Demand for Plants Sold in North Carolina Garden Centers

Banaga D. Abdelmagid, Michael K. Wohlgenant, and Charles D. Safley

Demand for selected nursery plants sold in North Carolina (i.e., begonia, dianthus, geranium, impatiens, marigold, petunia, and vinca) was found to be affected more by prices than by income, demographic, and other variables. By using cross-sectional data, a modified AIDS model, incorporating demographic variables and quadratic income response, was estimated. Inverse Mills's ratios were also included in the model to correct for selectivity bias, resulting from zero purchases. Significant own-price elasticities ranged from −0.71 to −1.65, and income elasticities ranged from −0.78 to 0.41.

The objective of this paper is to determine the economic factors affecting the demand for garden center plants. This is accomplished by estimating demand functions for selected nursery plants purchased in North Carolina using cross-sectional data. The demand estimates can be used by nurserymen to segment their markets and to help them in marketing their plants and services. A modified form of the Almost Ideal Demand System (AIDS) is used to estimate demand for seven annual plants: begonia, dianthus, geranium, impatiens, marigold, petunia, and vinca. The model extends the AIDS model by incorporating demographic variables and a quadratic income term. In estimation, a correction is also made for selectivity bias resulting from the use of household data with zero purchases.

There is a dearth of information on demand for nursery products. Johnson and Jensen analyze the impact of GNP and share of private residential construction on demand for U.S. nursery stock. Using household data, Gineo and Omamo develop Engel relationships to identify significant determinants of household expenditures on plants. While their study provides useful information on important variables to include in the analysis (e.g., household income, number of single-family home construction starts, educational level attained, and age composition of the population), it does not analyze demand at the individual product level and does not take into account the influence of product prices on demand behavior. There are only two studies that evaluate the influence of prices (Stegelin; Rhodus), but the elasticity estimates were obtained by comparing prices and quantities for a limited number of data points, rather than by estimating demand functions for individual plants via econometric methods that control for the influence of other factors on demand.

In what follows, we first describe the particular modeling approach used. Next, we describe the data and present the empirical results. Finally, we present and evaluate price and income elasticities for the seven individual plants (begonia, dianthus, geranium, impatiens, marigold, petunia, and vinca) purchased at garden centers.

Model Specification

The use of household data involves qualitative, or discrete, choices in addition to the traditional continuous choices in empirical demand analysis. Therefore, the two choices should be modeled in a mutually consistent manner. The mixed discrete/continuous choices can be regarded as switching regression models, and the estimation techniques developed by Amemyia, Heckman, and Lee and Trost can be applied to them. King and Dubin and McFadden constructed specific parametric demand models for discrete and continuous choices derived from an underlying theoretical model of utility maximization.

Tobin has shown that traditional least squares regression analysis on a sample characterized by a truncated dependent variable can lead to biased

---

Banaga D. Abdelmagid is a Ph.D. candidate, Michael K. Wohlgenant is a professor, and Charles D. Safley is an associate professor, Department of Agricultural and Resource Economics, North Carolina State University, Raleigh, North Carolina. The authors appreciate the helpful comments of the anonymous reviewers.

and inconsistent parameter estimates. Therefore, with a large number of nonconsuming households in the sample, the use of a limited (truncated) dependent variable model is indicated. As a result, estimators that correct for potential truncation bias are appropriate (Thraen, Hammond, and Buxton; Capps and Love; Cox and Wohlgenant).

The decision to buy a horticultural product can be considered a discrete choice because each individual either buys a plant or does not. The decision also has a continuous component because, given that he or she decides to buy a plant, the decision then becomes how many plants to purchase.

The general structure of discrete/continuous demand models described by Hanemann derives from the utility maximization model with a random component. The random utility model arises when we assume that, although the utility function is deterministic for the individual consumer, the utility function contains some components that are unobservable to the investigator. The unobservable components could be characteristics of the consumer and/or attributes of the commodities. The concept, therefore, combines two ideas that have a long history in economics: variation in tastes among individuals in a population and unobservable variables in the econometric model (Hanemann).

When the utility function is a well-behaved function, the derived demand functions possess all the standard properties. In particular, the function is quasiconvex, decreasing in price, increasing in income, and satisfies Roy's Identity. The quantities are known to the consumer, but because preferences are incompletely observed, they are random variables from the point of view of the econometric investigator. The discrete choice of whether to consume the commodity or not is indicated by a set of binary value indices, $Y_1, \ldots, Y_n$, where $Y_i = 1$ if the quantity consumed is positive and $Y_i = 0$ if nothing is consumed and n is the number of alternative commodities.

Consider the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer and the Transcendental logarithmic function (TL) by Swamy and Bingswanger. The latter model, called the modified AIDS model, is the AIDS model except for inclusion of an additional squared logged income term. This specification of income permits more flexibility in consumer behavior. The inclusion of the quadratic term allows for more variation in the income elasticities (Chung). Such an extension is important when estimation is conducted using cross-section data, where large variations in income occur.

