Title

Econometric Models of Discrete/Continuous Supply Decisions under Uncertainty

Authors

Hanemann, W. Michael
Tsur, Yacov

Publication Date

1982-03-01

Working Paper No. 195

ECONOMETRIC MODELS OF DISCRETE/CONTINUOUS SUPPLY DECISIONS UNDER UNCERTAINTY

by

W. Michael Hanemann and Yacov Tsur

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
March 1982

ECONOMETRIC MODELS OF DISCRETE/CONTINUOUS SUPPLY DECISIONS UNDER UNCERTAINTY

By W. Michael Hanemann and Yacov Tsur

1. Introduction

This paper is concerned with the theoretical specification and estimation of econometric models of discrete/continuous supply decisions by economic agents who face uncertainty with respect to output prices and/or yields. Until very recently, economic theorists and econometricians have paid very little attention to the discreteness of the choices made by individual agents; they have focused almost exclusively on purely continuous choices. Yet in reality discrete choices, and discrete choices interrelated with continuous choices, are a pervasive phenonomenon. For example, a producer may be able to produce many crops, but may choose to specialize in some of them; the decision to grow certain crops is the discrete choice, while the decision how much of the selected crops to plant is the continuous choice. Another example is where a producer is growing a certain crop and faces, say, a binary choice between two mutually exclusive techniques for growing the crop — e.g., whether to employ one type of irrigation technology or another, or whether to grow the crop under a federal government commodity program or not. The discrete choice is which technology to use or whether to participate in the program; the continuous choice is how much of the crop to plant. A third example is where an input to the production of a crop, say a drying or storage facility, is available in only two or three alternative sizes; the discrete choice is which size facility to invest in, and the continuous choice is how much of the crop to plant.

In all of these cases the discrete and the continuous choices are interrelated: the optimal continuous decision depends on the outcome of the discrete choice, and vice versa. Therefore, both choices must be modeled simultaneously. We face two challenges in seeking to do this. The first is the problem of developing theoretical models to explain the discrete and continuous choices in a manner which is consistent with a single underlying utility-of-profit maximization decision. The second is the problem of developing statistical techniques for estimating the resulting discrete and continuous supply functions. These issues are discussed, respectively, in sections 2 and 3.

In section 2 we extend Hanemann's [5 , 6] analysis of discrete and continuous consumer demand models to the context of producer supply under uncertainty. The key concept here is the notion of a "random supply" model, in which some components of the supplier's production or utility-of-profit functions are treated as being random from the viewpoint of the econometric investigator. It is this random component which generates the stochastic structure employed in the estimation of the discrete and continuous supply functions. The estimation is discussed in section 3 in a very general framework, which covers not only the specific supply models developed in section 2 but also other supply models and demand models such as those developed by Duncan [3] and McFadden [13]. We show that these models can be regarded as special types of switching regression models which involve an N-fold switching instead of the binary switching that is commonly assumed. Moreover, because the discrete and continuous supply functions both flow from the same underlying theoretical model of optimization by an economic agent, there are additional restrictions on the coefficients and disturbance terms appearing in these functions. We show how these restrictions can be exploited in an efficient estimation procedure. Finally, in section 4 we summarize our conclusions and suggest some ways in which the analysis of this paper can be further extended.

2. Theoretical Models

The purpose of this section is to explain the general structure of discrete/continuous models of supply under uncertainty, and to motivate their estimation by showing how they can be cast in the form of a statistical model of switching regression. In order to create a statistical framework for estimating these models it is necessary to postulate that some component of the supplier's production or utility function is random from the viewpoint of the econometric investigator. Before describing this random supply model, however, it is convenient to begin by summarizing a "deterministic" model of supply under uncertainty, where this random component is absent — this is done in section 2A. The corresponding random supply model is presented in section 2B.

A. Deterministic Models

We shall first summarize the standard deterministic model of purely continuous supply under uncertainty and then generalize this to the case of discrete/continuous supply choices under uncertainty. We focus on the special case of a supplier of a single product, who faces no explicit constraints on his production decision (such as a limit on the availability of land or credit). His profit, π, is given by

(2.1) $$\pi = pq - c(w, q) - b$$

where p is the product price, q is the amount of product supplied, $c(\cdot)$ is a variable total cost function generated by some production function, and b is fixed costs. We assume that the producer faces uncertainty with respect to the product price. His subjective density will be denoted f_p, with mean μ and variance σ_p^2.[1] The producer has a utility-of-profit function, $u(\pi)$, with $u' > 0$, and $u'' \gtrless 0$ depending whether he is risk-prone, riskneutral, or risk-averse. To allow for the possibility that his risk preferences depend in a parametric manner on his individual characteristics, s, we shall write $u = u(\pi; s)$.

The producer chooses an output level, q, so as to maximize his expected utility

$$(2.2) \qquad \max_q \upsilon(q) \equiv \max_q \int u(\pi(q, p); s)\, f_p dp \ .$$

The solution to the producer's maximization problem will be denoted $q(\mu, w, b; s)$. Substituting this into the maximand in (2.2) yields the <u>indirect</u> <u>expected</u> <u>utility-of-profit</u> <u>function</u>,

$$(2.3) \qquad v(\mu, w, b; s) = \upsilon(q(\mu, w, b; s)).$$

By a standard application of the envelope theorem it can be shown that

$$(2.4) \qquad \frac{\partial v(\mu, w, b; s)}{\partial \mu} = q(\mu, w, b; s)E\{u'\}$$

$$(2.5) \qquad \frac{\partial v}{\partial b}(\mu, w, b; s) = -E\{u'\}$$

where

$$E\{u'\} = \int u'[pq(\mu, w, b; s) - c(w, q(\mu, w, b; s)) - b; s]\, f_p dp \ .$$

Hence, we have the equivalent of Hotelling's lemma for production decisions under uncertainty

$$(2.6) \qquad q(\mu, w, b; s) = -\frac{\partial v(\mu, w, b; s)/\partial \mu}{\partial v(\mu, w, b; s)/\partial b} \ .$$

It follows that, as with the theory of supply under certainty, there are two methods for generating a particular parametric supply model. The direct (primal) approach is to specify a particular utility function and density, f_p, and then solve the resulting maximization problem (2.2) for q(•) and v(•). The indirect (dual) approach is to start by specifying an indirect expected utility-of-profit function, v(•), which satisfies the appropriate requirements for such a function, and then to derive the output supply function from (2.6). An alternative, related, approach is to start by specifying an output supply function, q(•), which satisfies the appropriate requirements for such a function, and then integrating (2.4) and (2.5) to obtain the corresponding indirect expected utility-of-profit function. The feasibility of this last approach remains to be seen.[2]

This completes our account of the standard, deterministic, purely continuous model of supply under uncertainty. Before proceeding to the corresponding discrete/continuous supply model, we pause to give an example based on the primal approach. This particular example will be continued throughout this paper.

Example.

We assume constant returns to scale with respect to the variable inputs, so that the total variable cost curve can be written

(2.7) $$c(w, q) = c(w)q$$

where c(w) is a unit variable cost function. We also assume constant absolute risk aversion:

(2.8) $$u(\pi; s) = 1 - e^{-\alpha(s)\pi}$$

where $\alpha(s)$, the absolute risk aversion coefficient, is allowed to vary with the characteristics of the producer. In particular, if the producer's wealth is one of these characteristics, this formulation allows for the possibility of, say, absolute risk aversion declining with wealth across individuals while being constant for a given producer making a given risky decision. It follows from (2.8) that expected utility is

(2.9) $\quad \upsilon(q) = 1 - M_p(-\alpha(s)q)e^{\alpha(s)c(w)q+\alpha(s)b}$

where $M_p(\cdot)$ is the moment generating function associated with f_p. If $f_p \sim N(\mu, \sigma_p^2)$, then

$$M_p[-\alpha(s)q] = \exp[-\alpha(s)\mu q + \frac{\alpha(s)^2q^2\sigma_p^2}{2}$$

and expected utility becomes

(2.10) $\quad \upsilon(q) = 1 - \exp[-\alpha(s)(\mu q - c(w)q - b) + (\alpha(s)^2q^2\sigma_p^2)/2].$

Alternatively, if f_p is the gamma distribution with parameters γ, λ, where $\mu = \gamma/\lambda$ and $\sigma_p^2 = \gamma/\lambda^2$, then

$$M_p[-\alpha(s)] = (\frac{\mu}{\sigma_p^2})^{(\mu/\sigma_p)^2}[\frac{\mu}{\sigma_p^2} + \alpha(s)q]^{-(\mu/\sigma_p)^2}$$

and

(2.11) $\quad \upsilon(q) = 1 - (\frac{\mu}{\sigma_p^2})^{(\mu/\sigma_p)^2}[\frac{\mu}{\sigma_p^2} + \alpha(s)q]^{-(\mu/\sigma_p)^2}e^{\alpha(s)c(w)q+\alpha b}$.

