AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

WP 2002-42
December 2002

Working Paper

Department of Applied Economics and Management
Cornell University, Ithaca, New York 14853-7801 USA

BAYES' ESTIMATES OF THE DOUBLE HURDLE MODEL IN THE PRESENCE OF FIXED COSTS

Garth Holloway, Christopher B. Barrett, and Simeon Ehui

It is the Policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

BAYES' ESTIMATES OF THE DOUBLE HURDLE MODEL IN THE PRESENCE OF FIXED COSTS

Garth Holloway[a,b], Christopher B. Barrett[c] and Simeon Ehui[a]

December 2002 Revised Version

Abstract

We present a model of market participation in which the presence of non-negligible fixed costs leads to non-zero censoring of the traditional double-hurdle regression. Fixed costs arise when household resources must be devoted a priori to the decision to participate in the market. These costs—usually a cost of time—motivate two-step decision-making and focus attention on the minimum-efficient scale of operations (the minimum amount of milk sales) at which market entry becomes viable. This focus, in turn, motivates a non-zero-censored Tobit regression estimated through routine application of Markov chain Monte Carlo Methods. (95 words).

Keywords: market participation, fixed costs, double-hurdle model, censored regression (8 words).

Journal of Economic Literature Classifications: O11, C34, O13.

---

[a]Livestock Policy Analysis Program, International Livestock Research Institute, PO Box 5689 Addis Ababa, Ethiopia, phone: (251)-(1)-46-32-15, fax: (251)-(1)-46-12-52, email: g.holloway@cgiar.org; [b]University of Reading, UK. [c]Cornell University, USA; We are grateful to Charles Nicholson who conceived the data collection while a Post-Doctoral Fellow at the International Livestock Research Institute in 1997 under a grant from the Rockefeller Foundation. We are also grateful to Bill Griffiths for several enlightening conversations concerning the estimation algorithm. Submitted: September 27, 2001.

BAYES' ESTIMATES OF THE DOUBLE HURDLE MODEL IN THE PRESENCE OF FIXED COSTS

Abstract

We present a model of market adoption (participation) where the presence of non-negligible fixed costs leads to non-zero censoring of the traditional double-hurdle regression. Fixed costs arise due to household resources that must be devoted a priori to the decision to participate in the market. These costs—usually a cost of time—motivate two-step decision-making and focus attentions on the minimum-efficient scale of operations (the minimum amount of milk sales) at which market entry becomes viable. This focus, in turn, motivates a non-zero-censored Tobit regression estimated through routine application of Markov chain Monte Carlo Methods. (95 words).

Keywords: market participation, fixed costs, double-hurdle model, censored regression (8 words).

Journal of Economic Literature Classifications: O11, C34, O13.

THE DOUBLE HURDLE MODEL IN THE PRESENCE OF FIXED COSTS

Households commonly incur fixed costs in making the decision to trade in a market. These costs can involve pecuniary expenditures, such as a fixed fee to enter a market in order to sell product. More commonly, the fixed costs of market participation involve time spent in search for and screening of counterpart transactors and in negotiating and enforcing contracts. Such costs are known to exist irrespective of transactions volume and surely affect the logically subsequent decision over how much quantity to supply to the market. Yet the standard estimation of market supply equations fails to account for these fixed costs. In this paper we demonstrate a method for estimating the double hurdle model of market participation and supply volume determination in the face of unobservable fixed costs.

The next section presents a simple model of the adoption decision and uses some familiar specifications to motivate the basic ideas. Section three presents the econometric model. Section four presents the estimation algorithm. Section five presents modifications incurred when fixed-costs are non-negligible. Section six discusses the application and section seven presents results. Conclusions are offered in section eight.

Market Participation As An Adoption Decision

Over the past decade or so, economists have begun to treat market supply decisions as a sequence of two steps, a market participation decision followed by a supply volume decision (Goetz 1992, Key et al. 2000). The notion of two-step decision-making can be motivated in the following way. Let i = 1, 2, .., N denote the households in question. Each household compares the level of utility derived from market participation, y_{pi}^*, against its reservation utility attainable

without market participation, y_{ri}^*. Here, we use the superscipt "*" to denote the fact that both levels of utility are latent (unobservable) random variables. We will follow this convention below.

We assume that the difference between the utility levels is determined by a vector of characteristics specific to each household, $\mathbf{x}_{pi}$. Without loss of generality, we set $y_{ri}^* = 0$ and denote the difference between the incurred and reserve utility levels y_{pi}, and their relationship to the characteristics by the function $f_i(\cdot)$. The condition characterizing the discrete choice about whether to participate in the market can then be written

$$y_{pi} = f_i(\mathbf{x}_{pi}), \tag{1}$$

with participation when $y_{pi} > 0$ and nonparticipation otherwise. We now let the indicator variable $\delta_i = 1$ when $y_{pi} > 0$ and the household participates in the market, with $\delta_i = 0$ under nonparticipation.

Statistical implementation depends on the information structure of this choice problem, in particular whether the discrete participation decision occurs <u>before</u> a corresponding quantity decision is undertaken about the <u>intensity</u> of participation, in this case, as to <u>how much</u> quantity to supply to the market. As is customary, we assume the participation decision is made first and that, conditional on that decision, the household now faces a corresponding quantity decision.