The modified AIDS models is given by

$$(1) \quad w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln(m/P) + \delta_i(\ln(m/P))^2, \; i = 1, \ldots, n$$

where $w_i = p_i q_i/m$ is the share of the ith commodity in total expenditures, p_i is the price and q_i is the quantity purchased, m is the total consumer expenditures and P is the price index of all commodities.

Equation (1) is written in unrestricted form. In a complete demand system, homogeneity of degree zero in prices in income implies $\Sigma\gamma_{ij} = 0$ for all i. Symmetry implies $\gamma_{ij} = \gamma_{ji}$. Also, because the shares for all commodities add up to one, $n - 1$ of the equations are linearly dependent so that one equation can be dropped from estimation.

Because this study focuses only on annual plants purchased by consumers, model (1) is best viewed as an incomplete demand system (La France and Hanemann). That is, our interest is only in, say, the first n_1 commodities as opposed to demand for all n commodities ($n = n_1 + n_2$). There is no loss in generality by focusing on a subset of n_1 commodities instead of all commodities. Moreover, if the n_1 commodities (plants) are assumed to be weakly separable from all other commodities (nonplants), then it is possible to write the demand functions in equation (1) in terms of just plant prices, real income (m/P) and a price index, say Pn_2, of nonplants; i.e.,

$$(2) \quad w_i = \alpha_i + \sum_{j=1}^{n_1} \gamma_{ij} \ln p_j + \gamma_{in_2} \ln P_{n_2} + \beta_i \ln(m/P) + \delta_i(\ln(m/P))^2, \; i = 1, 2, \ldots, n_1.$$

Because nonplant commodities account for a very large porportion of all consumer expenditures, it is reasonable to assume Pn_2 and P are nearly perfectly collinear. Therefore, upon imposing the zero homogeneity restriction, $\Sigma\gamma_{ij} = 0$, we see that model (2) can be expressed as

$$(3) \quad w_i = \alpha_i + \sum_{j=1}^{n_1} \gamma_{ij} \ln(p_j/P) + \beta_i \ln(m/P) + \delta_i(\ln(m/P))^2, \; i = 1, \ldots, n_1.$$

Finally, because we estimate this system of equations with cross-sectional data and include location dummies in the model, inclusion of price index P is redundant because the location dummies account for cost-of-living differences. Therefore, model (3) can be rewritten as

$$(4)\qquad w_i = \alpha_i + \sum_{j=1}^{n_1} \gamma_{ij} \ln p_j + \beta_i \ln m + \delta_i(\ln m)^2, \quad i = 1, \ldots, n_1$$

where the only parameter restriction to hold is $\gamma_{ij} = \gamma_{ji}$ for $i \neq j$.

Following Heien and Wessells, demographic variables can be incorporated into model (4) in the following manner:

$$(5)\quad \alpha_i = \rho_{i0} + \sum_{k=1}^{s} \rho_{ik} d_k, \quad i = 1, \ldots, n_1,$$

where ρ_{i0} and the ρ_{ik}'s are parameters to be estimated, and the d_k's are the demographic variables. These variables include work status (full time, part time, or retired), indicating the amount of time available for the household to do gardening, and number of years lived in the house, where more gardening would be expected to be done in newer than in older homes. Gender is also relevant, as women tend to have more human capital associated with shopping skills than men (Becker). Age of the customer can also affect plant purchases through time available for gardening or through the type of gardening, whether annual gardening or perennial and tree gardening. As indicated above, dummies, representing locations where the consumers reside and shop, were included to capture differences in cost of living across the state.

Given these specifications, equation (4) can be written as

$$(6)\qquad w_{ih} = \rho_{i0} + \sum_{k=1}^{s} \rho_{ik} d_{kh} + \sum_{j=1}^{n_1} \gamma_{ij} \ln p_{jh} + \beta_i \ln(m_h) + \delta_i(\ln(m_h))^2$$

where h denotes household, and where $\gamma_{ij} = \gamma_{ji}$ for $i \neq j$.

The main obstacle to implementing this model empirically is that, with cross-sectional data where there are numerous zero purchases, the dependent variables are censored. That is, we observe only positive purchases and therefore expenditure shares. Hence, application of least squares methods like seemingly unrelated regressions (SUR) will lead to biased and inconsistent estimates of the parameters of the system of equations because of the existence of nonzero means of the error terms. Two options are available to circumvent this problem. The first would be to develop a utility theoretic decision framework to model each of the sequential decisions involved in purchasing the commodity. The second approach would be to view the model as a tobit model and develop an estimator that provides consistent estimates of the parameter of model (6). This is the approach taken here, where we employ the methods developed by Heien and Wessells. The manner in which their censored regression approach is implemented here is described below.

The dependent variables, budget shares, are either zero or some positive amount for each household. Let the decision to buy or not to buy be indicated by a binary variable Y, which equals one if the household buys the good (i.e., if $w_h > 0$), and zero if the household does not buy the item in question. This decision is a function of latent variables, and for each commodity a probit model of the following form is estimated:

$$(7)\qquad \mathbf{Y}_{ih} = f(\mathbf{p}_h, m_h, \mathbf{d}_h)$$

where $\mathbf{p}_h$ is a vector of prices for the hth household and $\mathbf{d}_h$ is a vector of the demographic variables for the hth household. The same variables used in the probit equation above are included in the demand equation.

