Corner Solutions in Duality Models:

A Cross-Section Analysis of Dairy Production Decisions

by

Robert D. Weaver*

Daniel A. Lass

The authors are Assistant Professor and Graduate Assistant, Department of Agricultural Economics and Rural Sociology, The Pennsylvania State University. Senior authorship is not assigned. Presented at the American Agricultural Economics Association meetings in Ithaca, New York, August 1984.

1. Introduction

Design of effective dairy policy requires knowledge of short- and long-run elasticities of output supply and input demand by dairy farmers. However, little is known about such elasticities. Past studies have reported estimates of either directly estimated milk supply functions (Halvorson (1955), Hughes and Stanton (1965), Conneman (1967), Wipf and Houck (1967), Benson (1972), (Chen et al. (1972), Benson et al. (1974), and Stammer (1975)) or have derived choice response elasticities from directly estimated production functions, e.g., Hoch (1976), and Dahlgran (1980). Initial estimates of the direct function excluded input prices and found the effects of farm operator characteristics to be dominating determinants of milk supply. Results indicated that the supply elasticity with respect to milk price was probably zero. Such supply inelasticity was found in normative and positive studies of farm level as well as aggregate data. Wipf and Houck (1967) and Stammer (1975) included input prices and Stammer found milk supply to be responsive to feed prices, as well as calf, cull, and replacement cow prices, though inelastic with respect to its own price. Despite these findings, industry sources have consistently indicated that milk supply is responsive in the short-run to changes in the price milk through culling and changes in input mixes, e.g., see dairy economics proceedings U.S.D.A. (1982). Past econometric studies have presented models which were only intuitively linked to a behavioral hypothesis. As noted in Weaver (1982, 1983), it is often advantageous to make this link explicit. The primary purpose of this paper is to present a complete set of short-run elasticities of choice by dairy farmers which is consistent with the hypotheses of 1) short-run profit maximization, 2) multiple output and multiple input technology which may be non-joint (Weaver (1977, 1982)), and 3) existence of fixed input flows.

Rather than searching for such relationships in aggregate time series data, the present study presents results based on a sample of Pennsylvania dairy farms from the 1974 U.S.D.A. Cost of Production Study. This study was mandated by Congress[1] and was based on a survey of individual dairy farms which collected a comprehensive account of output and variable input prices and quantities produced or employed, and of fixed farm input flows and operator characteristics. A careful review of state level budget programs indicated that COPS presented without question the most comprehensive farm level data base available. The principal use of this data and its subsequent updates has been annual estimates of enterprise specific costs of production. These estimates are derived from the collected data by using engineering relationships and other enterprise specific budgets to calculate typical farm enterprise budgets for different scales of production and for various regions, see e.g., U.S. Congress (1975). A secondary purpose of this study is to demonstrate the value of such data sets for establishing econometric estimates of fundamental characteristics of production decisions such as elasticities with respect to prices or fixed input flows.

---

[1] The Agriculture and Consumer Protection Act of 1973 directed the Secretary of Agriculture to conduct a study of the cost of production of wheat, feed grains, cotton and dairy commodities. The Economic Research Service and Statistical Reporting Service conducted a survey of nearly 1600 dairy farmers in 24 major milk producing areas of the U.S.

The present paper focuses on cross-section data which allows explicit consideration of heterogeneity of production decisions within a geographical area. Although all farmers may face common technological possibilities, not all factors of production are variable in the short-run. This reality constrains each farmer to choices along particular dimensions of the production possibilities frontier. An immediate implication is some farmers may find corner solutions optimal and not use (or produce) particular inputs (or outputs). A third purpose of this paper is to consider the nature of this problem, and its implications for estimation of complete sets of choice functions and aggregation of such functions.

2. Theoretical Model

The theoretical foundation of a model of production decisions which is to be applied to cross-section data must explicitly incorporate a behavioral hypothesis which recognizes the possibility of corner solutions for some inputs or outputs on some farms. For example, in the data set to be analyzed here only 64% of the farms hired labor. For those farms which did not hire labor, no data for the wage rate is available, and more importantly, continuous relations do not exist between observed hired labor and other choices and the market wage rate for hired labor. The multiple output, multiple input profit function of Weaver [1982, 1983] can be adapted to consider this problem.