In introducing the multivariate econometric model, below, it will be useful to conserve on notation. Hence, in presenting the sales decision, we continue to use y to reference the endogenous variable of interest, but distinguish the sales quantity from the latent participation variable through subscripts, the former denoted y_{si} and the latter denoted y_{pi}. Let $\Phi_i(\cdot)$ denote the level of a maximand – e.g., profit or utility – defined over the supply quantity, y_{si}, and let $\varphi_i(\cdot)$ denote its first-order partial derivative with respect to this quantity. Naturally, this decision

will also be affected by a set of household characteristics, which may be the same or may differ from the ones affecting the participation action. Let $\mathbf{x}_{si}$ denote these characteristics. Across each of the households i = 1, 2, .. N, we are concerned with the problem:

$$\max_{y_{si}} \quad \Phi_i(y_{si} \mid \mathbf{x}_{si}) \qquad \text{subject to} \qquad y_{si} \geq 0 \tag{2}$$

and the associated first-order conditions for a maximum; namely the derivative condition on the objective function,

$$\varphi_i(y_{si} \mid \mathbf{x}_{si}) \leq 0; \tag{3}$$

the non-negativity restriction on choice,

$$y_{si} \geq 0; \tag{4}$$

and the complementary-slackness condition,

$$\varphi_i(y_{si} \mid \mathbf{x}_{si}) \, y_{si} = 0. \tag{5}$$

Equations (1)-(5) form the basis for a double-hurdle interpretation of the household's supply decision, on which we now expand.

A Standard Double-Hurdle Model Of The Supply Decision

Assume that the households, i = 1, 2, .., N generate a sample (of size N) independent supply decisions. For each household in the sample the decision as to how much quantity to supply is a double-hurdle problem with three components. Observed sales are

$$y_{si} = \delta_i \, y_{si}^{**}, \tag{6}$$

where δ_i is the market participation indicator variable and y_{si}^{**} refers to a potentially censored target sales quantity.. A linear version of the participation equation (equation (1)) has the form

$$y_{pi} = \boldsymbol{\beta}_p \, \mathbf{x}_{pi} + u_{pi}, \tag{7}$$

where $\delta_i = 1$ if $y_{pi} > 0$ and $\delta_i = 0$ otherwise, where $\boldsymbol{\beta}_p$ is a vector of unknown coefficients

controlling the relationship between household-specific characteristics and market participation, and u_{pi} is a random error. Finally, the model is completed by inclusion of a sales equation,

$$y_{si}^{*} = \boldsymbol{\beta}_s \mathbf{x}_{si} + u_{si}, \tag{8}$$

where we observe $y_{si}^{**} = \max\{0, y_{si}^{*}\}$; y_{si}^{*} is the latent (random) optimal sales volume, which is related to the household-specific covariates, $\mathbf{x}_{si}$, by the vector $\boldsymbol{\beta}_s$, with u_{si} a random error.

Equations (6)-(8), along with their restrictions, combine to yield the double-hurdle motivation for participation. This notion is exhibited clearly in equation (6), which states that two conditions must be met in order for positive sales to be observed. First, the indicator variable, δ_i, must be positive. In other words, the condition $y_{pi} > 0$ must prevail in equation (7). Second, the latent quantity y_{si}^{*} must exceed zero in equation (8). Hence, both the participation- and the sales-equations "constraints" must be satisfied in order for positive sales to arise.

Equation (7) is simply a linear, statistical interpretation of the participation decision in equation (1) and, when the error is normal, has the (important) connotation of a probit equation. Equation (8) follows from relaxing the non-negativity constraint in equation (4), ignoring the complementary-slackness condition in equation (5) and acknowledging that, when one does so, a latent, censored (Tobit) regression is implied in which observed sales are left-censored at zero.

Estimation

Because two conditions must be met in order for positive sales to arise, the likelihood of observing a positive observation is simply the conditional data density for that observation multiplied by the joint probability that the two events occur, or

$$\ell(y_{si} > 0) = f(y_{si} \mid \delta_i = 1 \text{ and } y_{si} > 0) \times \text{prob}(\delta_i = 1 \text{ and } y_{si}^{*} > 0). \tag{9}$$

Consequently, the likelihood for observing zero sales is the probability that neither of the two

conditions in question prevail, or

$$\ell(\, y_{si} = 0) \;=\; 1\text{- prob}(\delta_i = 1 \text{ and } y_{si}^{*} > 0). \tag{10}$$

If the errors in the participation and sales equations (u_{pi} and u_{si}, respectively) are independent, then the joint probability of the two events occurring ($\delta_i = 1$ and $y_{si}^{*} > 0$) can be factored into the product of marginal probabilities. Other recent work has used that simplifying restriction (Key et al., 2000). Less restrictively, one can assume that the errors in (7) and (8) follow a multivariate normal distribution. In this context equation (7) depicts a traditional probit regression, equation (8) depicts a traditional Tobit regression, and the multivariate-normal assumption allows correlation between the errors, as in Nelson (1977), Cogan (1981) or Goetz (1992). By combining results in Chib and in Albert and Chib, some algebra (available upon request) reveals that the full conditional distributions for the unknown quantities have simple forms, wherein a Gibbs-sampling, data-augmentation algorithm can be constructed in order to simulate from the joint posterior distribution for the system parameters.