The probit equation is used to compute the inverse Mills's ratio for each household; this is,

$$(8)\qquad \mathbf{R}_{ih} = \phi(\mathbf{p}_h, \mathbf{d}_h, m_h)/\Phi(\mathbf{p}_h, \mathbf{d}_h, m_h)$$

where ϕ and Φ are the density and cumulative probability functions, respectively. The inverse Mills's ratios for each item are then included in the second stage regressions. From equation (6), these equations have the following form:

$$(9)\qquad w_{ih} = \rho_{i0} + \sum_{k=1}^{s} \rho_{ik} d_{kh} + \sum_{j=1}^{n_1} \gamma_{ij} \ln p_{jh} + \beta_i \ln(m_h) + \delta_i(\ln(m_h))^2 + \lambda_i R_{ih},$$

where $\gamma_{ij} = \gamma_{ji}$ for $i \neq j$ and $i = 1, 2, \ldots, n_1$.[1]

[1] As an anonymous reviewer correctly points out, the Heien and Wessells procedure is strictly correct only if we assume that the error terms in the probit equations are pair-wise uncorrelated, and if we assume that the error terms between each of the probit equations and AIDS models are uncorrelated. The possible consequence of including only a single inverse Mills's ratio in each equation is that the parameter estimates still suffer from selectivity bias if the error assumptions are correct. Moreover, if the error terms are correlated, then the uncorrected standard errors reported in this paper understate the true standard errors.

Data and Estimation Results

A consumer survey was conducted at eighteen garden centers in five market areas throughout North Carolina in 1992. Four separate retail outlets participated each in the Charlotte area, the Triangle area,[2] the Triad area,[3] and Asheville-Hendersonville area. Two garden centers in the Wilmington market area cooperated. A total of 1,807 customers were surveyed, but only 1,519, or 84%, of the surveys were usable.

The survey was divided into two segments. The first segment was administered before the customers entered the garden center to determine what they planned to purchase and approximately how much they thought they would spend. The second part of the survey was conducted as the customers left the store to determine what they actually purchased and the amount they actually spent. In addition, the second part asked what services the customers considered to be important and how well the garden center met their demands.

With respect to frequency of purchases, the data collected reflect individual/household purchases on just one day. A total of 15% of the customers did not purchase anything while at the garden center, while 70.4% purchased annuals or perennials, 12.1% bought shrubs, and 2% purchased trees. Only 6.7% of the customers indicated that they were browsing, 5.4% stated that they could not find what they were looking for, and 1.4% indicated that they were comparison shopping.

The information collected includes socioeconomic and other demographic variables, including age, income, value of residence, number of years in residence, housing tenure, employment status, and marital status. Information on purchases includes plant prices, plant types, and plant sizes. The analysis here focuses on seven annual plants (i.e., begonia, dianthus, geranium, impatiens, marigold, petunia, and vinca) because their data are the most complete of all the data collected.

Although the items purchased by customers are likely purchased infrequently, we assume all individuals would behave the same way on repeat purchases. Replication does occur across sites within the state. There is also substantial price variation across locations, holding demographic and product characteristics constant.

The survey reported income and value of residence variables in intervals (discrete) form rather than continuous form. These variables are observed to fall only in a certain interval on a continuous scale, with actual remaining unobserved. Also, both end intervals are open-ended. Transforming the data from discrete to continuous would save degrees of freedom in estimation.

An ad hoc method of transforming the discrete data to continuous form might involve assigning the midpoint to observations in any given group, with the open-ended group being assigned values on an even more ad hoc basis. Because such methods do not generally result in consistent estimates, a least squares two-step estimator procedure developed by Stewart was used to transform the discrete data into continuous ones by assigning each observation its conditional expectation q_k:

$$q_k = \mu_y + \sigma_y \frac{[f(B_{k-1}) - f(B_k)]}{F(B_k) - F(B_{k-1})}$$

where q_k satisfies $A_{k-1} < q_k < A_k$, A_k is the boundary value for the category, $B_k = (A_k - \mu_y)/\sigma_y$, and f and F are the density function and cumulative distribution of the standard normal, respectively. Consistent estimates of μ_y and σ_y are obtained by fitting a normal distribution to the sample distribution of the partially observed variable. Tables 1 and 2 show the resulting transformations of the data for income and value of residence from discrete to continuous.[4]

Missing prices occur whenever there are zero purchases, but the model requires price data for all plants and all households. When missing prices were encountered, reported prices for the most frequently purchased items were used as a proxies for missing prices at the garden center where the customer shopped. Prices of other plants at the garden

---

[2] The Triangle area includes Raleigh, Durham, and Chapel Hill, North Carolina.

[3] The Triad area includes Greensboro, Winston-Salem, and High Point, North Carolina.