Assume firms base production decisions on the solutions of the following choice problem:

$$\max \Pi = PY' - RX'$$

s.t. $F(Y,X,\Theta) = 0$

where: P, and Y are 1 x m vectors of expected net output[2] prices and levels;

R, and X are 1 x n vectors of variable input flows;

Θ is a 1 x p vector of fixed input flows;

π is short-run profits, or equivalently Ricardian rents available as returns to Θ; and

F(.) is a production technology satisfying the usual neoclassical properties.

The Kuhn-Tucker conditions for this choice problem provide the basis for deriving different sets of choice functions and associated expected profit functions depending upon the occurrence of corner solutions. To illustrate, in the case of one truncated chioce, the following simplified set of first-order conditions are of interest:

1) $P_i + \mu \; \partial F/\partial Y^i = 0 \qquad i = 1,\ldots m$

2) $-R_h + \mu \; \partial F/\partial X_h = 0 \qquad h = 1,\ldots n-1$

3) $(-R_n + \mu \; \partial F/\partial X_n)X_n = 0$

4) $F(Y,X,\Theta) = 0$

---

[2] A net output is defined as production minus consumption.

Depending on the value taken on by X_n 1)-4) define two sets of choice functions written in implicit form. Recognizing each set is zero homogeneous in prices, we solve each set for explicit choice functions, and by substitution, the normalized expected profit function (NEPF). The derivative property links choices to respective elements of the gradient vector of the NEPF.

Summarizing these statements,

5) $\pi^* = \pi^*(\tilde{P}, \tilde{R}; \Theta, x_n^* > 0) = \Pi^*(P, R; \Theta, x_n^* > 0)/P.$

6) $Y_i^* = \partial\pi^*/\partial\tilde{P}_i = \partial\Pi^*/\partial P_i = Y_i^*(\tilde{P}, \tilde{R}; \Theta, X_n^* > 0) \qquad i = 2,\ldots m$

7) $-X^*_h = \partial\pi^*/\partial\tilde{R}_h = \partial\Pi^*/\partial R_h = X_h^*(\tilde{P}, \tilde{R}; \Theta, X^*_n > 0) \qquad h = 1,\ldots n$

8) $Y_1^* = \pi(.) - \tilde{P}\tilde{Y}^{*\prime} + \tilde{R}X^{*\prime}$

where $\tilde{P}_i = P_i/P_1$, $\tilde{R} = R/P_1$, and $\tilde{Y}$, $\tilde{P}$ are 1 x (m-1). Concavity of F(.) implies convexity of $\pi^*(.)$. A similar set of functions exists as the explicit form of 1) - 4) conditional on $X_n^* = 0$. These functions would relate optimal choices and profits conditional on $X^*_n = 0$ denoted (π^c, Y^c, X^c) to (P, R^c, Θ) where X^c and R^c are (n-1) * 1. The advantage of the profit function over a cost function as a conceptual tool for studying choice response of profit maximizing firms should be clear from 6) - 8). Given any choice problem (such as profit maximization) a subset of the choices can be determined conditionally on levels of the remaining choice variables (conditioning choice variables) and a conditional dual function (such as a cost function conditional on output levels) can be written.

However, only if the production function is homothetic in the conditioning choice variables can comparative-statics of unconditional choices be derived from estimates of the conditional dual and choice functions. By analogy to demand theory results, Hicksian demand function estimates do not allow identification of Marshallian comparative-statics. Although in the non-homothetic case unconditional comparative-statics can be determined with a suitable model specification, Weaver and Lass (1983), use of the multiple product profit function is clearly more direct and convenient.

The comparative-statics of choice conditional on $X^*_n > 0$ are easily derived from the system of choice functions 6) - 8) by further differentiation with respect to prices, Weaver (1982, 1983). For example,

9) $\partial Y^*_i/\partial P_j = (1/P_1)(\partial^2\pi^*/\partial \tilde{P}_i \partial \tilde{P}_j)$

10) $\partial X^*_h/\partial P_j = -(1/P_1)(\partial^2\pi^*/\partial \tilde{R}_h \partial \tilde{P}_j)$

11) $\partial Y^*_i/\partial P_1 = (1/P_1\{\Sigma\Sigma(\partial^2\pi^*/\partial \tilde{P}_i \partial \tilde{P}_j)\ \tilde{P}_i \tilde{P}_j + \Sigma_h\ (\partial^2\pi/\partial \tilde{P}_i \partial \tilde{R}_h)\ \tilde{P}_i \tilde{R}_h) + \Sigma\Sigma_{hj}$