More precisely, stacking (7) and (8) as

$$\mathbf{y} \;=\; \mathbf{x}\,\boldsymbol{\beta} \;+\; \mathbf{u}, \tag{11}$$

where $\mathbf{y} \equiv (\mathbf{y}_p', \mathbf{y}_s')'$, $\mathbf{y}_p \equiv (y_{p1}, y_{p2}, .., y_{pN})'$, $\mathbf{y}_s \equiv (y_{s1}, y_{s2}, .., y_{sN})'$; $\mathbf{x} \equiv (\mathbf{x}_1, \mathbf{x}_2)'$, $\mathbf{x}_1 \equiv (\mathbf{x}_p, \mathbf{0}_s)'$, $\mathbf{x}_2 \equiv (\mathbf{0}_p, \mathbf{x}_s)'$, $\mathbf{x}_p \equiv (\mathbf{x}_{p1}, \mathbf{x}_{p2}, .., \mathbf{x}_{pN})'$, $\mathbf{x}_{p1} \equiv (x_{p11}, x_{p12}, .., x_{p1kp})$, $\mathbf{x}_{p2} \equiv (x_{p21}, x_{p22}, .., x_{p2kp})$, .., $\mathbf{x}_{pN} \equiv (x_{pN1}, x_{pN2}, .., x_{pNkp})$, $\mathbf{x}_s \equiv (\mathbf{x}_{s1}, \mathbf{x}_{s2}, .., \mathbf{x}_{sN})'$, $\mathbf{x}_{s1} \equiv (x_{s11}, x_{s12}, .., x_{s1ks})$, $\mathbf{x}_{s2} \equiv (x_{s21}, x_{s22}, .., x_{s2ks})$, .., $\mathbf{x}_{sN} \equiv (x_{sN1}, x_{sN2}, .., x_{sNks})$; $\boldsymbol{\beta} \equiv (\boldsymbol{\beta}_p', \boldsymbol{\beta}_s')'$, $\boldsymbol{\beta}_p \equiv (\beta_{p1}, \beta_{p2}, .., \beta_{pkp})'$, $\boldsymbol{\beta}_s \equiv (\beta_{s1}, \beta_{s2}, .., \beta_{sks})'$; $\mathbf{0}_p$ and $\mathbf{0}_s$ are null vectors of dimensions $N\times k_s$ and $N\times k_p$, respectively; and the 2N vector $\mathbf{u} \equiv (\mathbf{u}_p', \mathbf{u}_s')'$, $\mathbf{u}_p \equiv (u_{p1}, u_{p2}, .., u_{pN})'$, $\mathbf{u}_s \equiv (u_{s1}, u_{s2}, .., u_{sN})'$, is assumed to have a multivariate normal distribution with mean the 2N null vector and covariance $\boldsymbol{\Sigma}\otimes\mathbf{I}_N$. The parameters of the 2×2 covariance matrix $\boldsymbol{\Sigma}$ are important because they indicate the degree to which errors in the discrete- and continuous-

choice components of the double-hurdle decision are correlated.

The system in (11) is in the form of Zellner's (1971) seemingly-unrelated regressions model (equations (8.72)-(8.78), p. 241). As such, the model plays an important role in another discrete-choice setting that has received considerable attention of late, the multinomial-probit model (see, for examples, Geweke et al. 1994 and 1997; McCulloch et al., and Dorfman). In those situations, a Gibbs-sampling, data-augmentation algorithm is used to simulate from the joint posterior. We demonstrate below that this estimation strategy also proves successful in the double-hurdle context. However, in the double-hurdle case, the two-step decision implies additional restrictions. In this regard, note that the 2Nx1 vector $\mathbf{y}$ contains both observed and latent components. The first N components, $\mathbf{y}_p$, are all latent and some proportion of the second component, $\mathbf{y}_s$, will also be unobserved. In particular, define $\mathbf{c} \equiv \{ i \mid y_{si} = 0 \}$ as the censor set corresponding to the households for which zero supply (market sales) is observed. For each household belonging to the censor set a latent (nonpositive) quantity of sales is implied. These quantities facilitate estimation (a point that is demonstrated to great effect in the seminal paper by Chib) but they are also interesting in a policy context, conveying the notion of a 'distance' at which these non-participating households stand from the market. But restrictions dictated by the double-hurdle representation must be placed on these latent quantities during estimation. There are several variants of these restrictions. The variants arise in correspondence to the investigator's interpretation of the hurdling sequence in the two-step decision-making process. The respective variants can be characterized with reference to the probability masses of the four, respective events: $E_1 \equiv$ the event ($\delta_i = 1$ and $y_{si}^* > 0$), $E_2 \equiv$ the event ($\delta_i = 1$ and $y_{si}^* \leq 0$), $E_3 \equiv$ the event ($\delta_i = 0$ and $y_{si}^* > 0$) and $E_4 \equiv$ the event ($\delta_i = 0$ and $y_{si}^* \leq 0$). These four events are mutually exclusive and exhaustive and motivate four, alternative specifications of the sampling model.

Model One

The first and most natural interpretation, due to its links with standard Tobit and probit formulations, is to consider the joint restrictions $\delta_i = 1$ and $y_{si}^* > 0$ as perfectly correlated. This interpretation, in effect, assigns zero probability to events E_2 and E_3 (prob($\delta_i = 1$ and $y_{si}^* \leq 0$) = prob($\delta_i = 0$ and $y_{si}^* > 0$) = 0). Then, according to the restrictions implied by the probit model (equation (7)) all N elements of $\mathbf{y}_p$ are latent with $\mathbf{y}_{pi}$ truncated to the positive (negative) orthant according to $\delta_i = 1$ ($\delta_i = 0$) and, in addition, the censored components of $\mathbf{y}_s$ are all constrained to be negative.

Model Two

The second model assigns zero mass to event E_2 but not to E_3. Here prob($\delta_i = 1$ and $y_{si}^* \leq 0$) = 0 but prob($\delta_i = 0$ and $y_{si}^* > 0$) $\neq 0$. Accordingly, we model this situation by simulating a draw from the probit model (as above) but now do not constrain the draws for the latent supplies to be negative.

Model Three

The third model assigns zero mass to event E_3 but not to E_2. Here prob($\delta_i = 0$ and $y_{si}^* > 0$) = 0 but prob($\delta_i = 1$ and $y_{si}^* \leq 0$) $\neq 0$. By analogy to the previous case, we simulate this situation by constraining the draws in the Tobit regression to be negative but do not constrain the corresponding draws in the probit regression. Other variants of the basic set-up are possible, but the three presented appear to be the ones that have attracted most attention in the literature (see, for examples, Cragg, Fin and Schmidt, and Jones and the references therein).

A particularly attractive feature of the estimation algorithm that we are about to present is the ease with which these variants of the basic model can be simulated and tested as part of a model selection exercise. Because the three variants imply a set of nested restrictions on the most

general specification, this comparison is performed robustly and intuitively by imposing the implied restrictions and computing at each round of the Gibbs sequence the relative number of violations.

Experiments in the present setting suggest that the first variant (model one) strongly dominated the other two variants (model two and model three) and, hence, reports are made only for the model 1 specification. In addition, further experimentation led to the conclusion that the same covariates were significant in explaining both the participation and the supply decisions.