[4] The way in which the continuous data were generated was as follows. The cumulative frequency normal distribution C_k can be expressed as

$$C_k = \int_{-\infty}^{(A-\mu_y)/\sigma_y} \Phi(U)du$$

where $\Phi(U)$ is the cumulative density function of the standard normal. The inverse of the cumulative frequency is

$$F^{-1}(C_k) = (A_k - \mu_y)/\sigma_y = Z_k$$

Where Z_k is the distance from the mean of a normal distribution expressed in units of standard deviation. Therefore the sample cumulative frequency C_k can be expressed as Z_k and

$$Z_k = (1/\sigma_y)A_k - (\mu_y/\sigma_y)$$

where $(1/\sigma_y)$ and $(-\mu_y/\sigma_y)$ act as slope and intercept of the regression of Z_k on A_k. Consistent estimates of μ_y and σ_y obtained from this regression are then used to form the variables. Values for the density function and cumulative distribution of the standard normal are obtained by evaluating B_k from tabulated values of the normal density function and cumulative distribution function of the standard normal distribution. The values of μ_y, σ_y, B_k, $f(B)$, and $F(B)$ are then used to obtain q_k, which is the conditional expectation of the discrete variable as shown in the text.

Table 1. Transformation of Household Income from Discrete to Continuous

Income Range	Discrete	Continuous
<$15,000	1	$9,251
$15,000–$35,000	2	$25,547
$35,001–$80,000	3	$44,584
>$80,000	4	$82,364

center were imputed by selecting the size with most frequent observations at a specific garden center. In some cases, garden centers had no records of prices on related plants. In these cases, prices recorded at other garden centers in the same market area were taken as proxies by assuming consumers face the same prices in a given area.

The explanatory variables of the model include: (1) location of the household, (2) value of the residence, (3) household income, (4) age of the respondent, (5) number of years residing in the current home, (6) housing tenure (owner or renter), (7) employment status, (8) marital status (single male, single female, married couple), (9) plant prices, and (10) plant sizes. (Definitions of variables are displayed in table 3.) Mean shares of the seven annual plants along with their respective number of positive observations are reported in table 4.

The probit model, equation (7), was estimated for each plant as a function of the following variables: prices for begonia, dianthus, geranium, impatiens, marigold, petunia, and vinca; income; value of residence; years lived in the house; age and dummy variables representing type of residence; male only shoppers; female only shoppers; individual working full time; individual working part time; and locations 1, 3, and 4. The parameter estimates for this model are not reported in this paper but can be obtained from the authors upon request. The Likelihood Ratio (LR) for the probit equation of each plant is used to test the null hypothesis that the coefficients of the probit model are jointly equal to zero (H_o: $\beta_1 = \beta_2 = \ldots = 0$). The results, shown in table 5, indicate that all

Table 2. Transformation of Value of Household Residence from Discrete to Continuous

Value of Residence	Discrete	Continuous
<$100,000	1	$76,460
$100,000–$150,000	2	$120,243
$150,001–$200,000	3	$164,946
>$200,000	4	$211,694

Table 3. Definitions of Variables

Variable	Definition
lnp1	log (price of begonia [cents/plant])
lnp2	log (price of dianthus [cents/plant])
lnp3	log (price of geranium [cents/plant])
lnp4	log (price of impatiens [cents/plant])
lnp5	log (price of marigold [cents/plant])
lnp6	log (price of petunia [cents/plant])
lnp7	log (price of vinca [cents/plant])
lncome	log (household income for 1992 [000$])
lncom2	log income squared
lnval	log (market value of household residence [000$])
lnyr	log (number of years household lived in the house)
lnage	log (customer age [in years])
s1	dummy = 1 if plant size in inches; 0 otherwise
s2	dummy = 1 if plant size in flat; 0 otherwise
males	dummy = 1 if male shopper; 0 otherwise
females	dummy = 1 if female shopper; 0 otherwise
type1	dummy = 1 if detached house; 0 otherwise
work	dummy = 1 if one adult in the house working full time; 0 otherwise
work2	dummy = 1 if two adults in the house working full time; 0 otherwise
local1	dummy = 1 if location 1 (Charlotte); 0 otherwise
local3	dummy = 1 if location 3 (Triad); otherwise
local4	dummy = 1 if location 4 (Ashville-Hendersonville); 0 otherwise
invmill	inverse Mills's ratio

NOTE: The intercept (constant) of the model reflects an individual who is retired (or not working or working part time), is married, lives in Wilmington, lives in an attached house (condominium, apartment, townhouse, duplex, or mobile home), and purchases plants in gallons. The observations of the Triangle area (lcoation 2) have been omitted because they are incomplete.

of them are significant at the 1% level of significance. This indicates that the explanatory variables used in the probit model estimation, excluding the intercept, are jointly significant. As indicated previously, the results of the probit model are used to compute the inverse Mills's ratios, which are then used as explanatory variables together with other variables in estimating the modified AIDS model.