$(\partial^2\pi^*/\partial \tilde{R}_h \partial \tilde{P}_j)\ \tilde{R}_h \tilde{P}_j) + \Sigma\Sigma\ (\partial^2\pi^*/\partial \tilde{R}_h \partial \tilde{R}_k)\tilde{R}_h \tilde{R}_k\}$

12) $\partial Y^*_1/\partial P_j = (1/P_1)\ \Sigma\ \tilde{P}_i\ \partial^2\pi/\partial \tilde{P}_i \partial \tilde{P}_j$

Continuity of the NEPF in prices implies these comparative-statics satisfy the symmetry property, e.g.,

13) $\partial Y^*_i/\partial P_j = \partial Y^*_j/\partial P_i \qquad i,j = 1, \ldots m,\ i \neq j$

The comparative-statics of short-run choices with respect to exogeneous changes in fixed factors can also be derived from the profit function as was demonstrated by Weaver (1978,1983). For the present case, differentiation of 6)-8) with respect to any fixed factor θ_r provides the basis for determining individual choice elasticities as well as the Hicksian biases (changes in product mixes) of such changes. Following Weaver (1978,1983), the allocative effect of a change in θ_r on the relative use of X_k and X_h can be summarized by the rule:

Changes in θ_r are Hicks'

$$14)\quad X_h \left.\begin{matrix}\text{saving}\\ \text{neutral}\\ \text{using}\end{matrix}\right\} \text{relative to } X_k \text{ if } B_{hk} = \partial \frac{X_h^*}{X_k^*} / \partial\Theta_r \gtreqless 0.$$

Using the derivative properties above, B_{hk} can be rewritten:

$$15)\quad B_{hk} = \frac{\partial^2 \pi^*}{\partial R_k \partial \Theta_r} \frac{1}{X_k^*} - \frac{\partial^2 \pi^*}{\partial R_h \partial \Theta_r} \frac{1}{X_h^*} \gtrless 0.$$

In addition to these comparative-statics, estimates of the full NEP function in 5) can be employed to determine the shadow value of additional units of any fixed factor θ_r, $\partial\pi^*/\partial\Theta_r$. In the long-run, if the firms objective were to maximize each period's total profit (where all factors are variable), then the shadow value of each fixed factor would be set equal to that factor's market price. In the absense of observable market prices, the shadow value reports the decision maker's willingness-to-pay for additional units.

A final descriptive statistic of interest in characterizing production technology and short-run choices is short-run returns to size. Following Weaver (1983), this can be written in terms of the profit function:

$$16)\quad RTSZ \equiv \sum_{h=i}^{n} \frac{R_h}{\pi^*} \frac{\partial \pi^*}{\partial R_h} \Big/ \sum_{i=1}^{m} \frac{P_i}{\pi^*} \frac{\partial \pi^*}{\partial P_i}$$

3. Functional Form

In general, a NEP function such as $\Pi^*(.)$ can be characterized by a parameter vector Γ, the elements of which will be related by duality to the elements of the vector (Δ) of parameters that characterize the production technology. If this relationship among elements of Γ and Δ is one-to-one, it is clear that the dimensions of, and the values taken on by, the elements of Γ will depend on which, if any, choices are optimal when truncated at zero. For example, in the case considered above, Γ is functionally related to Δ for $X_n^* > 0$, while Γ^c would be for $X_n^* = 0$. Consider the case where $X_n^* > 0$, and assume $\Gamma = [\alpha,\beta]$ for a quadratic form of $\Pi^*(.)$. Using the derivative properties and appending additive error terms, we can write the following dual system for the case where $X_n^* > 0$:

$$17)\quad \Pi^* = \gamma_o + (\gamma_P + 1/2\ \tilde{P}\ \beta_{PP} + 1/2\ \tilde{R}\ \beta_{RP} + 1/2\Theta\ \ \beta_{\Theta P})\ \tilde{P}'$$

$$+ (\gamma_R + 1/2\ \tilde{P}\ \beta_{PR} + 1/2\ \tilde{R}\ \beta_{RR} + 1/2\ \ \ \beta_{\Theta R})\tilde{R}'$$

$$+ (\gamma_\Theta + 1/2\ \tilde{P}\ \beta_{P\Theta} + 1/2\ \tilde{R}\ \beta_{R\Theta} + 1/2\ \Theta\ \beta_{\Theta\Theta})\Theta' + \varepsilon$$