For the normal case, the maximization of (2.10) yields

(2.12) $\quad q(\mu, w, b; s) = \frac{\mu - c(w)}{\sigma_p^2} \cdot \frac{1}{\alpha(s)}$.

Thus supply increases with expected unit profit $(\mu - c(w))$, and decreases with the variance of prices, σ_p^2, and the supplier's risk aversion, α. Substituting (2.12) into (2.10) yields the indirect expected utility-of-profit function

(2.13) $\quad v(\mu, w, b; s) = 1 - \exp[\alpha(s)b - \frac{(\mu - c(w))^2}{2\sigma_p^2}]$.

Alternatively, in the gamma case maximization of (2.11) yields

(2.14) $\quad q(\mu, w, b; s) = [\frac{\mu - c(w)}{\sigma_p^2}][\frac{\mu}{\alpha(s)c(w)}]$

$$(2.15) \qquad v(\mu, w, b; s) = 1 - \left(\frac{\mu}{c(w)}\right)^{-(\mu/\sigma_p)^2} \exp\left[\frac{\mu}{\sigma_p^2}(\mu - c(w)) + \alpha(s)b\right] .$$

We now introduce the possibility of a discrete choice by the producer in addition to the continuous supply decision discussed above. Specifically, we assume that the producer faces N mutually exclusive discrete choices. Examples of such discrete choices might be: which of N mutually exclusive production technologies to employ; in which of N mutually exclusive locations to produce; which of N mutually exclusive types of fixed equipment to use alongside of the variable inputs; or whether or not to participate in a federal government commodity program. In general, we can assume that each discrete alternative j presents the producer with a particular vector of variable input prices, w_j; a particular variable cost function, $c_j(w_j, q)$; a particular fixed cost, b_j; and a particular distribution of output prices, $f_{p_j}(p)$, with mean μ_j and variance σ_{jp}^2. We also allow for the possibility that the producer's individual characteristics, s, may vary with the discrete choice, and hence that his utility function u(•), may vary with j.

Suppose, for the moment, that the producer has decided to select the j^{th} discrete alternative. Conditional on this decision, his profit is

$$\pi_j = p_j q_j - c_j(w_j, q_j) - b_j ,$$

where q_j is his output under the j^{th} discrete alternative, and his expected utility is

$$(2.16) \qquad \bar{U}_j(q_j) = \int u_j(\pi_j(q_j, p_j); s_j) f_{p_j} dp_j .$$

His continuous supply decision conditional on this discrete choice is $\bar{q}_j(\mu_j, w_j, b_j; s_j)$ which is obtained by maximizing (2.16). His expected utility on making this supply decision is

(2.17) $$\bar{v}_j(\mu_j, w_j, b_j; s_j) = \bar{U}_j(\bar{q}_j(\mu_j, w_j, b_j; s_j)).$$

It is evident from this derivation that the conditional output supply function, $\bar{q}_j(\cdot)$, and the conditional indirect expected utility-of-profit function, $\bar{v}_j(\cdot)$, have all the standard properties of an output supply function and an indirect expected utility-of-profit function as outlined above. In particular, the relation (2.6) carries over

(2.18) $$\bar{q}_j(\mu_j, w_j, b_j; s_j) = -\frac{\partial\bar{v}_j(\mu_j, w_j, b_j; s_j)/\partial\mu_j}{\partial\bar{v}_j(\mu_j, w_j, b_j; s_j)/\partial b_j}.$$

All of the foregoing is conditional on the producer's selecting discrete alternative j. His discrete choice can be represented by a set of binary valued indices, $d_1, \ldots, d_N$, where $d_j = 1$ if alternative j is selected, and $d_j = 0$ otherwise. His overall continuous and discrete maximization problem is to select $q_1, \ldots, q_N$ and $d_1, \ldots, d_N$ so as to maximize

(2.19) $$\sum_1^N d_j\bar{U}_j(q_j)$$

subject to

$$d_j = 0 \text{ or } 1, \ \Sigma d_j = 1 .$$

The solution for the discrete choices, denoted $d_j = d_j(\mu_1, \ldots, \mu_N, w_1, \ldots, w_N, b_1, \ldots, b_N; s_1, \ldots, s_N)$ or, more compactly, $d_j = d_j(\mu, w, b; s)$, $j = 1, \ldots, N$, are functions of the full set of input costs and output prices. Similarly, the solution for the continuous choices — the unconditional supply functions — will be denoted $q_j = q_j(\mu, w, b; s)$. Finally, the unconditional indirect expected utility-of-profit function is $v(\mu, w, b; s)$, defined as

(2.20) $$v(\mu, w, b; s) = \sum_1^N d_j(\mu, w, b; s)\bar{U}_j(q_j(\mu, w, b; s)).$$

These unconditional functions are related to the conditional functions derived above in the following manner:

$$(2.21) \qquad d_j(\mu, w, b;\ s) = \begin{cases} 1 \text{ if } \bar{v}_j(\mu_j, w_j, b_j;\ s_j) \geq \bar{v}_i(\mu_i, w_i, b_i;\ s_i), \text{ all } i \\ 0 \text{ otherwise} \end{cases}$$

$$(2.22) \qquad q_j(\mu, w, b;\ s) = d_j(\mu, w, b;\ s)\ \bar{q}_j(\mu_j, w_j, b_j;\ s_j)$$

$$(2.23) \qquad v(\mu, w, b;\ s) = \max\{\bar{v}_1(\mu_1, w_1, b_1;\ s_1), \ldots, \bar{v}_N(\mu_N, w_N, b_N;\ s)\}.$$

These relations are of considerable practical importance, because they can be used to derive the properties of the unconditional functions from those of the conditional ones. However, this will not be explored further here. Instead we now turn to the random supply models, after first continuing the example started above.

Example.

For the model (2.7) - (2.9), under the assumption that the p_j's are independently distributed with $f_{p_j} \sim N(\mu_j, \sigma^2_{jp})$, the conditional output supply and indirect expected utility-of-profit functions are

$$(2.24) \qquad \bar{q}_j(\mu_j, w_j, b_j;\ s_j) = \frac{\mu_j - c_j(w_j)}{\alpha_j(s_j)\ \sigma^2_{jp}}$$

$$(2.25) \qquad \bar{v}_j(\mu_j, w_j, b_j;\ s_j) = 1 - \exp\{\alpha_j(s_j) b_j - \frac{[\mu_j - c_j(w_j)]^2}{2\sigma^2_{jp}}\}.$$

The corresponding unconditional functions can be obtained by direct application of (2.21) - (2.23).

B. Random Supply Models

A random supply model arises when one assumes that, although all the elements of the producer's decision — his cost function, his subjective probability density for the output price, and his own utility-of-profit function —

are known for sure to him, they contain some components which are unobservable to the econometric investigator, and are treated by the investigator as random variables. This formulation embodies two notions which, for practical purposes, are indistinguishable: the idea of a variation in technology, information or preferences among a population of individual economic agents, and the concept of unobserved variables in econometric models. These components will be denoted by ε^c, ε^p and ε^u, which may be scalars or vectors. In each case, they are fixed constants (or functions) for the individual producer, but for the investigator they are random variables. For example, because of unobservables or inter-agent variation in the production technology, the individual producer's cost function appears to the investigator to be of the form $c_j(w_j, q_j; \varepsilon_j^c)$; or, because of differences in perceptions among producers, the individual producer's subjective probability density for output prices appears to the investigator to be of the form $f_{p_j}(p_j; \varepsilon_j^p)$; or, finally, because of variations in risk preferences or unobservable components in profits (including fixed costs), the individual producer's utility-of-profit function appears to the investigator to be of the form $u_j(\pi_j; s_j, \varepsilon_j^u)$.[3]

One can generate different random production models depending on which of these sources of randomness one chooses to emphasize and on how one incorporates them. In order to avoid committing ourselves at this point to a specific random production model, we will refer to these random components collectively as ε_j; ε_j could be ε_j^c, ε_j^p, ε_j^u, or some combination of them. Accordingly, we write the direct expected utility-of-profit function associated with the j^{th} discrete alternative in general terms as $\bar{U}_j(q_j; \varepsilon_j)$. A similar set of random terms exists for each discrete alternative. We denote the overall set of random terms by $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)$. For the econometric investigator this is a multivariate random variable with some joint density function, denoted $f_\varepsilon(\varepsilon_1, \ldots, \varepsilon_N)$; for the individual producer, however, it is a set of fixed constants.

The individual producer's decision problem is to select $q_1, \ldots, q_N$ and $d_1, \ldots, d_N$ so as to maximize

$$\sum_1^N d_j \bar{\upsilon}_j(q_j;\ \varepsilon_j) \tag{2.26}$$

subject to $d_j = 0$ or 1, $\Sigma d_j = 1$.

