standard binary choice model and let Y be the observed variable that represents the set of alternatives (I = 2). Furthermore, assuming Y = 1 is the outcome that is observed when equation (2) is satisfied in the decision process. It follows that a PCM implied by equations (1) and (2) can be restated in a regression relation:

(4) $$Y_t = X_t^{k}\beta^{k} + \varepsilon_t, \qquad k = 1, \ldots, K,$$

where

$$Y_t = 1, \text{ if } \varepsilon_t \geq -X_t^{k}\beta^{k}, \Rightarrow U_{1t} \geq U_{2t}, \text{ and}$$
$$Y_t = 0, \text{ otherwise.}$$

While a PCM such as specified in equation (4) looks very similar to a traditional linear regression model, its underlying assumptions are distinctively different from that of linear regression models. Equation (4) is derived based on the assumption that discrete outcomes are generated by some explanatory variables that cross thresholds. In other words, embedded in equation (4) is a threshold concept that dictates the effects of an individual's choice and behavioral response. An individual responds to some exogenous stimuli with a certain choice when his utility function or 'degree of conviction' exceeds some threshold level. Furthermore, any changes in response are directly related to the estimated probability that a particular decision will be made. Therefore, estimation of equation (4) using the ordinary least squares procedure is inappropriate. Equation (4) is usually estimated by the probit or logit procedure.

The PCM of equation (4) can be extended to model consumer behaviors in cases where a sequence of choices rather than a single choice is selected in the decision-making process. Hamlett et al. postulated a sequential decision model to study the impact of socioeconomic variables on the likelihood of using natural Christmas trees. The decision on tree use is hypothesized to be made in two steps. The first step is to decide whether a Christmas tree is to be displayed. If the decision is to display a tree, then a choice is made between natural and artificial trees. Hamlett et al. estimated two probit models with the second model being estimated from a subset of tree users. In this study, a two-stage procedure similar to switching regression analysis (Maddala, pp. 223-228) is employed for estimation of a sequential probit model. Specifically, for a two-equation sequential model, the first equation is estimated by the probit procedure. The inverse mill's ratio obtained from the probit estimation is then included in the second equation as an additional regressor. This specification is designed to correct for sample selection bias since only a subset of the sample is used in the estimation of the second probit equation. In addition, Maddala (p. 225) also shows that the residuals in the second equation are heteroscedastic. Therefore, the second-stage estimation should be carried out with a weighted regression analysis.

The Data and Model Specification

The data for this research were collected from a mail survey of Georgia residents conducted in the summer of 1989. The survey was designed to assess consumers' risk perceptions toward food quality and safety with respect to use of chemicals on fresh produce, and their preferences and attitudes toward organically grown produce. Participants in the survey are members of the Georgia Consumer Panel. The panel constitutes a random sample of 580 Georgia households stratified by income class. The survey employed two follow-ups, scheduled three and six weeks after the first mailing. Prenotification letter and reminders were used prior to and during the survey period to elicit cooperation from the panel members. The survey resulted in a total of 389 completed questionnaires, a response rate of 67%. The final sample used for the analysis consisted of 381 observations.

In the survey, participants were asked a series of questions concerning their activities and attitudes toward fresh produce in the marketplace. With respect to preference for organic produce, respondents were first asked if they would prefer to buy fresh produce that is grown organically without using chemical pesticides or fertilizers. If the response were positive, the respondent would be queried about buying organic produce that had sensory defects such as insect holes, blemishes, and soft spots. Lockeretz suggested that consumers consider sensory quality as a more important attribute than price in their fresh produce purchase decisions. Given that the quality of organic produce is usually less consistent than conventionally grown produce, it is important to determine to what extent consumers would buy organic produce with low appearance and sensory quality.

Specifically, the first equation of the sequential probit model hypothesized that preference for organic produce is a function of respondents' risk perceptions and attitudes toward use of chemical pesticides on fresh produce, product attributes that affect their fresh produce purchasing decisions, and the socioeconomic characteristics. The second equation concerning acceptance of sensory defects was specified as a function of product attributes and the socioeconomic characteristics of the respondents. Thus,

(5) $$Porg_i = f(PA_i, ATTR_i, SE_i) + \zeta_i,$$

and

(6) $$Adef_i = g(ATTR_i, SE_i, IMR_i) + \xi_i, \text{ if } Porg_i = 1.$$

Where Porg and Adef represent preference for organic produce and acceptance of organic produce with sensory defects, respectively; PA is a vector of binary variables measuring respondents' risk perceptions and attitudes toward the use of chemicals on fresh produce; ATTR is a vector of product attributes that may influence purchase decisions; SE denotes a vector of socioeconomic characteristics associated with the respondents; IMR is the inverse Mill's ratio obtained from the estimation of equation (5); and ζ and ξ are the disturbance terms associated with equations (5) and (6), respectively. The definitions and summary statistics of the variables are presented in Table 1.