Table 4. Budget Shares and Positive Purchases of Selected Annuals

Annual	Budget Share (%)	Number of Positive Purchases
Begonia	.0215	75 observations
Dianthus	.0084	49 observations
Geranium	.0291	129 observations
Impatiens	.0264	113 observations
Marigold	.0165	49 observations
Petunia	.0156	49 observations
Vinca	.0266	53 observations

Table 5. Likelihood Ratio Test for Probit Equations

Annual	Zero	Positive	LR	P-Value
Begonia	784	75	132.304	0.0001
Dianthus	804	49	36.813	0.0122
Geranium	765	129	73.007	0.0001
Impatiens	766	113	87.503	0.0001
Marigold	801	49	60.489	0.0154
Petunia	789	49	86.081	0.0001
Vinca	791	53	55.959	0.0001

The modified AIDS model (9) is a demand system which can be estimated by seemingly unrelated regression (SUR) with symmetry on the cross-price parameters imposed. In estimating the modified AIDS model, note that the plants under consideration have unequal numbers of observations (see table 4). With unequal numbers of observations, generalized least squares (GLS) estimates of the coefficients from all equations can be obtained, but the estimation does not have the same form as the estimator when the numbers of observations are equal. Also the choice of an estimator for the disturbance covariance is problematic (Judge et al.). The system of seven equations estimated here uses the Schmidt and Sickles procedure.

In this case, the seemingly unrelated regression equation for the ith equation is given by

$$y_i = x_i\beta_i + e_i, \quad i, = 1, 2, \ldots 7$$

where y_i and e_i are of dimension $(N \times 1)$, x is $(N \times K)$ and β is $(K \times 1)$, and where N is the number of observations and K is the number of variables in the equation. Alternatively,

$$y = x\beta + e$$

where the dimensions of y, x, β, and e are, respectively, $[(7T + \Sigma n_i) \times 1]$, $[(7T + \Sigma n_i) \times \Sigma k_i]$, $(\Sigma k_i \times 1)$, and $[(7T + \Sigma n_i) \times 1]$; there are T observations in the equation with the smallest number of observations and $T + n_i$ observations in equation i, where n_i is the number of observations in equation i that exceeds the smallest number of observations. The disturbance vector (e_{1t}, e_{2t}) is assumed to be independently and identically distributed with zero mean and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{17} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \sigma_{71} & \sigma_{72} & \cdot & \sigma_{77} \end{bmatrix}.$$

The following steps are involved in estimating the variance covariance matrix: (1) OLS is applied separately to each function in equation (9) obtaining the vector of sample residuals $(e_1, e_2, \ldots e_7)$ where (2) the diagonal and off-diagonal elements of the matrix are found as follows: let $\hat{e}_1'\hat{e}_2$ be

$$e_i = I - x_i\,(x_i'x_i)^{-1}x_i'y_i \quad i = 1, \ldots, 7$$

residuals from the first and the second equations, respectively, and partition $\hat{e}_2$ *as*

$$\hat{e}_2 = \begin{bmatrix} \hat{e}^*{}_2 \\ \hat{e}^0{}_2 \end{bmatrix}$$

Also define

$$s_{12} = (\hat{e}_1'\hat{e}_2)/T$$

$$s_{11} = (e_1'\hat{e}_1)/T$$

$$s^*{}_{22} = (e^*{}_2'\hat{e}^*{}_2)/T$$

$$s^0{}_{22} = (e^0{}_2'\hat{e}^0{}_2)/n_2$$

$$s_{22} = (\hat{e}_2'\hat{e}_2)/(T + n_2)$$

where there are T observations for the first equation and $T + n_2$ for the second equation. The diagonal and off-diagonal estimators for the first and second equations are

$$\hat{\sigma}^{11} = s_{11}$$

$$\hat{\sigma}^{22} = s_{22}$$

$$\hat{\sigma}^{12} = s_{12}\,(s_{22}/s^*{}_{22})^{1/2}$$

The estimator of Σ is consistent and leads to an estimated GLS (EGLS) estimator of β that has the same asymptotic distribution as the GLS estimator.

The symmetry restrictions of the modified AIDS model, $\gamma_{ij} = \gamma_{ji}$, are tested using the statistic:

$$F = \frac{(r - Rb_*)(R(X'\,(\Sigma^{-1}\otimes I)X)^{-1}R')^{-1}(r - Rb_*)/q}{(e'(\Sigma^{-1}\otimes I)e)N - K}$$

where b_* is the GLS estimator, q is the number of restrictions embodied in the null hypothesis (H_o: $R\beta = r$), and $e = y - Xb_*$. Under the null hypothesis, this statistic follows the F-distribution with q and N–K degree of freedom. With an F value of 0.994, we fail to reject the null hypothesis that symmetry holds. These restrictions are, therefore, imposed in estimation.

The restricted SUR estimates of the modified AIDS model, equation (9), are displayed in table 6. Following Heien and Wessells, the inverse