The supply functions generated by this maximization problem parallel those developed in the previous section, except that they now involve a random component from the point of view of the econometric investigator. Suppose the producer has decided to select the j^{th} discrete alternative. If he maximizes $\bar{\upsilon}_j(q_j;\ \varepsilon_j)$ this yields the conditional supply function, $\bar{q}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j)$ and the conditional indirect expected utility-of-profit function, $\bar{v}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j) \equiv \bar{\upsilon}_j[\bar{q}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j), \varepsilon_j]$. These still have the properties mentioned in the previous section; in particular,

$$\frac{\partial \bar{v}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j)}{\partial \mu_j} = \bar{q}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j) \cdot E\{u'_j\} \tag{2.27}$$

$$\frac{\partial \bar{v}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j)}{\partial b_j} = -E\{u'_j\} \tag{2.28}$$

$$\bar{q}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j) = -\frac{\partial \bar{v}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j)/\partial \mu}{\partial \bar{v}_j(\mu_j, w_j, b_j;\ s_j, \varepsilon_j)/\partial b_j}\ . \tag{2.29}$$

The quantities $\bar{q}_j$ and $\bar{v}_j$ are known numbers to the producer but, because his decision is incompletely observed, they are random variables for the investigator. In particular, let $f_{\bar{v}}(\bar{v}_1, \ldots, \bar{v}_N)$ be the joint density of $\bar{v}_1, \ldots, \bar{v}_N$ induced by $f_\varepsilon(\cdot)$.[4]

Similarly, the unconditional discrete choice indices generated by the solution of (2.26), $d_j(\mu, w, b; s, \varepsilon)$, $j = 1, \ldots, N$, are random variables. Let $\bar{z}_i = \bar{v}_i - \bar{v}_j$, $i \neq j$ and let $F_{\bar{z}}(z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_N)$ be the joint c.d.f. of the $\bar{z}_i$'s. Then the mean of the expected value of the discrete choice indices, $E\{d_j\} \equiv P^j$, is

$$(2.30) \qquad P^j(\mu, w, b; s) = \Pr\{\bar{v}_j(\mu_j, w_j, b_j; s_j, \varepsilon_j) \geq \bar{v}_i(\mu_i, w_i, b_i; s_i, \varepsilon_i), \text{ all } i\}$$

$$= F_{\bar{z}}(0, \ldots, 0).$$

The unconditional supply functions generated by (2.26) denoted $q_j(\mu, w, b; s, \varepsilon)$, $j = 1, \ldots, N$, are also random variables, as is the unconditional indirect expected utility-of-profit function obtained by substituting these unconditional supply functions and the discrete choice functions into the maximand in (2.26); this will be denoted $v(\mu, w, b; s, \varepsilon)$. These unconditional functions are related to the conditional functions by formulas similar to those for the deterministic production model:

$$(2.31) \qquad q_j(\mu, w, b; s, \varepsilon) = d_j(\mu, w, b; s, \varepsilon)\bar{q}_j(\mu_j, w_j, b_j; s_j, \varepsilon_j)$$

$$(2.32) \qquad v(\mu, w, b; s, \varepsilon) = \max\{\bar{v}_1(\mu_1, w_1, b_1; s_1, \varepsilon_1), \ldots, \bar{v}_N(\mu_N, w_N, b_N; s_N, \varepsilon_N)\}.$$

In order to construct the probability distributions of these random variables, we introduce the sets $A_j = \{\varepsilon | \bar{v}_j(\mu_j, w_j, b_j; s_j, \varepsilon_j) \geq \bar{v}_i(\mu_i, w_i, b_i, s_i, \varepsilon_i), \text{ all } i\}$, $j = 1, \ldots, N$. Let $f_{\varepsilon|\varepsilon\in A_j}$ be the conditional joint density of $\varepsilon_1, \ldots, \varepsilon_N$ given that $\varepsilon \in A_j$; i.e., given that discrete alternative j is selected. Then the probability density of $\bar{q}_j$, i.e., the conditional probability $\Pr\{q_j = q | q_j > 0\}$, denoted $f_{q_j|q_j>0}(q)$, can be obtained by a change of variable from $f_{\varepsilon|\varepsilon\in A_j}$ using (2.29). The probability density of q_j, i.e., the unconditional probability $\Pr\{q_j = q\}$, denoted $f_{q_j}(q)$, therefore has the form

$$(2.33)\qquad f_{q_j}(q) = \begin{cases} 1 - P^j & q = 0 \\ f_{q_j|q_j>0}(q)\cdot P^j & q > 0 . \end{cases}$$

Thus, given a sample of T producers, where j^* is the index of the discrete choice selected by the t^{th} producer and q^*_t is his observed supply, the likelihood function of the sample is, from (2.33),

$$(2.34)\qquad L = \prod_{t=1}^{T} \{P_t^{j^*}\cdot f_{q_{j^*t}|q_{j^*t}>0}(q_t^*)\} .$$

Finally, let $f_{\bar{v}_j|\varepsilon\in A_j}(s)$ be the conditional density of $\bar{v}_j$, i.e., the conditional probability $\Pr\{\bar{v}_j = s|\bar{v}_j \geq \bar{v}_i$, all i}; this is obtained from conditional joint density $f_{\varepsilon|\varepsilon\in A_j}$ by a transformation of variables. The mean of this distribution is

$$(2.35)\qquad E\{\bar{v}_j|\bar{v}_j \geq v_i,\ \text{all } i\} = \int_{-\infty}^{\infty} s f_{\bar{v}_j|\varepsilon\in A_j}(s)ds.$$

From (2.32) the mean of the distribution of $v(\mu, w, b; s, \varepsilon)$ is given by

$$(2.36)\qquad E\{v\} = \sum_j E\{\bar{v}_j|\bar{v}_j \geq \bar{v}_i,\ \text{all } i\}\cdot P^j .$$

This completes our account of the general structure of random supply discrete/continuous choice models. The crucial ingredients in these models are the conditional indirect expected utility-of-profit functions, $\bar{v}_j(\mu_j, w_j, b_j; s_j, \varepsilon_j)$, $j = 1, \ldots, N$, and the joint density $f_\varepsilon(\cdot)$. With these one can construct the densities $f_v(\cdot)$, $f_{\varepsilon|\varepsilon\in A_j}(\cdot)$, which are used to form the discrete choice probabilities and the conditional and unconditional densities of the q_j's. As noted above, different random supply models can be generated by allowing the ε_j's to enter the conditional indirect expected utility-of-profit functions in different ways or by making different assumptions about their joint distribution, but these models will all conform to the general structure outlined above.

The estimation of these discrete/continuous supply models will be discussed in detail in the next section. However, we will make one general comment here. First we must draw attention to an alternative way of representing the unconditional supply functions, besides (2.31). Purely for notational convenience we consider the case where $N = 2$.[5] The unconditional supply functions may be written

$$(2.37) \qquad q = \begin{cases} -\dfrac{\partial \bar{v}_1(\mu_1, w_1, b_1; s_1, \varepsilon_1)/\partial \mu_1}{\partial \bar{v}_1(\mu_1, w_1, b_1; s_1, \varepsilon_1)/\partial b_1}, & \text{if } \bar{v}_1(\mu_1, w_1, b_1; s_1, \varepsilon_1) \geq \bar{v}_2(\mu_2, w_2, b_2; s_2, \varepsilon_2) \\ -\dfrac{\partial \bar{v}_2(\mu_2, w_2, b_2; s_2, \varepsilon_2)/\partial \mu_2}{\partial \bar{v}_2(\mu_2, w_2, b_2; s_2, \varepsilon_2)/\partial b_2} & \text{otherwise}. \end{cases}$$

Since the $\bar{v}_j$'s are functions of several variables — μ_j, w_j, b_j, s_j, etc. — it is convenient at this point to refer explicitly to the coefficients of these variables, which we denote by the vector β. Therefore, we now write the conditional indirect expected utility-of-profit functions as $\bar{v}_j(\mu_j, w_j, b_j; s_j, \varepsilon_j, \beta)$. Then (2.37) can be written symbolically as

$$(2.38) \qquad q = \begin{cases} g_1(\mu_1, w_1, b_1; s_1, \varepsilon_1, \beta) & \text{if } h(\mu_1, \mu_2, w_1, w_2, b_1, b_2; s_1, s_2, \varepsilon_1, \varepsilon_2, \beta) \geq 0 \\ g_2(\mu_2, w_2, b_2; s_2, \varepsilon_2, \beta) & \text{otherwise} \end{cases}$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are the ratios of the derivatives of $\bar{v}_1(\cdot)$ and $\bar{v}_2(\cdot)$, and $h(\cdot) \equiv \bar{v}_1(\cdot) - \bar{v}_2(\cdot)$; these functions will be linear or nonlinear in β depending upon the underlying structure of the $\bar{v}_j(\cdot)$ functions.