Empirical Results

Probit estimates of consumers' preference for organic produce and their willingness to accept sensory defects are presented in Table 2. Several goodness-of-fit measures are also reported. One measure is the log-likelihood ratio. A second measure used is the pseudo-R^2 (Maddala, p. 40). A third measure examines how well the model classifies the respondents correctly based on the estimated probabilities. All these measures indicate that the models had significant explanatory power and they fitted the data reasonably well. The signs on the estimated coefficients were consistent as might be expected. Results from the second-stage estimation, however, showed that the estimated coefficient for sample selection bias was not statistically significant.

The results from the first probit equation indicated that respondents who suggested a ban or greater restrictions on pesticides use are more likely to prefer purchase of organic produce. Similarly, those who considered testing and certification important, are more likely to prefer organic to conventional produce. Among the product attributes, nutritional value and freshness of the produce, and low price had significant impacts on the probability of preferring organic produce. Respondents who considered low price important in making their purchase decisions are less likely to indicate a preference for organic produce. This is to be expected since organic produce is typically priced at least 30% higher than that of conventionally grown produce.

Most socioeconomic characteristics except age and income had no significant effects on predicting preference for organic produce. The results suggested that respondents who were between 30 and 60 years of age are less likely to indicate a preference for organic produce. Respondents who had annual income less than $20,000 are more likely to prefer organic produce. When comparing the predicted outcomes with actual observations, the model predicted 192 out of 231 that indicated a preference to buy organic produce. However, the model was less successful in predicting those who would not prefer to buy organic produce. The number of correct predictions was only slightly greater than those of wrong predictions.

With respect to sensory defects, the attribute of produce appearance was a significant factor that influences the probability of accepting organic produce with low sensory quality. As might be expected, respondents who considered produce appearance important in their purchasing decisions are less willing to buy organic produce if it had sensory defects. Respondents who were of European origin are more likely to tolerate with sensory defects. In addition, educational level and household size were found to have positive and significant impacts on the acceptance of sensory defects. Results suggested that the probability of acceptance increases as respondents' educational level and household size increase. Based on the estimated results, the model predicted 163 negative responses out of 173 respondents who indicated that they would not purchase organic produce with low sensory quality. The model was heavily biased toward wrong predictions on acceptance. Only 19 positive responses were correctly predicted out of 58 cases.

Conclusions

A sequential probit model was formulated to estimate the probabilities of preference for organic produce and acceptance of organic produce with sensory defects. Given that the sample on acceptance of sensory defects was censored, the sequential probit model that accounts for potential selectivity bias was appropriate for analyzing the survey data collected for this study. In general, both equations had significant explanatory power and they proved to be good predictive tools toward predicting the probabilities of preferring organic produce and of not accepting organic produce with sensory defects.

The results suggested that the profile for potential organic produce buyers is: consumers who have an attitude against the use of chemical pesticides on fresh produce; who demand that produce should be tested and certified as residue-free; who tend to believe that organic produce offers better nutritional value; and who are younger than 30 years or older than 60 years of age, and have annual household income less than $20,000. Although the survey found that approximately 61% of the respondents said they would prefer to buy organic produce, only 25% of them indicated that they would buy organic produce even if it had sensory defects. The estimated model based on the subset of potential organic buyers clearly indicated that the attribute of produce appearance had a significant negative effect on that decision. Consumers who would buy organic produce with low sensory quality are likely to be white, and have higher education and larger family. For the organic producers, this information should be helpful in developing marketing strategies and for delineating market segments based on demographic projections.

Based on the empirical evidence, this study concludes that consumers have found organic produce attractive primarily due to concerns about food safety associated with pesticide residues. The demand is most likely to be limited to a small group of consumers who are determined and committed enough to buy organic produce despite high prices and low appearance quality. The most important factors for marketing potential of organically grown produce are testing and certification for freedom of chemical residues, good sensory qualities, and a reasonable price that is competitive with conventionally grown produce.

References

Bealer, R. C. and F. K. Willits. "Worriers and Non-Worriers Among Consumers About Farmers' Use of Pesticides," *Journal of Consumer Affairs* 2(1968):189-204.

Franco, J. "An Analysis of the California Market for Organically Grown Produce." *American Journal of Alternative Agriculture* 4(1989):22-27.