$$18)\quad Y_i^* = \frac{\partial \Pi^*}{\partial \tilde{P}} = \alpha_{P_i} + \tilde{P}\ \beta_{P_iP} + \tilde{R}\ \beta_{P_iR} + \ \Theta\ \beta_{P_i\Theta} + \varepsilon_{Yi} \qquad i = 1,\ldots m-1$$

$$19)\quad -X_h^* = \frac{\partial \Pi^*}{\partial \tilde{R}_h} = \alpha_{R_h} + \tilde{P}\ \beta_{R_hP} + \tilde{R}\ \beta_{R_hR} + \Theta\ \beta_{R_h\Theta} + \ \varepsilon_{X_h} \qquad h = 1,\ldots n$$

If the hypothesis is maintained that the NEP is quadratic, the parameters contained in α,β can be interpreted as first- and second-order derivatives of $\pi(\cdot)$ evaluated at the origin.

By Young's theorem, continuity of $\Pi^*(.)$ in $(\tilde{P}, \tilde{R}, \Theta)$ implies that its Hessian is symmetric. If the quadratic form is to be interpreted as a functional representation to $\pi^*(.)$, then β must be symmetric. Conditional upon this interpretation, comparative-statics can be written in terms of (α, β), e.g.,

$$20) \quad \eta_{ij} \quad \frac{\partial Y_i^*}{\partial P_j} \frac{P_j}{Y_i^*} = \beta_{ij} \frac{P_j}{P_1 Y_i^*}$$

$$21) \quad \eta_{hj} = \frac{\partial X_h^*}{\partial P_j} \frac{P_j}{X_h^*} = -\beta_{hj} \frac{P_j}{P_1 X_h^*}$$

$$22) \quad \eta_{1j} = \frac{Y_i^*}{P_i} \frac{P_i}{X_1^*} = [- \sum_{j=2}^{m} \beta_{ij} \frac{P_j}{P_1^2} + \sum_{h=1}^{n} \beta_{ih} \frac{R_h}{P_1^2}] \frac{P_i}{Y_1^*}$$

A similar dual system involving $\Gamma^c = [\alpha^c \beta^c]$, $(\tilde{P}, \tilde{R}^c, \Theta)$, (Y^c, X^c) and disturbances $\varepsilon_{Yi}^c, \varepsilon_{Xn}^c$ can be written for the case where $X_n^* = 0$.

Finally, as has been demonstrated in Weaver (1977, 1982, 1983), estimates of Γ or Γ^c can be used to describe the characteristics of production such as returns to size and technical change as well as the characteristics of choice response such as price elasticities of choice.

4. Estimation of Duality Models When Truncated Choices Are Observed

To consider estimation, it will be convenient to write the model in more compact notation and focus on the choice functions in 18) and 19):

$$23) \quad Y^* = Z\Gamma + U^*$$

where $Y^* = \begin{pmatrix} Y \\ -X' \end{pmatrix}$ a $MT_1 * 1$ vector, $M = m+n-1$, T_1 is the number of observations where $Y^*_m > 0$,

$Z = \begin{pmatrix} Z_1 & & 0 \\ & \ddots & \\ 0 & & Z_M \end{pmatrix}$, a $MT_1 * \sum_{i=1}^{m} K_i$ matrix,

$Z_i = [1 \; \tilde{P} \; \tilde{R} \; \Theta]$, a $T_1 \times K_i$ matrix of the exogenous determinants of the i^{th} choice function,

$\Gamma = [\Gamma'_1 \ldots \Gamma'_M]'$,

$\Gamma_i = [\alpha'_i \; \beta'_{i\,1} \ldots \beta'_{i\,M}]$,

$U^{*\prime} = (\varepsilon'_{y_1} \ldots, \varepsilon'_{y_m} \, \varepsilon'_{x_1} \ldots \varepsilon'_{x_n}]$.

Similarly, if we consider the last product (Y^*_M) to be truncated at zero for some observations, we can summarize the dual system of NEP and choice functions for this case by:

24) $Y^c = Z^c\Gamma^c + U^c$

where notation is analogous to that used in 20) however, Y^c is $T_2 \times M-1$.