The purpose of the formulation (2.38) is to demonstrate how our theoretical random supply model generates a statistical switching regression model. The general (binary) single-equation switching regression model can be written in the form

$$(2.39) \qquad Y = \begin{cases} G_1(X_1; \beta_1, \xi_1) & \text{if } H(Z; \gamma, \eta) \geq 0 \\ G_2(X_2; \beta_2, \xi_2) & \text{otherwise} \end{cases}$$

where Y is the dependent variable, X_1, X_2 and Z are exogenous variables, β_1, β_2, and γ are the coefficients to be esimated, and ξ_1, ξ_2, and η are random error terms. Our supply model (2.38) is clearly a special case of (2.39) where, because the discrete and continuous choices both flow from the same underlying expected utility-of-profit maximization problem, the variables X_1 and X_2 are known transformations of the variables in Z, the coefficients β_1 and β_2 are the same as the coefficients γ, and the random terms ξ_1 and ξ_2 are directly related to the random term η. We can therefore estimate the random supply model (2.38) by any of the techniques developed for the switching regression model (2.39), while taking advantage of the special structure of our model. This is the subject of the following section. To illustrate the ideas discussed above in a more concrete form we now continue and terminate the example started in section 2.A.

Example.

Our starting point is the deterministic discrete/continuous supply (2.24) and (2.25). To allow for the unobservable elements which are treated by the econometric investigator as random variables, we might in general write

$$c_j(w_j;\ \varepsilon_j^c) = \hat{c}_j(w_j) + \varepsilon_j^c \tag{2.40a}$$

$$\sigma_{jp}^2(\varepsilon_j^p) = \hat{\sigma}_{jp}^2 + \varepsilon_j^p \tag{2.40b}$$

$$\alpha_j(s_j;\ \varepsilon_j^u) = \hat{\alpha}_j(s_j) + \varepsilon_j^u \tag{2.40c}$$

where "^" signifies the nonstochastic variables or functions observed by the investigator. Substitution of (2.40) into (2.24) and (2.25) yields

$$\bar{q}_j(\mu_j,\ w_j,\ b_j;\ s_j,\ \varepsilon_j) = \frac{\mu_j - \hat{c}_j(w_j) - \varepsilon_j^c}{\hat{\sigma}_{jp}^2 + \varepsilon_j^p} \cdot \frac{1}{\hat{\alpha}_j(s_j) + \varepsilon_j^u} \tag{2.41}$$

$$\bar{v}_j(\mu_j, w_j, b_j;\ s_j,\ \varepsilon_j) = 1 - \exp\{\hat{\alpha}_j(s) b_j + \varepsilon_j^u b_j - \frac{[\mu_j - \hat{c}_j(w_j) - \varepsilon_j^c\ ]^2}{2\hat{\sigma}_{jp} + 2\varepsilon_j^p}\}. \tag{2.42}$$

The model is closed by specifying a joint distribution for $\varepsilon_1^c, \ldots, \varepsilon_N^c, \varepsilon_1^p, \ldots, \varepsilon_N^p, \varepsilon_1^u, \ldots, \varepsilon_N^u$.

We will work through these steps for a simplified version of this model in which the random terms ε_j^c and ε_j^p are omitted, leaving only the random term ε_j^u. That is, we assume that the random supply model arises from unobservable variation in the producer's risk preferences. Dropping the "u" superscript, we rewrite (2.40c) as

$$\alpha_j(s_j; \varepsilon_j) = S_j\beta_j + \varepsilon_j \tag{2.43}$$

where S_j is a row vector of K observed exogenous variables representing attributes of the individual producer or the discrete alternative which influence his degree of risk aversion, and β_j is the associated (K×1) vector of coefficients to be estimated — for the sake of generality we allow both S and β to vary with the discrete choice, j. We assume that $\varepsilon_1 \varepsilon_2$ have a bivariate normal distribution with mean zero and some covariance matrix Σ,

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ & \sigma_2^2 \end{bmatrix} \tag{2.44}$$

Our formulation thus allows for the possibility that $\mathrm{cov}(\varepsilon_1, \varepsilon_2) \neq 0$. The correlation of the random terms across the discrete choices could be generated by assuming that the coefficients $\beta_{j1}, \ldots, \beta_{jK}$ are themselves random, in the manner of Hausman and Wise [7]. More generally, it could arise because the same extraneous unobserved factors influence the producer's risk aversion in a similar way across different discrete choices. Finally, to simplify the model further, we assume that the investigator observes the unit costs, c_j, and so does not have to estimate the cost functions, $c_j(w_j)$. Thus, for each given producer, the observed variables are c_1, c_2, μ_1, μ_2, σ_{1p}^2, σ_{2p}^2, b_1, b_2, S_1, and S_2, as well as the producer's actual supply decision — both his discrete choice and his continuous choice. The unknowns are β_1, β_2, and the elements of Σ.

Accordingly, for a given producer the model can be written as:

$$q_1 = \left(\frac{\mu_1 - c_1}{\sigma_{1p}^2}\right) \frac{1}{S_1\beta_1 + \varepsilon_1} \qquad \text{if } I' \geq 0$$

$$q_2 = \left(\frac{\mu_2 - c_2}{\sigma_{2p}^2}\right) \frac{1}{S_2\beta_2 + \varepsilon_2} \qquad \text{if } I' < 0$$

$$I' = [1 - \exp\{b_1S_1\beta_1 + b_1\varepsilon_1 - \frac{(\mu_1 - c_1)^2}{2\sigma_{1p}^2}\}] - [1 - \exp\{b_2S_2\beta_2 + b_2\varepsilon_2 - \frac{(\mu_2 - c_2)^2}{2\sigma_{2p}^2}\}].$$

Define $Y_j \equiv (\mu_j - c_j)/q_j\sigma_j^2$, $j = 1, 2$. An equivalent way of formulating the model is

(2.45a) $\quad Y_1 = S_1\beta_1 + \varepsilon_1 \quad \text{if } I \geq 0$

(2.45b) $\quad Y_2 = S_2\beta_2 + \varepsilon_2 \quad \text{if } I < 0$

(2.45c) $\quad I = b_2S_2\beta_2 - b_1S_1\beta_1 + \frac{(\mu_1 - c_1)^2}{2\sigma_{1p}^2} - \frac{(\mu_2 - c_2)^2}{2\sigma_{2p}^2} + b_2\varepsilon_2 - b_1\varepsilon_1 .$

Alternatively, define

$$Z_1 = (b_1S_1, 0) \qquad Z_2 = (0, b_2S_2)$$

$$X_1 = (S_1, 0) \qquad X_2 = (0, S_2)$$

$$Z_1^+ = -\left[\frac{(\mu_1 - c_1)^2}{2\sigma_{1p}^2}\right] \quad Z_2^+ = -\left[\frac{(\mu_2 - c_2)^2}{2\sigma_{1p}^2}\right]$$

$$\beta = \binom{\beta_1}{\beta_2} .$$

Then, the model (2.45) can be written in its most general form as

(2.46a) $\quad Y_1 = X_1\beta + \varepsilon_1 \quad \text{if } I \geq 0$

(2.46b) $\quad Y_2 = X_2\beta + \varepsilon_2 \quad \text{if } I < 0$

(2.46c) $\quad W_j = Z_j\beta + Z_j^+\gamma + b_j\varepsilon_j, \quad j = 1, 2$

(2.46d) $\quad I = W_2 - W_1$

where $\gamma \equiv 1$. The estimation of switching regression models with the structure of (2.46) is discussed in the next section.