Table 6. Econometric Estimates of the Modified AIDS Model

	Expenditure Share						
Variable	Begonia	Dianthus	Geranium	Impatiens	Marigold	Petunia	Vinca
Constant	1.456	−0.168	0.136	1.482	−0.563	−2.709	3.822
	(1.051)	(0.928)	(1.482)	(1.143)	(0.638)	(2.19)	(0.993)
lnp1	0.004	−0.010	0.011	0.002	0.018	0.029	−0.054
	(0.010)	(0.008)	(0.002)	(0.042)	(0.015)	(0.026)	(0.017)
lnp2	−0.010	0.001	0.007	0.003	−0.003	0.010	−0.006
	(0.028)	(0.001)	(0.003)	(0.030)	(0.016)	(0.018)	(0.009)
lnp3	0.011	0.007	0.008	0.004	−0.008	−0.004	0.024
	(0.012)	(0.003)	(0.003)	(0.010)	(0.006)	(0.007)	(0.011)
lnp4	0.002	0.003	0.004	−0.017	−0.001	0.023	0.0003
	(0.011)	(0.008)	(0.003)	(0.009)	(0.002)	(0.050)	(0.004)
lnp5	0.018	−0.003	−0.008	−0.001	−0.009	0.0003	−0.008
	(0.024)	(0.008)	(0.019)	(0.026)	(0.010)	(0.041)	(0.022)
lnp6	0.029	0.010	−0.004	0.023	0.0003	−0.001	−0.008
	(0.090)	(0.017)	(0.087)	(0.102)	(0.067)	(0.023)	(0.036)
lnp7	−0.054	−0.006	0.024	0.0003	−0.008	−0.008	0.006
	(0.131)	(0.020)	(0.088)	(0.111)	(0.067)	(0.139)	(0.010)
lnval	−0.001	0.003	0.004	−0.010	0.006	0.006	0.015
	(0.013)	(0.004)	(0.008)	(0.010)	(0.009)	(0.015)	(0.015)
lnyr	0.0004	−0.001	0.001	0.0001	−0.003	−0.005	−0.004
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)
lnage	0.012	−0.017	0.012	0.004	0.027	0.028	−0.004
	(0.018)	(0.007)	(0.009)	(0.015)	(0.012)	(0.019)	(0.014)
lncome	−0.262	0.071	−0.030	−0.230	0.118	−0.533	−0.668
	(0.179)	(0.167)	(0.137)	(0.204)	(0.119)	(0.405)	(0.181)
s1	−0.004	−0.014	−0.007	0.030	—	—	—
	(0.047)	(0.009)	(0.014)	(0.021)			
s2	—	−0.009	—	—	—	—	—
		(0.009)					
males	−0.009	−0.005	0.001	−0.005	0.016	−0.0001	0.031
	(0.013)	(0.003)	(0.009)	(0.009)	(0.008)	(0.017)	(0.014)
females	−0.002	−0.007	−0.006	−0.010	0.004	0.014	−0.020
	(0.008)	(0.002)	(0.005)	(0.006)	(0.005)	(0.008)	(0.008)
type1	−0.001	—	0.001	0.002	−0.009	0.005	−0.001
	(0.015)		(0.008)	(0.011)	(0.015)	(0.015)	(0.016)
work	0.002	−0.0001	0.019	0.018	0.010	−0.004	−0.001
	(0.011)	(0.006)	(0.006)	(0.009)	(0.006)	(0.013)	(0.013)
work2	−0.001	0.002	0.013	0.016	0.012	0.011	0.007
	(0.013)	(0.005)	(0.006)	(0.010)	(0.007)	(0.013)	(0.014)
local1	0.043	−0.006	−0.002	0.020	—	—	—
	(0.059)	(0.020)	(0.020)	(0.023)			
local3	0.019	0.001	−0.008	0.007	0.013	−0.015	0.026
	(0.030)	(0.009)	(0.010)	(0.020)	(0.007)	(0.027)	(0.025)
local4	−0.0005	0.004	−0.010	0.016	0.003	0.006	0.017
	(0.028)	(0.010)	(0.027)	(0.033)	(0.019)	(0.088)	(0.021)
invmill	0.011	−0.132	−0.017	−0.158	−0.058	0.019	0.024
	(0.070)	(0.085)	(0.070)	(0.077)	(0.052)	(0.077)	(0.050)
lncom2	0.011	−0.004	0.00002	0.009	−0.006	0.023	0.030
	(0.009)	(0.008)	(0.006)	(0.010)	(0.006)	(0.019)	(0.009)

NOTE: Definition of variables provided in table 3. Equations estimated from model (9) with symmetry imposed on cross-price parameters. Values in parentheses are standard errors.

Mills's ratio for each commodity is included to correct for selectivity bias. Only dianthus and impatiens show significant coefficients for the inverse Mills's ratios, while the rest of the plants show no significance. The negative sign implies the expenditure distribution is truncated from above. This was expected because people who buy annual plants likely have a threshold price, above which they will not purchase the plant.

The value of residence variable is not significant in any of the equations. The size of the plant, whether in inches, in flats, or in gallons, is included to capture differences in household response to plant size and preferences over particular plant sizes. The results indicate that customers have no special preference for any particular size.

The age variables are positive in most cases, indicating that household customers tend to spend more on annual plants as they grow older. This result is consistent with the assertion that retired and older people have more time available for gardening. The number of years lived in the house variable was not found to be significant in any of the equations.

Dummy variables representing single male shoppers only, single female shoppers only, and couples shopping together (the intercept) were included to capture the effect of gender on demand for annuals. With the exception of petunia and marigold, the results indicate that female shoppers only have a lower demand than both males and females shopping together. This diverges from the expectation that female shoppers only would buy more plants. The male shoppers only dummy was not found to be significant except for marigold and for vinca, where the male shoppers have larger demand than males and females shopping together.