The values taken on by (Y^*,U^*) in 23) are conditional on $Y^*_M > 0$ and those taken on by (Y^c,U^c) in 24) are conditional on $Y^c_M = 0$. To define the stochastic properties of these models, we assert Y^* and Y^c are drawn from a common multivariate normal distribution of the vector Y. Similarly, U^* and U^c are drawn from the distribution of U. We further assume $E(U) = 0$, $E(UU') = \Sigma \otimes I_T$ where $T = T_1 + T_2$. Condensing 23) and 24) we have:

$$25) \; Y = \begin{cases} Z + U^* & \text{if } Y_M \geq 0 \\ \begin{pmatrix} Z^c \; \Gamma^c + U^c \\ 0 \end{pmatrix} & \text{if } Y_M = 0. \end{cases}$$

It follows that

$$26)\quad E(Y^*) = E(Y|Z, Y_M^* > 0) = Z \;\; + E(U| \, Y_M > 0).$$

Since $E(U \; Y_M^* > 0) \neq 0$, $Z\Gamma$ would not provide an unbiased estimator of $E(Y^*)$ if all observations were drawn conditional on $Y^*_m > 0$. On the other hand, if $Y^*_m = 0$, then $Z^c\Gamma^c$ would similarly fail as an unbiased estimator of $E(Y^*)$.

By an extension of Heckman's (1976) suggestion, the conditional nature of the distributions of Y* and Y^c can be summarized with an unobservable index L*. Using the first-order conditions 1)-4) and previous definitions, the following rule can be written

$$27)\quad Y_M > 0 \text{ if } -Z_M + \frac{\partial F}{\partial Y_M} = L^* > 0,$$

$$Y_M = 0 \text{ if } -Z_M + \frac{\partial F}{\partial Y_M} = L^* \leq 0.$$

Recognizing Y^c is determined by $(\tilde{P}, \tilde{R}, \Theta)$ the indicator L* can be approximated by

$$28)\quad L^* = Z\delta + \varepsilon_L$$

Although the index L* is unobservable, an observable binary indicator L can be defined as $L = 1$ if $L^* > 0$, $L = 0$ if $L^* \leq 0$.

Equations 23) and 28) fully describe choices made by the firm. To proceed, we assume the vector $[U^* \; \varepsilon_L]$ is multivariate normal and

$$E\,(U^{*\prime}\ \varepsilon_L') = 0$$

$$E \begin{vmatrix} U^{*\prime} \\ \varepsilon_L{}' \end{vmatrix} [U^*\ \varepsilon_L] = \Omega \otimes I_{T_1} = \begin{bmatrix} \Sigma_* & \sigma^* \\ \sigma & 1 \end{bmatrix}$$

where

$$\Sigma^* = E(U^{*\prime}\ U^*)$$

$$\sigma^{*2} = (\sigma^2_{U_1\varepsilon_L} \cdots \sigma^2_{U_M\varepsilon_L})$$

and by definition, $E(U^*U^c) = E(U^*)E(U^c) = 0$. A convenient estimation method follows from Heckman (1976, 1979) and Maddala, and Trost (1980), who noted that:

29) $E(U|Y^*_m > 0) = E(U|E_L > -Z\delta) = \Lambda^*\sigma^*$

where $\Lambda^* = I_m \otimes \lambda^*$, λ^* is $T_1 \times 1$ with $\lambda^*_t = \phi(-Z_t\tilde{\delta})/[1-\Phi(-Z_t\tilde{\delta})]$ and $\phi(.)$ and $\Phi(.)$ are the standard and cumulative normal density functions.