3. Estimation

In this section we discuss the estimation of the following statistical model:

(3.1) $Y_{jt} = X_{jt}\beta + X^{+}_{jt}\theta_j + \varepsilon_{jt}$ $j = 1, \ldots, N$

(3.2) $W_{jt} = Z_{jt}\beta + Z^{+}_{jt}\gamma_j + \eta_{jt}$ $j = 1, \ldots, N$

(3.3) $I_{(j\ell)t} = W_{jt} - W_{1t}$ $1, j = 1, \ldots, N, \quad 1 \neq j$

(3.4) Y_{jt} is observed if $I_{(j\ell)t} > 0$ for all $1 \neq j$, $j = 1, \ldots, N$,

where, for each observation $t = 1, \ldots, T$, X_{jt} and Z_{jt} are each K-dimensional vectors of known constants, X^{+}_{jt} is a K^{+}-dimensional vector of known constants, Z^{+}_{jt} is a K^{++}-dimensional vector of known constants, β, θ_j and γ_j are, respectively, K-, K^{+}-, and K^{++}-dimensional vectors of unknown parameters, and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})$ and $\eta_t = (\eta_{1t}, \ldots, \eta_{Nt})$ are N-dimensional multivariate normal random variables with mean zero and covariance matrices $\Sigma = (\sigma_{ij})$ and $\Xi = (\xi_{ij})$, respectively, which are nonsingular and independent of t.[6] For each t we observe the exogenous variables X_{jt}, X^{+}_{jt}, Z_{jt}, and Z^{++}_{jt}, $j = 1, \ldots, N$. We do not observe W_{jt} or $I_{(j)t}$, but we do observe the indicator variables

(3.5) $$d_{jt} = \begin{cases} 1 \text{ if } I_{(j\ell)t} > 0 \text{ all } j \neq \ell \\ 0 \text{ otherwise} \end{cases} \qquad j = 1, \ldots, N$$

and, if $d_{jt} = 1$ we also observe Y_{jt}. Finally, the random vectors ε_t and η_t are related to one another in some known way. An example is where

(3.6) $\eta_{jt} = a_{jt}\varepsilon_{jt} + \nu_{jt}$ $j = 1, \ldots, N$

where the a_{jt}'s are known constants and the ν_{jt}'s are i.i.d. $N(0, \sigma^2_\nu)$, but other relations are possible. This type of statistical model arises whenever one has a theoretical microeconomic model of discrete/continuous supply or demand decisions, where the discrete and continuous choices both flow from the same underlying random profit or random utility maximization process.

A producer supply example is given in the previous section; a consumer demand example is given in Hanemann [5].

In the terminology of Amemiya [2] and Lee et al. [11], the model (3.1)-(3.5) is a generalized probit regression model. It represents a generalization to N-dimensions of the binary switching regression models presented in Lee and Trost [12], Lee [10], and Heckman [8,9].[7] However, (3.1)-(3.5) is not the most general possible formulation of an N-fold switching regression model. It has two distinctive features: (i) some of the coefficients which appear in the continuous choice equations (3.1) also appear in the discrete choices, (3.3)-(3.4), namely β; and (ii) the disturbance terms in the discrete choice equations are a known function of those in the continuous equations. Both of these features result from the underlying theoretical model of simultaneous optimization of discrete and continuous choices by an economic agent. This special structure is exploited in the estimation procedure described below. In particular, it enables us to identify and estimate the off-diagonal elements in Σ, which is not generally possible when the disturbance terms in (3.1) and (3.2) are unrelated — see, for example, [12, p. 365]. Duncan [3] presents an N-fold discrete/continuous supply model with a structure similar to that of (3.1)-(3.5). However, for the purposes of estimation he assumes that the disturbances in the discrete choice equations are independent of those in the continuous choice equations. As will be shown below, practical procedures exist for estimating the model consistently without invoking this assumption.

Before continuing, it is convenient to introduce some new notation. For each j, define the (N-1) dimensional vector $\upsilon_{(j)t} = (\upsilon_{(j1)t}, \ldots, \upsilon_{(j,\, j-1)t}, \upsilon_{(j,\, j+1)t}, \ldots, \upsilon_{(jN)t})$, where $\upsilon_{(j\ell)t} \equiv \eta_{\ell t} - \eta_{jt}$. In matrix notation we can write

(3.7) $$\upsilon_{(j)t} = D_j \eta_{jt}$$

where D_j is an (N-1) by N matrix with +1 and -1 in the ℓ^{th} and j^{th} cells of each row, and zeros elsewhere.[8] Under the above assumptions, $\upsilon_{(j)}$ is an (N-1) dimensional multivariate normal random vector with mean zero and covariance matrix $\Omega_j \equiv (\omega_{j,\ \ell k}) = D_j \Xi D_j'$. The diagonal elements of this matrix are $\omega^2_{j,\ \ell} = \text{var}(\upsilon_{(j\ell)})$, and the off-diagonals are $\omega_{j,\ell k} = \text{cov}(\upsilon_{(j\ell)}, \upsilon_{(jk)})$. Note that there are some restrictions on the covariance matrices $\Omega_1, \ldots, \Omega_N$ because by definition, $\upsilon_{(j\ell)} = -\upsilon_{(\ell j)}$. In particular, for all ℓ and j, $\omega^2_{j,\ \ell} = \omega^2_{\ell,\ j}$. Hence, although there are N covariance matrices Ω_j and each has (N-1) diagonal elements, there are only N(N-1)/2 distinct $\omega^2_{j,\ \ell}$ terms.

From (3.2)-(3.5) the condition for Y_{jt} to be observed is that

(3.8) $$\upsilon_{(j\ell)t} \le (Z_{jt} - Z_{\ell t})\beta + Z^+_{jt}\gamma_j - Z^+_{\ell t}\gamma_\ell \equiv \Delta_{(j\ell)t}(\beta, \gamma_j, \gamma_\ell), \quad \text{all } \ell \ne j.$$

Alternatively, define the (N-1) dimensional vector of normalized differences $\upsilon^*_{(j)t} = (\upsilon^*_{(j1)t}, \ldots, \upsilon^*_{(j,\ j-1)t}, \upsilon^*_{(j,\ j+1)t}, \ldots, \upsilon^*_{(jN)t})$, where

(3.9) $$\upsilon^*_{(j\ell)} = \upsilon_{(j,\ \ell)} / \sqrt{\omega^2_{j,\ \ell}}.$$

Then, $\upsilon^*_{(j)}$ is multivariate normal with mean zero and covariance matrix $\Omega^*_j \equiv (\omega^*_{j,\ \ell k})$, where the diagonal terms are all unity and the off-diagonals are

(3.10) $$\omega^*_{j,\ \ell k} \equiv \text{cov}\,(\upsilon^*_{(j\ell)}, \upsilon^*_{(jk)}) = \omega_{j,\ \ell k} / \sqrt{\omega^2_{j,\ell}\,\omega^2_{j,\ k}}.$$

In terms of these normalized differences, the condition for Y_{jt} to be observed is that

(3.11) $$\upsilon^*_{(j\ell)t} \le (Z_{jt} - Z_{\ell t}) \frac{\beta}{\sqrt{\omega^2_{j,\ \ell}}} + Z^+_{jt} \frac{\gamma_j}{\sqrt{\omega^2_{j,\ \ell}}} - Z^+_{\ell t} \frac{\gamma_\ell}{\sqrt{\omega^2_{j,\ \ell}}} \equiv \Delta^*_{(j\ell)t}(\beta, \gamma_j, \gamma_\ell)$$

all $\ell \ne j$.

For future reference, the density of $\upsilon^*_{(j)}$ will be denoted $\phi_{N-1}(\upsilon^*_{(j)};\ 0,\ \Omega^*_j)$. For each j, the joint density of ε_j and $\upsilon^*_{(j)}$ is N-dimensional multivariate normal with mean zero and covariance matrix

$$(3.12)\qquad E\left\{\begin{pmatrix}\varepsilon_j \\ \upsilon^*_{(j)}\end{pmatrix}(\varepsilon_j,\ \upsilon^*_{(j)})\right\} = \begin{bmatrix}\sigma_j^2 & \tau_j' \\ \tau_j & \Omega_j^*\end{bmatrix}$$

where $\tau_j' = (\tau_{(j1)}, \ldots, \tau_{(j,\ j-1)}, \tau_{(j,\ j+1)}, \ldots, \tau_{(jN)})$ and $\tau_{j\ell} = \text{cov}(\varepsilon_j, \upsilon^*_{(j\ell)}) = \text{cov}(\varepsilon_j, \upsilon_{(j\ell)})/\sqrt{\omega^2_{j,\ell}}$. It follows that the conditional density of ε_j given $\upsilon^*_{(j)}$ is univariate normal with mean $\tau_j'\ \Omega_j^{*-1}\upsilon^*_{(j)}$ and variance $(\sigma_j^2 - \tau_j'\ \Omega_j^{*-1}\tau_j)$, which we denote $\phi(\varepsilon_j;\ \tau_j'\Omega_j^{*-1}\upsilon^*_{(j)},\ \sigma_j^2 - \tau_j'\Omega_j^{*-1}\tau_j)$.