In this case, seemingly-unrelated regressions model (equation (12)) reverts to the traditional multivariate regression system (Zellner, equation (8.1), p. 224) and estimation is slightly simplified. In terms of equations (11), the modifications implied are $\mathbf{y} \equiv (\mathbf{y}_p, \mathbf{y}_s)$; $\mathbf{x} \equiv \mathbf{x}_p \equiv \mathbf{x}_s$; $\mathbf{x}$ has dimensions N×k; $\boldsymbol{\beta} \equiv (\boldsymbol{\beta}_p, \boldsymbol{\beta}_s)$; and $\mathbf{u} \equiv (\mathbf{u}_p, \mathbf{u}_s)$ is now assumed to have a multivariate normal distribution with mean the N×2 null vector and covariance $\boldsymbol{\Sigma} \otimes \mathbf{I}_N$. Additionally, due to the facts that the vector $\mathbf{y}_p$ is latent, and a subset of the components of $\mathbf{y}_s$ is also latent, we use the symbols $\mathbf{z}_p$ and $\mathbf{z}_s$ to signify the corresponding observed vectors with the latent components included. Hence, $\mathbf{z} \equiv (\mathbf{z}_p, \mathbf{z}_s)$. Finally, in a conventional notation, we note that there are m = 2 equations in the system.

With this notation at hand, under a conventional, non-informative prior $\pi(\boldsymbol{\Sigma}, \boldsymbol{\beta}, \mathbf{z}_p, \mathbf{z}_s) \propto |\boldsymbol{\Sigma}|^{-(m+1)/2}$, the full conditional distributions comprising the joint posterior for the unknown parameters and the latent data, $\pi(\boldsymbol{\Sigma}, \boldsymbol{\beta}, \mathbf{z}_p, \mathbf{z}_s \mid \mathbf{y}, \mathbf{x})$, have the following forms:

$$
\begin{aligned}
\mathbf{z}_p \mid \boldsymbol{\Sigma}, \boldsymbol{\beta}, \mathbf{z}_s &\sim \text{truncated-normal}(E\mathbf{z}_p, V\mathbf{z}_p), \\
\mathbf{z}_s \mid \mathbf{z}_p, \boldsymbol{\Sigma}, \boldsymbol{\beta} &\sim \text{truncated-normal}(E\mathbf{z}_s, V\mathbf{z}_s), \\
\boldsymbol{\beta} \mid \mathbf{z}_s, \mathbf{z}_p, \boldsymbol{\Sigma} &\sim \text{normal}(E\boldsymbol{\beta}, V\boldsymbol{\beta}), \qquad (12) \\
\boldsymbol{\Sigma} \mid \boldsymbol{\beta}, \mathbf{z}_s, \mathbf{z}_p &\sim \text{inverted-Wishart}(\mathbf{W}, v);
\end{aligned}
$$

where $E\mathbf{z}_p \equiv \mathbf{x}\,\boldsymbol{\beta}_p + \Sigma_{ps}\,\Sigma_{ss}^{-1}\,(\mathbf{z}_s - \mathbf{x}\,\boldsymbol{\beta}_s)$, $\mathbf{V}\mathbf{z}_p \equiv \Sigma_{pp} - \Sigma_{ps}\,\Sigma_{ss}^{-1}\Sigma_{sp}$; $E\mathbf{z}_s \equiv \mathbf{x}\,\boldsymbol{\beta}_s + \Sigma_{sp}\,\Sigma_{pp}^{-1}\,(\mathbf{z}_p - \mathbf{x}\,\boldsymbol{\beta}_s)$, $\mathbf{V}\mathbf{z}_s \equiv \Sigma_{ss} - \Sigma_{sp}\,\Sigma_{pp}^{-1}\Sigma_{ps}$; $E\boldsymbol{\beta} \equiv (\mathbf{x}'\mathbf{x})^{-1}\mathbf{z}$, $V\boldsymbol{\beta} \equiv \boldsymbol{\Sigma} \otimes (\mathbf{x}'\mathbf{x})^{-1}$; $\mathbf{W} \equiv (\mathbf{z} - \mathbf{x}\,\boldsymbol{\beta})'(\mathbf{z} - \mathbf{x}\,\boldsymbol{\beta})$, $v \equiv N\text{-}k\text{+}m\text{+}1$; and the 2×2 matrix $\boldsymbol{\Sigma}$ has (scalar) components Σ_{pp}, Σ_{ps}, Σ_{sp} has Σ_{ss}. Consequently, simulations from the joint posterior can be undertaken through the following algorithm:

Step 1: Select starting values $\mathbf{z}_p^{(s)}$, $\mathbf{z}_s^{(s)}$, $\boldsymbol{\beta}^{(s)}$.

Step 2: Draw $\boldsymbol{\Sigma}^{(s)}$ from the inverted-Wishart($\mathbf{W}^{(s)}$, v) distribution, where $\mathbf{W}^{(s)}$ implies conditioning on $\mathbf{z}_p^{(s)}$, $\mathbf{z}_s^{(s)}$, $\boldsymbol{\beta}^{(s)}$ from Step 1.

Step 3: Draw $\boldsymbol{\beta}^{(s+1)}$ from the multivariate-normal($E\boldsymbol{\beta}^{(s+1)}$, $V\boldsymbol{\beta}^{(s+1)}$) distribution, where $E\boldsymbol{\beta}^{(s+1)}$ and $V\boldsymbol{\beta}^{(s+1)}$ denote conditioning on $\mathbf{z}_p^{(s)}$, $\mathbf{z}_s^{(s)}$ and $\boldsymbol{\Sigma}^{(s)}$ from Steps 1 and 2.