By substitution,

30) $Y^* = Z\Gamma + \Lambda^*\sigma^* + v$

where $v = U - \Lambda^*\sigma^*$

$$E(v|Y^*_m > 0) = E(v) = 0.$$

Written in this form, it is apparent that δ in 28) can be estimated by maximum likelihood probit methods. Using $\tilde{\delta}$, we can determine $\tilde{\Lambda}^*$ and estimate 30) using iterative Zellner methods. By extension of Barten's results, this method will produce MLH estimates of (Γ, σ^*) conditional on $\tilde{\Lambda}^*$. Such estimators are easily shown to be consistent. As Lee et al. (1980) have noted, variances of these estimators are conditional upon the use of $\tilde{\delta}$, covariance matrices which ignore this will underestimate the correct asymptotic variances. This follows from the fact that given $\tilde{\delta}$ we obtain residuals $U = \tilde{v} + \sigma^*\tilde{\lambda}^*$, not $U = v + \sigma^*\lambda^*$. Estimators and their properties are derived in Appendix 2. Addition of the net expected profit function 17) to the system represented by 23) involves only an appropriate generalization of notation. Finally, a similar estimation method can be motivated for the system 24) augmented with the appropriate NEP function. The independence of the drawings of U ensure that U^* and U^c will be independent, a result which allows each system to be estimated independently. An alternative maximum likelihood method of estimation can be derived as an extension of Amemiya (1974); however, the above procedure will be adopted here due to its operational convenience. By extension of the Hausman & Wise (1977) results, the consistency of the proposed estimators implies they represent attractive starting values for a MLH estimation. Defining $L^c = 1-L^*$, and by analogy with 27) a selection rule for the system of optimal choices and profit conditional on $X^*_n = 0$ as summarized in 24) can be written.

The above discussion suggests that the interpretation of the estimated coefficients of the system 30) depends upon whether β is symmetric and consistent with convexity. Secondly, there exists uncertainty prior to estimation concerning the presence of sample selection bias. If such bias does not exist, our model can be simplified by dropping the matrix Λ^*. Since our priors concerning each of these issues are diffuse, the best approach is to attempt to resolve them through data learning. Joint F-tests will be employed in the section 5 to consider variation in the sum of squared errors induced by the symmetry restriction and exclusion of the matrix Λ^* from the model.

5. The Pennsylvania Dairy Sample

Dairy farming in Pennsylvania is typically a family-operated business with an average labor force of two persons, and an average herd size of approximately 40 cows. The majority of labor is supplied by the family. Most farms produce hay, silage and maintain pastures for their own use, although home production only accounts for about 70% of the value of all feed fed. The sample of data to be employed was chosen in part to demonstrate the value of total farm revenue and expense accounts for econometric modeling. Traditionally, the presence of multiple outputs, or enterprises, has led researchers to seek enterprise specific expense data. However, the above methodology allows a complete characterization of production and choice response forcing the researcher to allocate expenses to particular outputs. The data represents a sample of 117 Pennsylvania dairy farms which were individually enumerated in the 1974 USDA Cost of Production study solicited by Congress. Detailed data were collected on prices, quantities and qualities of inputs and outputs. For example, brand specific data were collected for agricultural chemicals. Products were aggregated using static forms of the Divisia index.

As noted by Sims (1974) and more recently by Weaver (1983), the specification of which products are variable within the observation period is critical for the interpretation of dual choice models. For the present sample of dairy farms we maintain the hypothesis that net milk and net grain crop outputs are variable in the short-run. Variable inputs are hypothesized to be commercial fertilizer and lime, herbicides, commercial feeds, hired labor, capital services (buildings and machinery), other livestock inputs and other crop inputs. The size of the dairy herd is assumed to directly affect production possibilities and is hypothesized to be variable in the short-run through sales or purchases of dairy cows. Acres operated and milk storage capacity were hypothesized to be fixed in the short-run due to the absence of short-run rental markets. Finally, production possibilities are hypothesized to be conditional upon farm operator characteristics measured by age of operator, experience and education. Definitions of these categories and characteristics of the sample are described in more detail in Appendix 1.

6. Results

First consider the results for the subsample of farms that employed hired labor. Elasticities of choice are derivable from the profit function on the condition that we cannot reject the estimated choice functions as derivatives of a quadratic approximation to $*(.)$. Before investigating whether the estimated system of choice functions are consistent with profit maximization, the hypothesis of the existence of sample selection bias was tested by restricting $\pi^* = 0$. As can be seen in Table 2, the restriction is rejected at the 99% level of significance. Conditional upon the inference that selection bias exists in the present sample * was retained in the model. As noted above in section 2, if the estimated choice functions are consistent with profit maximization, then they should be monotonic in prices and consistent with a profit function which is convex in prices, continuous in (P,R,Θ), and therefore, exhibits symmetry. Alternative specification tests were reviewed in detail in Weaver (1982, 1983). For present purposes, monotonicity and convexity were checked at each observation and the estimated choice functions were found to be consistent with these properties. Symmetry was tested by restricting to be symmetric. Since this amounts to imposing a set of linear restrictions on β, the percentage change in the system's weighted sum of squared error is distributed as an F-statistic. As seen in Table 2, symmetry cannot be rejected at the 95% level of significance though it can be at the 99% level. We chose to proceed by maintaining the hypothesis of symmetry and interpreting the estimated coefficients as first- and second-derivatives of the profit function evaluted at the origin.