The type of house in which the household lives is represented by a dummy variable taking the value one if it is a detached house and zero otherwise. This variable was included to account for the effect of type of house on demand for annuals. Geranium, impatiens, and marigold have positive signs, while the rest have negative signs. This indicates that owners of detached houses prefer geranium, impatiens, and marigold plants.

Dummies representing full-time working adults in the household were not found to be significant, except for geranium, impatiens, and petunia, where full-time workers show higher intercept values relative to part-time workers and retired adults. The location dummies, which were included to capture differences in cost of living between locations and other fixed differences between locations in North Carolina, indicate that location has no significant effect on purchasing annual plants.

Elasticities

To evaluate the economic structure implied by the estimated SUR modified AIDS model, marshallian own-price elasticities and cross-price elasticities, respectively, are computed at the means of the sample data using the following formulae:

$$\eta_{ii} = \frac{\gamma_{ii}}{\overline{w}_i} - 1$$

where η_{ii} is the own-price elasticity, γ_{ii} is the coefficient of equation i on price of commodity i, and $\overline{w}_i$ is the mean of the share of commodity i;

$$\eta_{ij} = \frac{\gamma_{ij}}{\overline{w}_i}$$

where η_{ij} is the cross-price elasticity, and γ_{ij} is the coefficient of equation i on price of commodity j.

The formula for standard errors (s.e.) of the elasticities is as follows:

$$\text{s.e. }(\eta_{ij}) = \frac{\text{s.e. }(\gamma_{ij})}{\overline{w}_i}$$

where s.e. (η_{ij}) is the standard error of price elasticity, and s.e. (γ_{ij}) is the standard error of the price coefficient.

Table 7 shows calculated own-price and cross-price elasticities. As expected, all own-price elasticities are negative, and demand ranges from being elastic for some commodities (i.e., impatiens,

Table 7. Own-Price and Cross-Price Elasticities of Annual Plants

	Price						
Quantity	Begonia	Dianthus	Geranium	Impatiens	Marigold	Petunia	Vinca
Begonia	−0.799*	−1.202	0.361	0.088	0.672	1.849	−2.034
	(0.484)	(0.903)	(0.984)	(1.602)	(0.552)	(1.672)	(0.630)
Dianthus	−0.471	−0.929*	−0.235	0.104	−0.130	0.652	−0.223
	(1.303)	(0.163)	(0.741)	(1.144)	(0.618)	(1.173)	(0.333)
Geranium	0.490	0.814	−0.710*	0.157	−0.309	−0.252	0.913
	(0.563)	(0.389)	(0.117)	(0.378)	(0.240)	(0.476)	(0.417)
Impatiens	0.108	0.327	0.142	−1.652*	−0.025	1.480	0.011
	(0.504)	(0.993)	(0.099)	(0.334)	(0.069)	(3.186)	(0.146)
Marigold	0.831	−0.409	−0.281	−0.025	−1.328*	0.021	−0.305
	(1.117)	(0.964)	(0.641)	(1.001)	(0.379)	(2.627)	(0.841)
Petunia	1.348	1.213	−0.135	0.878	0.012	−1.069	−0.311
	(4.182)	(2.062)	(2.987)	(3.879)	(2.509)	(1.467)	(1.340)
Vinca	−2.517	−0.703	0.832	0.011	−0.305	−0.528	−0.772*
	(6.118)	(2.321)	(3.007)	(4.210)	(2.210)	(8.908)	(0.379)

NOTE: Values in parentheses are standard errors. Asterisk denotes significance at 0.05 level.

marigold, and petunia) to inelastic for other commodities (i.e., begonia, dianthus, geranium, and vinca). The signs of the cross-price elasticities indicate that the plants are either gross substitutes when the sign is positive or gross complements when the sign is negative. Most of the cross-price elasticities are not significant. This would indicate that these annuals are neither substitutes nor complements.

The formula used to compute the income elasticity can be written as:

$$(11) \qquad \eta_m = \frac{\beta_i + 2\delta \log(m_h)}{w_i} + 1$$

where η_m is the income elasticity, β_i is the coefficient on log income variable, δ is the coefficient on log income squared, and log (m_h) is the mean of log income.

The formula for standard error of the income elasticity is:

$$\text{s.e.}(\eta_m) = \frac{1}{\bar{w}_i} [\text{Var}(\beta) + 4(\log(m))^2 * \text{Var}(\delta) + 4\log(m) * \text{Cov}(\beta,\delta)]^{\frac{1}{2}}$$

where s.e. (η_m) is the standard error of the income elasticity, Var (β) is the variance of the coefficient of income variable, Var (δ) is the variance of the coefficient of log income squared, and Cov (β,δ) is the covariance between coefficients β and δ.

The values of the parameters of the modified AIDS model (equation [9]) allow for variation in income elasticities over alternative income levels, implying that the income elasticities can vary widely over an extended income range. Table 8 shows the calculated income elasticities for the seven selected annuals at the sample means. With the exception of dianthus, geranium, impatiens, and petunia, the plants are normal goods. However, most of the estimates are insignificant. Negative income effects could be confounded with household time as suggested by Becker, and therefore could be indicative of spurious correlation.