Step 4: Draw $\mathbf{z}_p^{(s+1)}$ from the multivariate-normal($E\mathbf{z}_p^{(s+1)}$, $V\mathbf{z}_p^{(s+1)}$) distribution, where $E\mathbf{z}_p^{(s+1)}$ and $V\mathbf{z}_p^{(s+1)}$ denote conditioning on $\mathbf{z}_s^{(s)}$, $\boldsymbol{\Sigma}^{(s)}$ and $\boldsymbol{\beta}^{(s+1)}$ from Steps 1, 2 and 3. (13)

Step 5: Draw $\mathbf{z}_s^{(s+1)}$ from the multivariate-normal($E\mathbf{z}_s^{(s+1)}$, $V\mathbf{z}_s^{(s+1)}$) distribution, where $E\mathbf{z}_s^{(s+1)}$ and $V\mathbf{z}_s^{(s+1)}$ denote conditioning on $\boldsymbol{\Sigma}^{(s)}$, $\boldsymbol{\beta}^{(s+1)}$ and $\mathbf{z}_p^{(s+1)}$, from Steps 2, 3 and 4.

Step 6: Repeat steps 1-5 a large number of times, S^1, until convergence is attained.

Step 7: Repeat steps 1-5 a large number of times, S^2, and collect samples $\{\boldsymbol{\Sigma}^{(s)}$ s =

1, 2, .. S}, {$\boldsymbol{\beta}^{(s)}$ s = 1, 2, .. S}, {$\mathbf{z}_p^{(s)}$ s = 1, 2, .. S} and {$\mathbf{z}_s^{(s)}$ s = 1, 2, .. S}.

Three additional features of the algorithm are necessary for convergence. First, due to identification problems, the draw from the inverted-Wishart in step 2 is normalized on the parameter Σ_{pp} so that the variance implied in the probit equation is one. This is the traditional restriction imposed in univariate settings. Second, only a component of the vector $\mathbf{z}_s$, corresponding to the households in the censor set, are drawn from the conditional normal distribution and the draws for both $\mathbf{z}_p$ and $\mathbf{z}_s$ in steps 4 and 5 are made in accordance with the restrictions implied by the various models. Finally, the samples collected in the last step can be used to draw inferences about any of the unknown quantities of interest. In the results reported below, the algorithm is run for a "burn-in phase" of S^1 = 2,000 observations followed by a "collection phase" of S^2 = 2,000 observations.

In closing this section it seems natural to ask the extent to which the well-known problem of <u>sample selection bias</u> (see, for example, Greene, pp. 926-33.) may be problematic and whether there is need to apply correction procedures, such as those outlined in Heckman (1976, 1979) and applied in Goetz. Sample selection could arise in our context, in considering the effect upon sales of an increase in a level of a covariate, where some individuals who possess the covariate do not sell product. Had those individuals who do not sell been excluded from the sample then a <u>selection bias</u> exists due to the fact that only those respondents selling product are used to form an estimate of the response to the covariate. For example, if the covariate in question is related positively to sales, then only those respondents with a relatively strong response to the covariate will be included, leading to an upwards bias in the corresponding parameter estimate. But, because a latent (negative) sales quantity is simulated for each of the non-selling households and used as the dependent variable in a subsequent estimation step, no such bias exists. In short, the

problem of sample selection bias is conveniently circumvented through the data-augmentation step in the Gibbs-sampled double-hurdle model. In addition, related identification problems arising in frequentist applications, like the need to include non-identical covariate matrices in the probit and Tobit equations (as, for example, in Goetz) are similarly circumvented. Hence the algorithm (13) appears to offer a number of attractive features compared with more traditional methodology.

The Complicating Presence of Fixed Costs

Until now, we have said very little about the issue of fixed costs nor about their impact on the sales decision and an appropriate estimation strategy. With the layout for the traditional model firmly in place, these issues can now be handled with relative ease.

Basic theory of the firm tells us that in the presence of fixed costs there is some minimum quantity below which it is unprofitable for any economic unit – be it a firm or a household – to supply to the market. This implies that the true censoring point in the Tobit regression will not be zero but, rather, some unknown, positive quantity, $\theta > 0$. This quantity is important in the context of household's decisions to enter the market because it circumscribes a minimum-efficient scale of operations measured in terms of a sales quantity. This quantity can be conceptualized in the context of the decision-making model (equations (1)-(5)), the statistical description of the hurdle model (equations (6)-(8)) and the estimation equations ((9)-(13)), as follows.

The presence of fixed costs, may or may not influence the participation decision but, we conjecture, they are likely to influence the quantity decision. This is perhaps most apparent in the observation that at household level, trade is commonly discontinuous in time, with individual

households selling some periods and not selling in others. Plainly, such a household is a market participant, although it opts for zero sales volume in some periods. Put differently, the good it sells is tradable from its perspective even if it is not always traded. This is conceptually akin to households adopting a new technology, then discontinuing its use at some future date(s) when it proves unprofitable (Cameron, 1999).

Hence, in the sales optimization problem (equation (2)), the constraint $y_{si} \geq 0$ is replaced by the condition $y_{si} \geq \theta$. This modification leads, in turn, to the notion that the observed data on sales, y_{si}^{**}, are actually the maximum of the latent sales quantity, y_{si}^{*}, as specified in (8), and the unknown quantity $\theta > 0$. Consequently, θ is now the censoring point in the Tobit regression. As such θ becomes an additional parameter in the model and must be estimated, along with the system parameters $\boldsymbol{\Sigma}$ and $\boldsymbol{\beta}$, the latent $\mathbf{z}_p$ and the latent components of $\mathbf{z}_s$.

Devising the fully conditional distribution for θ would appear to be a difficult task were it not for its development in an apparently unrelated work by Albert and Chib. In that work the authors consider a problem that has an almost identical structure to the model in (9)-(13) and a little algebra (available upon request) indicates that the full conditional distribution for the unknown θ is uniform on the interval $[\max\{z_{si}, i \in \mathbf{c}\}, \min\{z_{si}, i \notin \mathbf{c}\}]$. The bounds on the interval of this uniform distribution are quite intuitive. The left bound is simply the greatest value of latent sales from the non-participating household and the right bound is the minimum quantity of sales observed by the participating households. Intuitively, because all of the households are the same (except for their endowments of the market-precipitating covariates), the unknown censoring point (the minimum efficient scale of operations) should lie between these values.