Given a sample of T observations, let j* be the index of the discrete choice mode in the t^{th} observation (this will vary with t), so that $d_{j^*t} = 1$ and $d_{\ell t} = 0$ all $\ell \neq j^*$, and let the observed continuous choice be Y_{j^*t}. The likelihood function for this sample is

$$(3.13)\qquad L = \prod_{t=1}^{T} L_t$$

where

$$(3.14)\qquad L_t = \Pr\{\varepsilon_{j^*t} = Y_{j^*t} - X_{j^*t}\beta - X^+_{j^*t}\theta_{j^*} \cap \upsilon^*_{(j^*\ell)t} \leq \Delta^*_{(j^*\ell)t} \quad \text{all } \ell \neq j^*\}$$

$$= \Pr\{\varepsilon_{j^*t} = Y_{j^*t} - X_{j^*t}\beta - X^+_{j^*t}\theta_{j^*} | \upsilon^*_{(j^*)t} \leq \Delta^*_{(j^*)t}\} \cdot \Pr\{\upsilon^*_{(j^*\ell)t_j} \leq \Delta^*_{(j^*\ell)t},\ \text{all } \ell \neq j^*\}$$

$$= \int_{-\infty}^{\Delta_{(j^*1)t}} \cdots \int_{-\infty}^{\Delta^*_{(j^*,j^*-1)t}} \int_{-\infty}^{\Delta^*_{(j^*,j^*+1)t}} \cdots \int_{-\infty}^{\Delta^*_{(j^*N)t}} \phi(Y_{j^*t} - X_{j^*t}\beta - X^+_{j^*t}\theta_{j^*};\ \tau_j'\Omega_{j^*}^{*-1}\upsilon^*_{(j^*)},\ \sigma^2_{j^*} - \tau'_{j^*}\Omega_{j^*}^{*-1}\tau_{j^*}$$

$$* \ \phi_{N-1}(\upsilon^*_{(j^*)};\ 0;\ \Omega^*_{j^*})\ d\upsilon^*_{(j^*1)}, \ldots, d\upsilon^*_{(j^*,\ j^*-1)}, d\upsilon^*_{(j^*,\ j^*+1)}, \ldots, d\upsilon^*_{(j^*N)}\ .$$

In principle, all of the unknowns in the model — β, $\theta_1, \ldots, \theta_N$, $\gamma_1, \ldots, \gamma_N$, Σ and Ξ — could be estimated by full information maximum likelihood based on (3.13)-(3.14). Following the argument of Amemiya [1], it can be shown that the MLE is consistent and asymptotically normal and efficient. In practice, however, this will be computationally burdensome unless N is smaller than 5 or 6. Moreover, the normal equations will generally have multiple roots because of the nonlinearity of (3.14) and, unless one starts from an initial consistent estimator, there is no guarantee of convergence to the global MLE.

Accordingly, we will employ a multi-stage estimation procedure originally developed by Amemiya [1], Heckman [9] and Lee [10], modified to allow for the special structure of (3.1)-(3.5). The first step involves a probit model applied to the purely discrete choice represented by (3.8) or (3.11). For each observation t, the probability that Y_{j*t} is observed is

$$(3.15) \quad P_t^{j*} = \int_{-\infty}^{\Delta^*_{(j*1)t}} \cdots \int_{-\infty}^{\Delta^*_{(j*,j*-1)t}} \int_{-\infty}^{\Delta^*_{(j*,j*+1)t}} \cdots \int_{-\infty}^{\Delta^*_{(j*N)t}} \phi_{N-1}(\upsilon^*_{(j*)}; 0, \Omega^*_{j*}) d\upsilon^*_{(j*1)}, \ldots, d\upsilon^*_{(j*,j*-1)}, d\upsilon^*_{(j*,j*+1)}, \ldots, d\upsilon^*_{(j*N)}$$

The likelihood function for the probit model is

$$(3.16) \qquad L' = \prod_{t=1}^{T} P_t^{j*}.$$

Using maximum likelihood or, if there are grouped data, weighted least squares, one can obtain estimates of the parameters in (3.15) which are consistent but not efficient, since they ignore the information contained in the data on the continuous choices.

It is important to note that, because the conditions (3.8) and (3.10) are observationally indistinguishable, we do not obtain estimates of β and $\gamma_1, \ldots, \gamma_N$ from the probit model. Rather we obtain estimates of the N(N-1)/2 terms $\beta/\sqrt{\omega^2_{j,\ell}}$, the N(N-1)/2 terms $\gamma_1/\sqrt{\omega^2_{j,\ell}}$, the N(N-1)/2 terms $\gamma_2/\sqrt{\omega^2_{j,\ell}}$, ..., the N(N-1)/2 terms $\gamma_N/\sqrt{\omega^2_{j,\ell}}$. Similarly, we obtain estimates of the off-diagonal terms in $\Omega^*_1, \ldots, \Omega^*_N$, not $\Omega_1, \ldots, \Omega_N$. However, as Duncan points out, if there is at least one non-homogeneous exact restriction on the elements of β or γ_j that is available exogenously, it can be employed to identify the N(N-1)/2 terms $\sqrt{\omega^2_{j,\ell}}$ and, hence, to identify β and $\gamma_1, \ldots, \gamma_N$. Failing this, we would have to impose the normalization

$$(3.16) \qquad \omega^2_{j,\ell} \equiv \mathrm{var}(\upsilon_{(j,\ell)}) = \xi^2_\ell + \xi^2_j - 2\xi_{j\ell} = 1 \quad \text{all } j \neq \ell$$

in order to identify β, the γ_j's and $\Omega_1, \ldots, \Omega_N$ from the probit analysis of the purely discrete choice data. This would be unattractive because it entails N(N-1)/2 restrictions on N(N+1)/2 free elements of the covariance matrix Ξ. Moreover, if Ξ is a function of Σ, as is implied by most theoretical random supply models, it entails restrictions on Σ, which further reduce the generality of the model. Fortunately, however, the normalization (3.16) can be avoided because we also have information from the continuous choices, Y_1, ..., Y_N, which serve to identify the $\omega^2_{j,\ell}$'s. This occurs in the second step of the estimation procedure.

Recall that for each t we observe only <u>one</u> of the Y_j's, Y_{j*t}; the one observed is determined by the conditions (3.8) or (3.10). Therefore, following the argument of Amemiya [2] and Heckman, the first two moments of the observed continuous choice variable are

$$(3.17)\qquad E\{Y_{j*t}|Y_{j*t}\text{ observed}\} = X_{j*t}\beta + X^{+}_{j*t}\theta_j + E\{\varepsilon_{j*t}|\upsilon^*_{(j*\ell)t} \leq \Delta^*_{(j*\ell)t},\ \text{all } \ell \neq j*\}$$

$$= X_{j*t}\beta + X^{+}_{j*t}\theta_j - \sum_{\ell\neq j*}\tau_{(j*\ell)}\lambda_{(j*\ell)t}$$

$$(3.18)\qquad V\{Y_{j*t}|Y_{j*t}\text{ observed}\} = V\{\varepsilon_{j*t}|\upsilon^*_{(j*\ell)t} \leq \Delta^*_{(j*\ell)t},\ \text{all } \ell \neq j*\}$$

$$= \sigma^2_{j*} + \tau'_{(j*)}\Lambda_{(j*)}\tau_{(j*)}$$

where

$$(3.19a)\qquad \lambda_{(j*\ell)t} = [\partial\Phi_{N-1}(\Delta^*_{(j*)t};\ 0, \Omega^*_j)/\partial\Delta^*_{(j*\ell)t}]/\ \Phi_{N-1}(\Delta^*_{(j*)t};\ 0, \Omega^*_j),$$

$$(3.19b)\qquad \Lambda_{(j*)} = \frac{\partial^2\Phi_{N-1}(\Delta^*_{(j*)};\ 0, \Omega^*_j)/\partial\Delta^*_{(j*)}\partial\Delta^{*\prime}_{(j*)}}{\Phi(\Delta^*_{(j*)};\ 0, \Omega^*_j)} - \lambda_{(j*)}\lambda'_{(j*)},$$

$\Phi_{N-1}(\cdot;\ 0,\ \Omega^*_j)$ being the c.d.f. associated with $\phi_{N-1}(\cdot;\ 0,\ \Omega^*_j)$. Hence, the appropriate regression model for the observed continuous choices is

$$(3.20)\qquad Y_{j*t} = X_{j*t}\beta + X^{+}_{j*t}\theta_j - \sum_{\ell\neq j}\tau_{(j*\ell)}\lambda_{(j*\ell)t} + \xi_{j*t}\qquad t = 1, \ldots, T$$

where $\xi_{j*t} = Y_{j*t} - E\{Y_{j*t}|Y_{j*t}\text{ observed}\}$ is a normal disturbance, independent of the regressors in (3.20), with a zero mean and a variance given by the right-hand side of (3.18). Therefore, following Heckman [9], one could use the fitted probit model to form consistent estimates of $\lambda_{(j*\ell)t}$, insert these as regressors in (3.20) and fit it by least squares — by OLS or, taking account of the heteroscedasticity implied by (3.18), by GLS. This yields consistent estimates of β, $\theta_1, \ldots, \theta_N$, and $\tau_{(11)}, \ldots, \tau_{(NN)}$. However, because the regression involves estimated regressors, the usual formulas for the covariance matrix of these coefficient estimates would need to be modified.