Table 8. Income Elasticities of Annual Plants

Annual Plant	Income Elasticity	Standard Error
Geranium	−0.056	0.210
Impatiens	−0.054	0.344
Begonia	0.153	0.549
Vinca	0.024	0.457
Dianthus	−0.367	0.462
Petunia	−0.784	0.782
Marigold	0.412	0.231

Summary and Conclusions

The study utilizes cross-sectional data to estimate market demand functions for nursery products purchased in North Carolina. In addition to individual household and product characteristics, the data include prices at various locations within North Carolina. A modified Almost Ideal Demand System (AIDS) model was used to estimate the demand parameters. The modified AIDS model incorporates a quadratic income term, which allows for more variation in income elasticities. In addition to these variables, demographic variables were also included in the model. Following Heien and Wessells, inverse Mills's ratios were also included in the AIDS model to correct for selectivity bias, resulting from households consuming zero amounts of the commodities under consideration.

Overall, the demand equations estimated were found to be consistent with our prior expectations. However, only prices were consistently found to be significant across the set of commodities analyzed. For the most part, the estimated own-price elasticities indicate that sales of each plant are quite sensitive to price. Indeed, three of the commodities (impatiens, marigold, and petunia) were found to be price elastic. This suggests that for these commodities, higher total revenue could be achieved by lowering price, not by raising price.

These demand estimates for nursery products should provide retail nurserymen with information that can help them identify market segments and improve marketing of their plants and services. In addition to knowing how price and income affect sales, estimates of demographic variables should be helpful in targeting potential customers. The results indicate that a major target audience for the plants analyzed in this paper are older or retired males or couples, who have more time available and who enjoy gardening. However, the results especially underscore the importance of sensitivity of plant sales to prices, and the importance of developing pricing strategies to maximize long-run profit.

References

Amemyia, Takeshi. "Regression Analysis When the Dependent Variable Is Truncated Normal." *Econometrics* 41(1973):997–1016.

Becker, B.S. *A Treatise on the Family.* Cambridge: Harvard University Press, 1981.

Capps, Oral, and John M. Love. "Determinants of Household Expenditure on Fresh Vegetables." *Southern Journal of Agricultural Economics* 15(1983):127–32.

Chung, Ching-Fan. "A Cross-Section Demand Analysis of Spanish Provincial Food Consumption." *American Journal of Agricultural Economics* 76(1994):513–21.

Cox, Thomas L., and Michael K. Wohlgenant. "Price and Quality Effects in Cross-Sectional Demand Analysis." *American Journal of Agricultural Economics* 68(1986): 908–19.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70(1980):312–26.

Dubin, Jeferey A., and Daniel L. McFadden. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52(1984):345–62.

Gineo, W.M., and S.W. Omamo. "An Analysis of Household Expenditures on Nursery Products in the United States." *Southern Journal of Agricultural Economics* 22(1990): 199–208.

Hanemann, W. Michael. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52(1984):541–61.

Heckman, J.J. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1979):153–61.

Heien, Dale, and Cathy Rohein Wessells. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *Journal of Business and Economic Statistics* 8(1990):365–71.

Johnson, L.A., and K.H. Jensen. "Economic Factors Affecting Sales of U.S. Nursery Stock." *Journal of Environmental Horticulture* 10(1992):108–10.

Judge, G.G., W.E. Griffiths, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. 2d ed. New York: John Wiley and Sons, 1985.

King, A. Mervyn. "An Econometric Model of Tenure Choice and Demand for Housing as a Joint Decision." *Journal of Public Economics* 14(1980):137–59.

La France, J.T., and W. Michael Hanemann. "The Dual Structure of Incomplete Demand Systems." *American Journal of Agricultural Economics* 71(1989):262–74.

Lee, L.F., and R.P. Trost. "Estimation of Some Limited Dependent Variable Models with Application to Housing Demand." *Journal of Econometrics* 8(1978):357–82.

Pollak, R.A., and T.J. Wales. "Demographic Variables in the Demand Analysis." *Econometrica* 49(1981):1533–58.

Rhodus, W. Timothy. "Estimating Price Elasticity for Fresh Flower Bouquets Sold in Supermarkets." *Hortscience* 24(1989):386–87.

Schmidt, P., and P. Sickles. "Some Further Evidence on the Use of the Chow Test under Heteroskedasticty." *Econometrica* 45(1977):1293–98.

Stegelin, F.E. "Consumer Demand Elasticities of the Selected Annuals and Perennials." *Proceedings of SNA Research Conference* 39(1994):388–90.

Stewart, W. Mark. "On Least Squares Estimation When the Dependent Variable Is Grouped." *Review of Economic Studies* 50(1983):737–53.

Swamy, G., and H.P. Binswanger. "Flexible Consumer Demand Systems and Linear Estimation: Food in India." *American Journal of Agricultural Economics* 65(1983): 675–84.

Thraen, C.S., J.W. Hammond, and B.M. Buxton. "Estimating Components of Demand Elasticities from Cross-Sectional Data." *American Journal of Agricultural Economics* 60(1978):674–77.

Tobin, James. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26(1958):24–36.