The censoring value, θ, can be estimated with a few basic modifications to the algorithm in (13). Essentially, three modifications are required. The first modification is to select, in Step 1,

a starting value $\theta^{(s)}$. We select the minimum sales quantity observed, i.e., the upper boundary of the feasible range for θ. Second, the draws in steps 2-4 are now conditional on the chosen value $\theta^{(s)}$. Third, below step 5, insert the additional step: Step 5a: Draw $\theta^{(s+1)}$ from the uniform distribution with bounds $[\max\{z_{si}^{(s+1)}, i \in \mathbf{c}\}, \min\{z_{si}, i \notin \mathbf{c}\}]$, where $\max\{z_{si}^{(s+1)}, i \in \mathbf{c}\}$ implies conditioning on the maximum component of $\mathbf{z}_s^{(s+1)}$ in step 5 and where $\min\{z_{si}, i \notin \mathbf{c}\}$ denotes the minimum sales quantity observed in the data.

The Application

We apply this method to data on milk marketing by Ethiopian dairy farmers in two sites close to the capital city, Addis Ababa. The sites were identified in 1997 as potentially useful for examining the impacts of transactions costs on participation in peri-urban milk marketing. We focus attentions on a subset of the farmers that have crossbreed cattle and make fluid milk sales to two milk cooperatives. There are two reasons.

First, private milk sales in Ethiopia are often impeded by high fixed transactions costs. Among the more prominent of these are costs associated with equipment for manufacturing easily transportable dairy products (butter, cheese and yogurt) from fluid milk; pecuniary transport costs, such as the purchase of a cart or a donkey for haulage (a fixed cost); non-pecuniary transport costs, such as the reallocation of household labor toward hand transportation of products (which has both fixed and variable interpretations); and the inevitable time and risks associated with searching, negotiating and enforcing a sale, irrespective of the volume transacted (again, with both fixed and variable components).[4] In this context, cooperative sales organizations purchasing fluid milk and manufacturing derivative products (butter, cheese and yogurt) are thought to be important catalysts stimulating participation into markets currently

constrained by considerable thinness. Hence, cooperative selling is thought to be a significant transactions cost reducing innovation.

The orientation towards crossbreed animals at the study sites is motivated by the fact that crossbreed animals generate potential production increases (over indigenous breeds) of one-hundred percent (milk-fat per metabolic weight of animal) have been recorded in station trials and these results are replicated to various degrees in field situations (Kiwuwa et al.). Production gains of this magnitude are an obvious stimulus to marketable surplus in the household and thereby to overcoming the fixed costs to market participation. .

The Data

Early in the 1997 production year a sample of 68 households was selected based on their stratification of cross-bred cow ownership and their physical location relative to two milk cooperatives. Three visits were made to each household during the year, and at each visit weekly sales of fluid milk to the milk cooperatives were obtained from co-op records. Demographic, nutritional and socioeconomic characteristics of the households were recorded.

The analysis focuses on the determinants of weekly sales of fluid milk at each of the 3 visits—a sample size of 204 observations. Preliminary analysis with the data suggests that seven covariates are particularly influential in explaining milk production and marketing from these households. Hence, estimation is conducted on a parsimonious choice of these seven effects, namely, (1) numbers of indigenous milking cows, (2) numbers of crossbred milking cows, (3) minutes, return time, to transport bucketed fluid milk to the milk cooperative, (4) years of formal schooling by household members, (5) the number of total visits by an extension agent discussing production and marketing practices, (6) a site-specific dummy variable corresponding to the 'Ilu-

Kura' sample site (about 60 mile south-west of Addis Ababa) and (7) and a site-specific dummy variable corresponding to the Mirti sample site about (about 140 miles north-east of the capital city).

Results

Results of the Gibbs-sampling, data-augmentation algorithm applied to these 204 observations are presented in table 1. The first column presents definitions and the remaining columns present the posterior means of the parameters in the multivariate probit-Tobit systems under traditional and non-zero censoring, respectively. Auxiliary statistics are reported in the lower portion of the table. The mnemonics in the first column refer, respectively, to θ ('Censor value'); minutes return time to transport bucketed-fluid milk to the milk cooperative ('Distance'); years of formal schooling by the household head ('Education'); the number of crossbreed cows being milked at the survey date ('Crossbred'); the number of indigenous-breed cows milked at the survey date ('Local'); the total number of visits in the twelve months prior to the survey date by an extension agent discussing production and marketing practices ('Extension'); a binary variable corresponding to the Ilu-Kura survey site (equals 1 if respondent is from Ilu-Kura and equals 0 otherwise); and a binary variable corresponding to the Mirti survey site (equals 1 if respondent is from the Mirti survey site and equals 0 otherwise). Numbers in parentheses below the parameter estimates are lower and upper bounds for the 95% highest-posterior density regions.

Considering, first, the traditional formulation with zero censoring in the Tobit regression, each of the parameter estimates are significant at the 5% significance level. (None of the 95% highest posterior density regions contains zero.) The signs of the posterior means all have the expected impact. Participation is promoted by education, cow ownership and the level of

extension services, but is mitigated by distance to market. Sales are also increased by the intellectual capital stock (education and extension visitation) and the animal stock (local and crossbreed animals) but is reduced by distance to market.

An important result in the context of two-step decision-making is the possibility that errors are correlated. Previous work (most notably, Key et al., 2000) assumes independence The estimated covariance parameters suggest strongly that the participation and the sales decisions are highly correlated. Other features of the traditional model are the relatively large degree of variability in the sales equation error variance (posterior mean estimate of 1047.40 liters of milk per household per week); outstanding predictive performance among the non-participating 'households' (179 of the 204 total observations); but less satisfactory fit in the participating sample (25 observations in total). Because 85% of the sample observations are censored, the poor prediction in the participating sample is somewhat expected due to small sub-sample size. But the large error variance in the sales equation suggests that a number of other omitted factors may be responsible for weekly sales variability.