Table 3 reports estimated parameters and associated one-tailed t-statistics based on estimated asymptotic variances. An immediately apparent conclusion is that own price effects were in general highly significant and had signs consistent with profit-maximization; however, cross-price effects and the effects of fixed factors were also highly significant.

Elasticities of choice with respect to prices and fixed factor changes were calculated and their values at the means of the data set are reported in Table 4. These represent the first complete set of dairy production choice elasticities based on micro level data making validation by comparison with past results is only of limited usefulness. The short-run elasticity of milk is estimated to be 31.39%. Lime and fertilizer demand shows substantial own-price elasticity. All other own price elasticities of input demand are less than one. It is of interest to note the apparent strong inelasticity of milk with respect all prices except its own. This would suggest cull prices and feed prices may be weak instruments with which to control milk supply.

Table 3 indicated significant effects of fixed factors on production choices. Weaver (1977, 1978, 1982) has reported a variety of results concerning the effects of changes in fixed factors on relative product choice. As Weaver (1978) noted, contrary to Binswanger's (1974) claim in the multiple product case single choice elasticities cannot reveal biases or changes in product mix introduced by changes in fixed factors. Nevertheless, absolute effects are of interest if properly interpreted. As seen in Table 3, herd size has a substantial positive effect on livestock sales, fertilizer use, and commercial feed use. Large positive elasticities of net crop output with respect to experience and education were found. Following Weaver (1983) the effect of changes in fixed input levels on product mixes can be thought of in Hicksian terms. Measures of biases are reported in Table 4.

Appendix 2. A Consistent Estimation Method of Mixed Systems of Truncated and Continuous Choice Functions

Consider the first regime where $Y_m > 0$. Consistent parameter estimates are obtained by applying iterative Zellner's methods to 20), Barten (1969). The proper covariance matrice for the parameters is obtained by using 20), 27) and 29) to note:

$$\text{(A.1)} \quad U^* = v^* + \Lambda^* \sigma^*.$$

However, since the probit results $\tilde{\delta}$ are used as estimates of δ in Λ, we have:

$$\text{(A.2)} \quad U^* = \tilde{v}^* + \tilde{\Lambda}^* \sigma^*.$$

(A.1) and (A.2) imply

$$\text{(A.3)} \quad \tilde{v}^* = v^* - (\tilde{\Lambda}^* - \Lambda)\sigma^*.$$

The effects of using estimates of is clearly indicated by (A.3). Following Lee, et al., the covariances for the residuals are obtained from:

$$\text{(A.4)} \quad E(\tilde{v}^*\tilde{v}^*) = \{\text{Var}\ (v^*) + (\sigma^*)^2\ AZ\ [\text{Var}(\tilde{\delta})]\ Z'A$$
$$-\sigma^*\ AZ\ \text{Cov}\ (\tilde{\delta}, v^{*\prime}) - \sigma^*\ \text{Cov}\ (\tilde{\delta}', v^*)\ Z'A\}$$

where $A = \text{diag.}[-Z_t \tilde{\delta}(\tilde{\lambda}^*_t) - (\tilde{\lambda}^*_t)^2] \qquad (t = 1,\ldots,T_1)$

Using the results (Lee, et al., p. 500):

(A.5) $\text{Cov}\ (\tilde{\delta}, v^*) = 0$

(A.6) $\text{Var}\ (\tilde{\delta}) = (Z_T' S Z_T)^{-1}$

(A.7) $\text{Var}\ (v^*) = \Sigma_{v^*v^*} + (\sigma^*)^2 A$

we can write (A.4) as:

(A.8) $E(\tilde{v}^*\tilde{v}^{*\prime}) = \{\Sigma_{v^*v^*} + (\sigma^*)^2 A + (\sigma^*)^2\ AZ(Z_T' S Z_T)^{-1} Z' A\}$

where Z_T is the matrice of regressors for the entire sample used in probit estimation,

$$\Sigma_{v^*v^*} = \begin{bmatrix} E(V_1V_1) & E(V_1V_2) \cdots & E(V_1V_M) \\ & E(V_2V_2) \cdots & E(V_2V_M) \\ & & \ddots E(V_MV_M) \end{bmatrix} \otimes I_{T_1}$$

$$S = \text{diag.}[\phi_t(.)/\Phi^* s' t.(.)(1-\Phi_t(.)] \qquad (t = 1, \ldots, T)$$