Hamlett, C. A., R. O. Herrmann, R. H. Warland and F. Zhao. "Christmas Tree Consumption Behavior: Natural vs. Artificial." *Northeastern Journal of Agricultural and Resource Economics* 18(1989):135-139.

Hammitt, J. E. *Estimating Consumer Willingness to Pay to Reduce Food-Borne Risk*, Santa Monica, CA: The Rand Corporation, 1986.

Jolly, D. A., H. G. Schutz, K. V. Diaz-Knauf and J. Johal. "Organic Foods: Consumer Attitudes and Use." *Food Technology* 43(1989):60-66.

Lockeretz, W. "Urban Consumers' Attitudes Towards Locally Grown Produce." *American Journal of Alternative Agriculture* 1(1986):83-88.

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

McFadden, D. "Econometric Models of Probabilistic Choice." In C. F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: The MIT Press, 1981, pp. 198-272.

Penner, K. P., C. S. Kramer and G. L. Frantz. *Consumer Food Safety Perceptions*, Manhattan, KS: Kansas State University Cooperative Extension Service, 1985.

Sachs, C., D. Blair and C. Richter. "Consumer Pesticide Concerns: A 1965 and 1984 Comparison." *Journal of Consumer Affairs* 21(1987):96-106.

Zellner, J. A. and R. L. Degner. "Consumer Willingness to Pay for Food Safety." Paper presented at the Southern Agricultural Economics Association Meetings, Nashville, TN, 1989.

Zind, T. "Fresh Trends 1990: A Profile of Fresh Produce Consumers." *The Packer Focus 1989-90*. Overland Park, KS: Vance Publishing Co., 1990, pp. 37-68.

Table 1. Variable Definitions and Sample Statistics

Variable	Definition	Mean	Standard Deviation
Porg	Prefer to buy organically grown fresh produce = 1; otherwise = 0	0.606	0.489
Adef	Would buy organic produce even if it had sensory defects = 1; otherwise = 0	0.251	0.435
CONPEST	If the respondent ranked 'food grown using pesticides' as one of the top three food concerns = 1; otherwise = 0	0.449	0.498
BANPEST	If the respondent indicated that chemical pesticides should be banned or subjected to greater restrictions = 1; otherwise = 0	0.433	0.496
TEST	If it is important that produce be tested and certified = 1; otherwise = 0	0.871	0.335
NUTR/FR	If nutritional value and freshness of the produce is considered important = 1; otherwise = 0	0.827	0.379
LOOK	If the appearance of fresh produce is very important = 1; otherwise = 0	0.811	0.392
LOWPR	If low produce price is considered an important attribute = 1; otherwise = 0	0.827	0.379
WHITE	If the race of the respondent is of European origin = 1; otherwise = 0	0.774	0.419
FEMALE	If the respondent is female = 1; otherwise = 0.	0.685	0.465
AGE	If the age of the respondent is between 30 and 60 years = 1; otherwise = 0	0.486	0.500
EDUC	Years of formal education	12.357	4.331
HSIZE	Number of persons in the household	2.732	1.567
INCOME	If annual household income is less than $20,000 = 1; otherwise = 0	0.281	0.450

Table 2. Probit Estimates of Porg and Adef Equations

Variable	Porg		Adef	
	Coefficient	t-ratio	Coefficient	t-ratio
Intercept	-0.583	-1.494	-3.319	-4.217**
CONPEST	0.198	1.353		
BANPEST	0.611	4.057**		
TEST	0.877	3.905**		
NUTR/FR	0.478	2.295*	0.390	1.045
LOOK	-0.295	-1.491	-1.024	-3.898**
LOWPR	-0.590	-2.755**	0.503	1.566
WHITE	-0.150	-0.861	1.915	4.205**
FEMALE	0.007	0.049	-0.246	-1.060
AGE	-0.333	-2.293*	0.201	0.909
EDUC	0.007	0.388	0.059	2.093*
HSIZE	0.072	1.477	0.174	2.616**
INCOME	0.348	1.966*	0.370	1.401
IMR			-0.529	-1.081
Summary Statistics:				
Number of observations	381			231
-2 x Log-likelihood ratio	75.399[a]			57.821[b]
Percent correctly classified	70.34			78.79
Pseudo-R^2	0.243			0.328

*Indicates the asymptotic *t*-ratio is at the 0.05 significance level.
**Indicates the asymptotic *t*-ratio is at the 0.01 significance level.
[a,b]The likelihood ratio statistic is distributed as Chi-square with 13 and 10 degrees of freedom, respectively, and is significant at the 0.01 significance level.