Before turning to examine differences between the first formulation and the formulation that does not restrict the censoring value to be zero, a word about the covariate 'Distance' seems in order. Recall that the purpose of relaxing the zero-restriction on the censoring value is to attempt to capture the importance of fixed costs and their affect on the minimum efficient supply quantity. But there may be grounds for suspecting double counting with reference to some of the covariates. For example, it is certainly true that there is a fixed cost related to distance (e.g., the cost of transporting the individual, not the milk, to market). In this case, it may be argued that the covariate 'Distance' is capturing both proportional and fixed transactions costs. Put differently, θ understates the fixed cost of market participation because of the distance-related

fixed cost. Identification of proportional costs and separating them out from their corresponding contributions to fixed costs is problematic. This point is made by Key et al. (2000) who attempt to distinguish between the two components empirically. Whether it is possible to perform a similar decomposition using the current estimation strategy remains an interesting issue for possible extensions of the current effort.

Turning to the second, non-zero censoring formulation, the most interesting comparisons are three. First, the posterior mean estimate of the censor value suggests that the minimum efficient scale of operations for the household is a resource base consistent with delivery of 5.26 liters of milk per week for a household located at the market delivery point. Note, also that this estimate is measured at a considerable degree of precision (with 95% highest-posterior-density bounds of 3.75 and 5.97, respectively). Hence, one important conclusion emerging from the exercise is that a significant bias could result from restricting the censor value to zero. Evidence of this potential bias is encountered in comparisons of the covariate estimates between the two models, which is the second important feature of comparison. In both the participation and supply equations, each of the continuous covariate (i.e., other than the site dummies) coefficient estimates has the same sign across the two models. But the magnitudes of the means estimates in the two equations exhibit an interesting pattern. In the participation equation each of the estimates in the random-censor model is <u>greater</u> (in absolute value) than the corresponding estimate in the traditional model and in the supply equation each of the estimates is <u>smaller</u> (in absolute value) than the corresponding estimate in the traditional, zero censoring model. Further, in both the participation and supply equations, the site-specific dummy coefficients are <u>greater</u> under random censoring than in the traditional formulation. Hence, having concluded that the true point of censoring is not zero, these results suggest that ignoring the importance of potential

fixed costs in the supply decision has three impacts on the double-hurdle estimates. First, it biases downwards both estimates of the impact of the covariates on participation and the impact of 'other factors' as depicted by the constant terms. Second, it biases upwards estimates of the impacts of the covariates on supply but biases downwards estimates of the impacts of 'other factors' on supply as evidenced in reports of the coefficients of the site-specific dummies. In short, the net impacts of ignoring fixed costs are a lower prediction about likelihood of participation and a higher prediction about supply potency. Further evidence that the second formulation is a better description of the data is evidenced by the reports of dramatically lower error variances and the improved predictive statistics in the lower part of the table. This is not just an idle methodological point. The practical implication is that increasing market participation is central to expanded aggregate supply, so traditional price policy prescriptions that rest upon the assumption of ubiquitous market participation may not be the most effective means of increasing market supply.

Conclusions

Collectively, these results demonstrate the importance of allowing for non-negligible fixed costs in market participation (adoption) studies. When these costs are ignored but are non-negligible, a significant bias in participation and supply estimation appears to exist. In the context of examining this issue, we have presented a Bayesian approach to estimation of the double-hurdle model, which is popular because it allows for a potentially diverse set of factors to influence participation and supply decisions. Our analysis, however, suggests that in these data on highland Ethiopian milk producers, the same factors influence both participation and supply and that the intellectual capital stock (education and extension visitation) is a vital complement to the

physical capital stock (both local and crossbred animals) in effecting market entry among formerly subsistence households. With the intent of expanding the density of milk-market participation in peri-urban settings, extension agents and policy makers should target these inputs with a view to expanding household capacities above a minimum of 5.26 liters of milk per household per week.

Footnotes

[1] The entire procedure took approximately ten minutes of real time on a DELL™ Optiplex G1 machine running a Pentium™ II processor at 330 megahertz with commands executed in MATLAB™ version 5.1.0.421. All computer code is available upon request.

[2] See Albert and Chib, equation (18), for a similar development in the context of the ordered probit specification.

[3] Experiments with the non-zero censoring algorithm suggest that these additional steps consumed negligible additional time.

[4] Proportional, variable costs plainly matter as well, but these have no effect on the censoring point, they merely adjust the net price received per unit sold, like an ad valorem tax. Fixed costs, by contrast, create the non-zero censoring point of interest here.

[5] These are the Bayesian equivalents to the traditional confidence intervals encountered in sampling theory.

[6] Key et al. (2000) make similar observations in their work on Mexican maize markets. They find, for example, that 60 percent of the increase in marketed maize supply in response to maize price increases is due to increased market participation, only 40 percent due to expanded sales by existing market participants.

References

Albert, J. and S. Chib. (1993). 'Bayesian Analysis of Binary and Polychotomous Data.' Journal of the American Statistical Association Vol. 88, pp. 669-79.

Brokken, R. F. and S. Seyoum (eds.). (1990) Dairy Marketing in Sub-Saharan Africa. Proceedings of a Symposium at ILCA, Addis Ababa, Ethiopia.

Cameron, L (1999). 'The Importance of Learning in the Adoption of High-Yielding Variety Seeds.' American Journal of Agricultural Economics Vol. 81, pp. 83-94.

Cogan, J.F. (1981). 'Fixed Costs and Labor Supply.' Econometrica Vol. 49, pp. 945-963.

Cragg, J. (1971). 'Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods.' Econometrica Vol. 39, pp. 829-44.

Chib, S. (1992). 'Bayes Inference in the Tobit Censored Regression Model.' Journal of Econometrics Vol. 52, pp. 79-99.

___. (1995). 'Marginal Likelihood From the Gibbs Output.' Journal of the American Statistical Association Vol. 90, pp. 1313-21.