The disadvantage of this approach in the present context is that, when one compares the estimates of β obtained from the estimation of (3.20) with the estimates of $\beta/\sqrt{\omega^2_{j,\ell}}$ obtained from the probit model, there will be K estimates of each of the N(N-1)/2 terms $\omega^2_{j,\ell}$.[9] An alternative procedure which combines the information from the discrete choices with that from the continuous choices and yields unique estimates of β and the $\omega^2_{j,\ell}$'s is the following. For each observation t, write the regression model (3.20) as N(N-1)/2 equations (for simplicity, we illustrate this for the case where N = 3):

$$Y_{j*t} = (X_{j*t}\frac{\beta}{\sqrt{\omega^2_{1,2}}})\sqrt{\omega^2_{1,2}} + X^+_{j*t}\theta_{j*} - \sum_{\ell\neq j*}\tau_{(j*\ell)}\lambda_{(j*\ell)t} + \xi_{j*t}$$

$$\text{(3.21)} \qquad Y_{j*t} = (X_{j*t}\frac{\beta}{\sqrt{\omega^2_{1,3}}})\sqrt{\omega^2_{1,3}} + X^+_{j*t}\theta_{j*} - \sum_{\ell\neq j*}\tau_{(j*\ell)}\lambda_{(j*\ell)t} + \xi_{j*t}$$

$$Y_{j*t} = (X_{j*t}\frac{\beta}{\sqrt{\omega^2_{2,3}}})\sqrt{\omega^2_{2,3}} + X^+_{j*t}\theta_{j*} - \sum_{\ell\neq j*}\tau_{(j*\ell)}\lambda_{(j*\ell)t} + \xi_{j*t}\,.$$

Use the probit estimates of $\beta/\sqrt{\omega^2_{j,\ell}}$ to form the regressors $(X_{j*t}\beta/\sqrt{\omega^2_{j,\ell}})$ and run the regression (3.22) with TN(N-1)/2 "observations," treating the $\sqrt{\omega^2_{j,\ell}}$ terms as the coefficients to be estimated, together with $\theta_1, \ldots, \theta_N$ and $\tau_{(11)}, \ldots, \tau_{(NN)}$. One can then multiply the N(N-1)/2 coefficient estimates from the probit model by the consistent estimates $\widehat{\sqrt{\omega^2_{j,\ell}}}$ from the regression model (3.21) to obtain unique and consistent estimates of β and $\gamma_1, \ldots, \gamma_N$. Similarly, one can multiply the estimates of the off-diagonal terms in Ω^*_j, ..., Ω^*_N obtained from the probit model by the estimates $\widehat{\sqrt{\omega^2_{j,\ell}}}$ to obtain consistent estimates of the off-diagonal terms in $\Omega_1, \ldots, \Omega_N$. Since $\Omega_j = D_j \Xi D'_j$, where D_j is a known matrix, one can then obtain consistent estimates of the covariance matrix Ξ. Further, by applying the relations in (3.18) to the second moments of the estimated residuals from (3.21) one can obtain consistent estimates of $\sigma_1^2, \ldots, \sigma_N^2$. Consistent estimates of the off-diagonal terms in Σ can be obtained using the estimates of Ξ and τ_j, as well as the relations (3.6).

It is worth emphasizing why we care about the uniqueness of the estimates of β and $\Omega_1, \ldots, \Omega_N$. The standard procedure for estimating switching regression models is to obtain a set of consistent estimates for all the unknowns and to employ these as initial values for an iterative maximization of the full likelihood function for the discrete and continuous choices combined. If one were following this procedure, it would be unimportant whether or not a unique set of coefficient estimates could be obtained from the probit and regression models. However, when N is at all large, the likelihood function (3.13)-(3.14) is computationally intractable. This arises not just because an (N-1) integral is involved — after all, the same is true of the likelihood function for the probit model (3.15) — but also because the integrand in (3.14) has a particularly complex structure, considerably more complex than that of the integrand in (3.15). Therefore, following Duncan, we assume that even a single Newton Raphson iteration of the normal equations for the full model is impractical and the only coefficient estimates available are those from the probit and regression models.

If $\text{cov}(\varepsilon_j, \upsilon^*_{(j)}) \equiv \tau_j = 0$, then all the terms in $\lambda_{(j^*, \ell)}$ drop out from (3.20) or (3.21), and the coefficient estimates obtained from these regression models are MLE's and, hence, fully efficient. If $\tau_j \neq 0$, as seems more plausible, the estimates are consistent but they are not fully efficient. In this context it may be worth treating the estimates of $\beta/\sqrt{\omega^2_{j,\ell}}$ and $\sqrt{\omega^2_{j,\ell}}$ obtained in the manner described above as stochastic prior information and applying the Theil-Goldberger [15] mixed regression procedure, as an additional step after the estimation of (3.21). For simplicity, we illustrate this for the case where $N = 3$ and, say $K = 3$; the additional regression model is

$$Y_{j*1} = X_{j*1}\beta + X^{+}_{j*1}\theta_{j*} - \sum_{\ell \neq j*} \tau_{(j*\ell)} \hat{\lambda}_{(j*\ell)} + \zeta_{j*1}$$

$$\vdots$$

$$Y_{j*T} = X_{j*T}\beta + X^{+}_{j*T}\theta_{j*} - \sum_{\ell \neq j*} \tau_{(j*\ell)} \hat{\lambda}_{(j*\ell)T} + \zeta_{j*T}$$

$$\sqrt{\hat{\omega}^2_{1,2}}\left(\frac{\hat{\beta}_1}{\sqrt{\omega^2_{1,2}}}\right) = 1\beta_1 \qquad + \psi_1$$

(3.22)
$$\sqrt{\hat{\omega}^2_{1,2}}\left(\frac{\hat{\beta}_2}{\sqrt{\omega^2_{1,2}}}\right) = 1\beta_2 \qquad + \psi_2$$

$$\sqrt{\hat{\omega}^2_{1,2}}\left(\frac{\hat{\beta}_3}{\sqrt{\omega^2_{1,2}}}\right) = 1\beta_3 \qquad + \psi_3$$

$$\sqrt{\hat{\omega}^2_{1,2}}\left(\frac{\hat{\beta}_2}{\sqrt{\omega^2_{1,3}}}\right) = 1\beta_2 \qquad + \psi_4$$

$$\sqrt{\hat{\omega}^2_{2,3}}\left(\frac{\hat{\beta}_3}{\sqrt{\omega^2_{2,3}}}\right) = 1\beta_3 \qquad + \psi_5$$

where β_k is the k^{th} element of β. In the general case, the regression model (3.22) involves $T + [\frac{N(N-1)}{2} + K - 1]$ "observations." The disturbance terms associated with the last $[\frac{N(N-1)}{2} + K - 1]$ observations have the form:

(3.23)
$$\psi_1 = \sqrt{\hat{\omega}^2_{1,2}}\,(\hat{\beta_1}/\sqrt{\omega^2_{1,2}}) - \beta_1$$

$$= \sqrt{\omega^2_{1,2}}\,[(\hat{\beta_1}/\sqrt{\omega^2_{1,2}}) - (\beta_1/\sqrt{\omega^2_{1,2}})] + (\hat{\beta_1}/\sqrt{\omega^2_{1,2}})[\sqrt{\hat{\omega}^2_{1,2}} - \sqrt{\omega^2_{1,2}}] .$$

Hence, plim $(\psi_1) = 0.$[10] Application of OLS — or alternatively, GLS — to (3.22) will yield consistent estimates of β which fully exploit the information contained in the constraint that these coefficients are common to both the discrete and the continuous choices.[11]

4. Conclusions

In section 2 of this paper we have shown how to construct theoretical models of discrete and continuous supply decisions, where both sets of decisions flow from a single underlying utility-of-profit maximization problem. In addition to describing the general procedure for creating such models, we have developed a specific model based on the exponential utility-of-profit function and the normal distribution for output prices which is suitable for empirical application. One area for future research is the development of other models based on alternative utility-of-profit functions and/or output price probability distributions. The key issue here, which is still unresolved, is whether the duality relationship between the indirect expected utility-of-profit function and the output supply function, (2.4), can be effectively exploited to generate a variety of parametric supply models.

In order to create a statistical framework for estimating the discrete/continuous supply models, it is necessary to postulate that some component of the supplier's utility or production function is random from the viewpoint of the econometric investigator. By introducing this random term in different ways, or by making different assumptions about its probability distribution, one can generate different discrete/continuous supply models. In this paper we have assumed that the random element is normally distributed, which leads to a probit model of the discrete choices. We could alternatively have employed the extreme value distribution, which would have yielded a logit model of the discrete choices. Moreover, rather than introducing the random element in an arbitrary manner, we have sought to give it an economic interpretation by identifying it with a specific parameter of the model — the producer's coefficient of risk aversion, in our specific model. Other ways of defining an economically meaningful random element deserve to be explored, and the same model-building philosophy can be applied to the profit-maximization models of Duncan [3] and McFadden [13].