Using (A.8) the proper covariance for the Zellner's estimators can be shown to be:

(A.9) $$\text{Cov}\begin{bmatrix} \hat{\Gamma} \\ \hat{\sigma}^* \end{bmatrix} = \{[Z\Lambda]'(\Sigma_{v^*v^*}^{-1})[Z\Lambda]\}^{-1} + (\hat{\sigma}^*)^2\ \{([Z\Lambda]'(\Sigma^{-1}_{v^*v^*})[Z\Lambda])^{1}$$
$$\times\ [Z\Lambda]'[A + AZ(Z_T' S Z_T)^{-1}\ Z'A]\ [Z\ ]\ ([Z\Lambda]'(\Sigma^{-1}_{v^*v^*})[Z\Lambda])\}.$$

The first term in the braces is the covariance matrix typically reported by software packages. The remainder represents the amount by which the asymptotic covariances are understated.

The exact prior symmetry restrictions ($R\beta = 0$) can be introduced to achieve the following restricted estimators:

$$\text{(A.10)} \quad \begin{bmatrix} \hat{\Gamma}^R \\ \hat{\sigma}^{*R} \end{bmatrix} = \begin{bmatrix} \hat{\Gamma} \\ \hat{\sigma}^* \end{bmatrix} + B[Z|\Lambda|R]'G(G'[Z\Lambda R]B[Z\Lambda R]'G)^{-1}(-G'[Z\Lambda R] \begin{matrix} \hat{\Gamma} \\ \hat{\sigma}^* \end{matrix})$$

where $\begin{bmatrix} \hat{\Gamma} \\ \hat{\sigma}^* \end{bmatrix}$ are the unrestricted estimators,

R = the appropriate restriction matrix,

$B = ([Z\Lambda R]' \,{}^{+}[Z\Lambda R])^{-1}$ is the unrestricted covariance matrix,

Ψ^+ = the generalized inverse of the covariance matrix $E(\tilde{e}\tilde{e}') =$ $E([\begin{smallmatrix}\tilde{v} \\ 0\end{smallmatrix}][\tilde{v}0])$ (See: Judge, et al., pp. 278-280), and

$U = [FG]$ is a matrix of characteristic vectors corresponding to the roots of $E(\tilde{e}\tilde{e}')$.

The covariance matrix for the restricted estimators can be derived in analogy to A.G:

$$\text{(A.11)} \quad \text{Cov}\begin{bmatrix} \hat{\Gamma}^R \\ \hat{\sigma}^{*R} \end{bmatrix} = \underline{M}B\underline{M} + \underline{M}B[Z\Lambda R]'\Psi^+ \{(\underline{\hat{\sigma}^*})^2 A + (\underline{\hat{\sigma}^*})^2 AZ(Z_T' S Z_T)^{-1} Z'A\} \times \Psi^+ [Z\Lambda R]B\underline{M}$$

where $\underline{M} = I - B[Z\Lambda R]'G(G'[Z\Lambda R]B[\;R]^{-1}(G'[Z\Lambda R])$

$\underline{\hat{\sigma}^*} = \begin{bmatrix} \hat{\sigma}^* \\ 0_q \end{bmatrix}$ and q is the number of restrictions.

Table 1. Specification Tests of Sample Selection Bias and Symmetry

Hypothesis	Sample Selection	Symmetry Conditional on the Existence of Selection Bias
General Models	$H_o: Y = X\beta + U$ $H_A: Y = X\beta + \Lambda\gamma + V$	$H_o: Y = X\beta + \Lambda\gamma + V$ $H_A: Y = X\beta + R\beta + \Lambda\gamma + W$
Test Statistic	$\frac{U'(\Sigma^{-1})U - V'(\Sigma^{-1})V}{V'(\Sigma^{-1})V} \cdot \frac{590}{10}$ $\sim F(10,590)$	$\frac{W'(\Sigma^{-1})W - V'(\Sigma^{-1})V}{V'(\Sigma^{-1})V} \cdot \frac{590}{45}$ $\sim F(45,590)$
Value of Test Statistic	2.4733	1.5784
Critical Value	$F_{01}(10,\infty) = 2.32$	$F_{01}(45,\infty) = 1.59$ $F_{05}(45,\infty) = 1.39$
Inference	Reject H_o	Fail to Reject H_o Reject H_o