Dorfman, J. H. (1996). 'Modeling Multiple Adoption Decisions in a Joint Framework.' Am. J. Agr. Econ. Vol. 78, pp. 547-57.

Fin, T. and P. Schmidt. (1984). 'A Test of the Tobit Specification Against An Alternative Suggested by Cragg.' Review of Economics and Statistics Vol. 66, pp. 174-77.

Gelfand, A. and A. Smith. (1990). 'Sampling-Based Approaches to Calculating Marginal Densities.' Journal of the American Statistical Association Vol. 85, pp. 972-985.

Geweke, J. (1992). "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds.), Proceedings of the Fourth Valencia International Conference on Bayesian Statistics.

New York: Oxford University Press, pp. 169-93.

Geweke, J., M. Keane and D. Runkle. (1994). 'Alternative Computational Approaches to Inference in the Multinomial Probit Model.' Review of Economics and Statistics Vol. 76, pp. 609-32.

Geweke, J. M. Keane and D. Runkle. (1997). 'Statistical Inference in the Multinomial, Multiperiod Probit Model.' Journal of Econometrics Vol. 80, pp.125-66.

Gelman, A., J. B. Carlin, H. S. Stern and D. B. Rubin. (1995). Bayesian Data Analysis. London: Chapman and Hall.

Goetz, S. J. (1992). 'A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa.' American Journal of Agricultural Economics Vol. 74, pp. 444-452.

Greene, W. H. (2000). Econometric Analysis. New Jersey: Prentice Hall, 4th ed.

Heckman, J. J. (1976). 'The Common Structure Of Statistical Models Of Truncation, Sample Selection and Limited Dependent Variables And A Simple Estimator For Such Models.' Annals of Economic and Social Measurement Vol. 5, pp. 475-92.

_____. (1979). 'Sample Selection Bias As A Specification Error.' Econometrica Vol. 47, pp. 53-161.

Jones, A. (1989). 'A Double-Hurdle Model of Cigarette Consumption.' Journal of Applied Econometrics Vol. 4, pp. 23-39.

Key, N., E. Sadoulet and A. de Janvry. (2000). 'Transactions Costs and Agricultural Household upply Response.' American Journal of Agricultural Economics Vol. 82, pp. 245-59.

Kiwuwa, G. H., J. C. M. Trail, M. Y. Kurtu, G. Worku, F. M. Anserson and J. Durkin. (1983). Crossbred Dairy Cattle Productivity in Arsi Region, Ethiopia. International Livestock Center for Africa (ILCA) Research Report No. 11. Addis Ababa: ILCA.

McCulloch. R., P. Rossi and N. Polson. (2000) 'Bayesian Analysis of the Multinomial Probit Model With Fully Identified Parameters.' Journal of Econometrics Vol. 99, pp. 173-93.

Nelson, F. (1977). 'Censored Regression Models with Unobserved, Stochastic Censoring Thresholds.' Journal of Econometrics Vol. 6, pp. 309-327.

Nicholson, C. F. (1997). The Impact of Milk Groups in the Shewa and Arsi Regions of Ethiopia: Project Description, Survey Methodology, and Collection Procedures. Mimeograph, Livestock Policy Analysis Project, International Livestock Research Institute, Addis Ababa.

Puhani, P. A. (2000). "The Heckman Correction for Sample Selection and Its Critique." Journal of Economic Surveys Vol. 14, pp. 53-69.

Staal, S., C. Delgado and C. F. Nicholson. (1977). 'Small-Holder Dairying Under Transactions Costs in East Africa.' World Development Vol. 25, pp. 779-94.

Stiglitz, J. E. 'Markets, Market Failures and Development.' (1989). American Economic Review Vol. 79, pp. 197-203.

Zellner, A. (1971). An Introduction To Bayesian Inference In Econometrics. New York: Wiley and Sons.

Table 1. Double-Hurdle Equation Estimates.

	Model			
	Zero Censoring		Non-Zero Censoring	
	Participation	Sales	Participation	Sales
Censor Value				5.26
				(3.75, 5.97)
Distance	-0.02	-0.46	-0.02	-0.31
	(-0.03, -0.01)	(-0.76, -0.17)	(-0.05, -0.01)	(-0.51, -0.12)
Education	0.17	4.21	0.22	2.59
	(0.08, 0.26)	(1.60, 7.35)	(0.08, 0.40)	(0.94, 4.53)
Crossbred	0.80	28.61	1.02	21.68
	(0.48, 1.20)	(20.45, 39.00)	(0.58, 1.64)	(16.18, 29.00)
Local	0.29	12.75	0.40	10.00
	(0.04, 0.55)	(5.59, 19.77)	(0.07, 0.80)	(5.64, 14.81)
Extension	0.16	4.39	0.20	2.87
	(0.06, 0.27)	(1.58, 7.37)	(0.09, 0.35)	(1.24, 4.49)
Ilu-Kura	-1.68	-64.82	3.12	-38.12
	(-2.53, -0.87)	(-98.00, -38.51)	(1.65, 4.31)	(-58.71,-22.51)
Mirti	-3.08	-102.57	1.33	-61.95
	(-3.97, -2.18)	(-150.09,-67.92)	(-0.98, 2.70)	(-91.09, -41.36)
	Covariance			
Participation	1.00	9.42	1.00	6.29
		(4.60, 14.99)		(3.46, 9.64)
Sales	(symmetric)	1047.40	(symmetric)	345.08
		(475.38,2045.15)		(154.72, 686.32)
	Auxiliary Statistics			
	Non-Participants			
R^2	0.97	0.91	0.98	0.87
Pos. pred.	3.00	4.00	2.00	8.00
Neg. pred.	176.00	175.00	177.00	171.00
	Participants			
R^2	0.92	0.33	0.84	0.39
Pos. pred.	11.00	11.00	25.00	13.00
Neg. pred.	14.00	14.00	0	12.00

Note: 95% highest posterior density values are reported in parentheses.