As for the estimation of discrete/continuous choice models, we have developed a very general statistical model which applies not only to the specific supply models developed in section 2 but also to some of the demand and supply models of Hanemann [6], Duncan, and McFadden. We have shown that these discrete/continuous choice models can be regarded as instances of a multivariate switching regression model with an N-fold switching, which generalizes the binary switching model that has appeared in the literature. Moreover, because the discrete and continuous choices both result from the same underlying optimization decision by an economic agent, there are additional restrictions on the coefficients and disturbance terms of the equations for the discrete and the continuous choices. The main focus of our discussion in section 3 has been how to exploit these common restrictions in an efficient estimation procedure.

Duncan, who discusses this issue, assumes that only the coefficients are common to the discrete and continuous choice equations while the disturbance terms are uncorrelated. However, we find this assumption unsatisfactory — given that both sets of choices result from a single optimization decision, it seems more plausible to assume that the random terms which influence the agent's discrete choice are related to those which influence his continuous choice. As we have shown, practical procedures exist for consistently estimating the discrete and continuous choice equations without invoking Duncan's assumption. In order to obtain efficient estimates, it is important to incorporate the results of fitting the discrete choice equation in the estimation of the continuous choice equations. Our innovation here is to treat this as a problem of mixed estimation with endogenous stochastic information. Most previous discussions of the value of information in regression analysis, such as [15], have only considered the case of exogenous prior information. However, it can be shown that there is still an efficiency gain with endogenous information. The magnitude of this gain remains to be tested in an empirical application, which will be reported separately.

FOOTNOTES

1. This assumption of output price uncertainty can also be extended to include the notion of yield uncertainty: interpret q as the ex-ante anticipated output and p as the "effective price" — i.e., actual price times the ratio of actual to anticipated output. It is necessary under this interpretation to assume that variable production costs, $c(\cdot)$, depend on planned output rather than actual output, which is not unreasonable. It is not possible, however, to include the notion of input price uncertainty in this formulation.

2. The dual approach to the generation of continuous supply models under uncertainty is investigated by Hallam, Just and Pope [4] , who describe the requirements for the functions $v(\cdot)$ and $q(\cdot)$.

3. In all these examples we can actually assume intra-agent as well as inter-agent variability — i.e., although an individual's technology, information and preferences are fixed at the point of each decision, they may vary between decisions in a manner which is partly unobservable to the investigator and is taken by the investigator to be random.

4. We assume that $f_{\bar{v}}(\cdot)$ and the other derived probability distributions described below exist and are well defined. Note that, if the random terms c_j enter the conditional indirect expected utility-of-profit functions in an appropriate manner, they might disappear from the ratio on the right-hand side of (2.29). In that case the conditional supply, $\bar{q}_j$, would not be a random variable for the investigator; in effect, the existence of an unobsrrvable component of the producer's decision affects his discrete choice but not his continuous choice.

5. The following development involving (2.37)-(2.39) can readily be extended to the case of $N > 2$. This will be discussed in section 3.

6. Actually we could allow Σ to vary with t as long as we imposed some further structure on $\Sigma_1, \ldots, \Sigma_N$, for example by adopting the random coefficient specification of Hausman and Wise [7]. Similarly, we could allow Ξ to vary with t.

7. Amemiya [2], Lee [10] and Heckman [8] each offer a version of a multivariate switching model, but their models have a different structure from ours: their models involve essentially a binary discrete choice, whereas our formulation involves an N-fold discrete choice.

8. At this point we will suppress the observation index, t, unless this causes an ambiguity.

9. Duncan, who assumes the presence of exogenous information which uniquely identifies β and the $\omega^2_{j,\ell}$'s from the probit estimates, faces a different over-identification problem. When one fits the regression model (3.20), there are then two consistent estimates of β, one from the probit model and the other from the regression. Denote these two estimates by $\hat{\beta}_1$ and $\hat{\beta}_2$, and their respective estimated covariance matrices by $\hat{V}_1$ and $\hat{V}_2$. Duncan proposes to resolve this over-identification by taking a weighted average of the two estimators, which itself is consistent: $\bar{\beta} = (\hat{V}_1^{-1} + V_2^{-1})^{-1}$ $(\hat{V}_1^{-1}\hat{\beta}_1 + \hat{V}_2^{-1}\hat{\beta}_2)$.

10. The reason why we add $[\frac{N(N-1)}{2} + K - 1]$ observations in (3.22) instead of TK observations is to avoid a dependence relation among the disturbance terms associated with the additional observations, ψ_1, ψ_2, etc. Suppose, for example, that we added the following row to (3.22):

$$\sqrt{\hat{\omega}^2_{1,3}}\left(\frac{\hat{\beta}_1}{\sqrt{\omega^2_{1,3}}}\right) = 1\beta_1 + \psi_6.$$

(10.) Then, since

$$\frac{\sqrt{\hat{\omega}^2_{1,2}}\,(\widehat{\beta_1/\sqrt{\omega^2_{1,2}}})}{\sqrt{\hat{\omega}^2_{1,3}}\,(\widehat{\beta_1/\sqrt{\omega^2_{1,3}}})} = \frac{\sqrt{\hat{\omega}^2_{1,2}}\,(\widehat{\beta_2/\sqrt{\omega^2_{1,2}}})}{\sqrt{\hat{\omega}^2_{1,3}}\,(\widehat{\beta_2/\sqrt{\omega^2_{1,3}}})},$$

the disturbances would satisfy the identity $\psi_1\psi_4 \equiv \psi_2\psi_6$.

11. Strictly speaking, the regression model (3.22) is not the same as Theil-Goldberger's mixed regression because it involves endogenous stochastic information, in the sense that $\text{cov}(\xi_{j*t}, \psi_\ell) \neq 0$. Nevertheless, using the results in Rothenberg [14, p. 47-8], it can be shown that the estimation of (3.22) must bring some gain in efficiency, at least in finite samples. The magnitude of this gain is currently being investigated. A further possibility is to estimate (3.22) iteratively: after fitting (3.22), compare the estimate of β with the estimate of $(\widehat{\beta/\sqrt{\omega^2_{j,\ell}}})$ from the probit model to obtain a new estimate of $\sqrt{\omega^2_{j,\ell}}$ and then refit (3.22), until the estimate of β converges.

REFERENCES

1. Amemiya, Takeshi, "Regression Analysis When the Dependent Variable is Truncated Normal," *Econometrica*, 41 (1973), 997-1016.

2. ________, "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," *Econometrica*, 42 (1974), 999-1012.

3. Duncan, Gregory M. "Formulation and Statistical Analysis of the Mixed, Continuous/Discrete Dependent Variable Model in Classical Production Theory," *Econometrica*, 48 (1980), 839-852.

4. Hallam, J. Arne, Richard E. Just, and Rulon D. Pope, "Positive Economic Analysis and Risk Considerations in Agricultural Production," in *New Directions in Econometric Modeling and Forecasting in U.S. Agriculture*, ed. by Gordon C. Rausser. New York: Elsevier North-Holland, 1982.

5. Hanemann, W. Michael, "Discrete/Continuous Models of Consumer Choice," Berkeley: University of California Department of Agricultural and Resource Economics, February 1982.

6. ________, "Quality and Demand Analysis," in *New Directions in Econometric Modeling and Forecasting in U.S. Agriculture*, ed. by Gordon C. Rausser. New York: Elsevier North-Holland, 1982.

7. Hausman, Jerry A., and David A. Wise, "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46 (1978), 403-426.

8. Heckman, James J., "Dummy Exogenous Variables in a Simultaneous Equation System," *Econometrica*, 46 (1978), 931-959.

9. ________, "Sample Selection Bias as a Specification Error," *Econometrica*, 47 (1979), 153-161.

10. Lee, Lung-Fei, "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables," *Econometrica*, 47 (1979), 977-996.

11. Lee, Lung-Fei, G.S. Maddala, and R.P. Trost, "Asymptotic Covariance Matrices of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equation Models with Selectivity, *Econometrica*, 48 (1980), 491-503.

12. Lee, Lung-Fei, and R.P. Trost, "Estimation of Some Limited Dependent Variable Models With Application to Housing Demand," *Journal of Econometrics,* 8 (1978), 357-382.

13. McFadden, D. "Econometric Net Supply Systems for Firms with Continuous and Discrete Commodities," Cambridge: Massachusetts Institute of Technology Department of Economics Working Paper, 1979.

14. Rothenberg, Thomas J. "Efficient Estimation with A Priori Information," New Haven: Yale University, Cowles Foundation Monograph 23, Yale University Press, 1973.

15. Theil, H., and A.S. Goldberger, "On Pure and Mixed Statistical Estimation in Economics," *International Economic Review*, 2 (1961), 65-78.